

**SOCIAL MEDIA GOVERNANCE:
PLATFORMS IN CONFLICT —
THE STRATEGIES OF OPERATORS AND
THE TACTICS OF USERS.**

Andrew Quodling

B Creative Industries (Honours I)

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Digital Media Research Centre

Creative Industries Faculty

Queensland University of Technology

2019

Keywords

Platform Governance

Social Media

Platform Politics

User Tactics

Content Moderation

Online Harassment

Facebook

Twitter

Google+

Abstract

The growth in availability, affordability and accessibility of internet services in recent years has seen a dramatic increase in the use of social media platforms like Facebook and Twitter. The political conflict and activism both on and against social media platforms have become a more frequent occurrence as their user bases grow and diversify. Conflicts between users and platform operators have involved their attitudes and policies towards issues as diverse as censorship, privacy and surveillance, images depicting breastfeeding, images depicting nudity, pornography, hateful speech and threats of physical harm — with many conflicts garnering coverage in mainstream media.

This project investigates conflicts in which users of social media platforms have dramatically different political views and expectations than the operators and designers of their platforms by analysing the different ways in which users have negotiated differences of power and political disputes on social media platforms in several case studies. It has highlighted regular areas of conflict or tension, the strategies that platforms use and the tactics that users deploy to negotiate these conflicts. By interrogating the power structures and political conflicts of social media platforms, this research has discovered how platform operators attempt to govern users through policy and enforcement strategies, and how users interact with platforms during conflict and by identifying themes in user actions during conflicts. Through analysing the interplay of these strategies and tactics, this thesis also illustrates some of the challenges of governance in globalised, digital spaces like internet platforms.

Table of Contents

Keywords	i
Abstract	ii
Table of Contents	iii
Statement of Original Authorship	v
Acknowledgements	vi
Content Notice	viii
Chapter 1: Introduction	1
Context	2
Significance and Scope	4
Thesis Outline	4
Chapter 2: Literature Review	7
Understanding Unequal Power Conflicts	7
Contextualising Platforms and their Politics	11
Platforms and Strategies	32
Users and Tactics	61
Platform Governance	84
The Role of this Research Project	93
Chapter 3: Research Design	95
Chapter 4: Case Study One: Facebook and Breastfeeding Photographs	119
Platform–Political Tensions	119
Platform History and the Conflict	133
Governance Context	145
Perspectives of Conflict and Governance	160
Evaluation	169
Chapter 5: Case Study Two: Google+ and “Real Names”	173
Platform–Political Tensions	173
Platform History and the Conflict	188
Governance Context	194
Perspectives of Conflict and Governance	198
Evaluation	202
Chapter 6: Case Study Three: Twitter, Abuse and Harassment	204
Platform–Political Tensions	204
Platform History and the Conflict	214
Governance Context	226

Perspectives of Conflict and Governance.....	240
Evaluation.....	250
Chapter 7: Discussion	263
Strategies	263
Tactics	267
The Interplay of Strategies and Tactics	275
Chapter 8: Conclusions.....	290
What happened next? Addenda to the Case Studies	291
The Limitations of this Research, and the Timeliness of this Investigation.....	305
Shocks & Exceptions	306
Bibliography	313

Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature: QUT Verified Signature

Date: March, 2019

Acknowledgements

To my supervisors, Axel Bruns, Stephen Harrington, and Nicolas Suzor: Thank you. Along this academic journey, you've each been invaluable sources of support and guidance, and have been essential to my development as a researcher. This thesis would not have been possible without your generous insights and your enduring support.

To Emma Potter-Hay: Thank you. I couldn't possibly have had a better colleague to share offices with. I have valued our far-ranging discussions, our arguments and our moments of shared vulnerability in the great, challenging and often daunting PhD journey. It has been an honour to share mine with you.

To Scarlett Hawkins: Thank you. Your care, candour, support, and friendship has been invaluable to me on this academic journey.

To my colleagues in the Digital Media Research Centre and Creative Industries Faculty: Thank you, to all of you. You've each enriched this PhD journey with your interest, your support and your enthusiasm. It's been a delight to develop as a researcher and an academic in the company of this academic cohort.

To the participants and organisers of the Oxford Internet Institute's 2015 Summer Doctoral Programme: Thank you. I had the privilege to participate in this programme at a particularly introspective point in my Doctoral research, and a formative point my development as a researcher. Your care, candour, excitement and encouragement are something I hope to remember fondly for many years to come.

To my family, Barbara, Peter and David:

Thank you, most of all.

Content Notice

This thesis engages with issues relating to online governance and the behaviour of platforms and users in digital public spaces. As a result of this research, this thesis includes both textual descriptions and graphic depictions of material that readers may find acutely distressing.

The following thesis shows detailed threats of abuse and harm, incitements to suicide, some graphic representations that accompanied these threats. In addition to this, some quotations and examples given in this thesis include white supremacist rhetoric, racist slurs, sexist slurs, and a host of other plainly offensive language that could be *charitably* described as hateful and abusive.

These examples are reproduced in this thesis not to be gratuitous or offensive, but to give context and perspective to the discussion of the policies cultures, and technological structures that support the creation and publication of these texts. A term like ‘hate speech’ can be more easily dismissed in the abstract, so to more clearly understand the dynamics of its effect on people and on discourse, we must be prepared to critically engage with these texts and challenge the systems that support their propagation.

Chapter 1: Introduction

When is it appropriate to use the word ‘dyke’? As an epithet long since reclaimed by those it had been hurled against, the word is complex and laden with meaning. After efforts to re-appropriate (Galinsky et al., 2013) the word, it’s perhaps most appropriate to defer to the judgement of the people who were harmed by the term, and the people who have re-appropriated the term as an identity. In June 2017, users of the Facebook social media platform reported a significant uptick in Facebook’s policing and punishment of lesbian users, correlating this in particular with their usage of the word ‘dyke’ (Trigger, 2017). Similarly, Facebook pages associated with the ‘Dykes on Bikes’ motorcycle groups have been removed from the platform. For the users who were affected by this policing, these platform interventions came as a particular source of frustration. Facebook was not only policing their language, telling them that they could not use a term that they have claimed for their own, but it was also impeding their ability to socialise and seek support from other people, other users of the platform who were likely to have similar concerns. These moderation interventions from Facebook highlight the platform’s influence in everyday life for some of its users and point to a punishment and severity that seems unjust for users affected by it.

After some public criticism, Facebook responded via their Vice President of Public Policy, Richard Allan. Allan urged that Facebook moderators had a range of guidelines that are intended to consider content, context, country, and intent — noting that the word “dyke” could be used in a self-referential way and not fall afoul of any of Facebook’s rules — but the use of “dykes” as a collective group would be more likely seen by moderators as an epithet. Allan lamented that machine-learning technologies are not significantly enough advanced to be meaningfully used in content

moderation (Allan, 2017), although spokespeople for Facebook have been at pains to assure users that content is reviewed by moderation teams, and is not automatically dealt with by secretive algorithms. Allan conceded that mistakes had been made by the platform, offering that Facebook would continue to refine their processes.

This issue presents a microcosm for this research thesis more broadly — the tension, disparate power and political conflict between platform operators and users.

Firstly, it highlights a lack of Transparency in the processes of these platforms — showing that moderation decisions are often made without consultation or negotiation with stakeholders. Secondly, it shows that platforms are governed spaces, and the ways in which these spaces are governed can come into conflict with user expectations.

Facebook’s policing here demonstrates the capabilities platforms have to police acts of user expression and speech acts, and user expressions of identity. Explored further in this thesis is the influence that this policing (or the absence of this policing) can have on user feelings of Safety on social platforms. These issues each illustrate how these platforms can be profoundly important to the users who consider them a part of their everyday life.

Context

This thesis and research project are by necessity an interdisciplinary investigation. Issues relating to this research have risen across a range of disciplines, including communication studies, as well as studies of the rule of law and governance. Given this, I have engaged with the work of theorists across a broad range of fields, as well as engaging with the works of thinkers outside of formal academic research disciplines. Key amongst these disciplines is the study of the politics of ‘platforms’

(Gillespie, 2010), and an engagement with these complex digital spaces as places of both labour and leisure, of civic participation, of interpersonal communication, and of a range of other uses — as determined in an informal negotiation of power between users and platform operators.

This study uses conflicts between users and platform operators on social media platforms like Facebook, Twitter and Google+ to explore issues of governance and power. It does all of this by chronicling a handful of ‘platformed conflicts’ in close detail, using case study, close reading and desk-based research practices. Through critical interrogation of the behaviours of platform operators and users in these conflicts, we can discern key strategies and tactics involved in the unbalanced power conflict of these digital spaces. Moreover, investigating how the interplay of these strategies and tactics might influence platform governance may provide the foundations for a theoretical framework for better understanding the power relationship between platform operators and platform users.

Through this thesis, I present a study that has charted the political interplay of users and platform operators in the fraught, complicated social and political conflicts that affect users in day-to-day life. This work has revealed the existing frameworks of governance to be inadequate and provides an insight into the political efforts of users seeking reform on the platforms they use. The efforts users make not simply in participating in online platform spaces, but in acting as voices for reform, and sometimes even impromptu custodians of the spaces are indicative of the growing need to more fully include users in the machinations of the governance that so regularly affects their lives.

Significance and Scope

Given the growth of not only social media platforms, but other intermediary platforms like ‘sharing economy’ apps and massively multiplayer video games, it is critical that researchers form a strong understanding of these platforms, not only on a technical or theoretical level, but an understanding that can also engage with and chronicle the lived experiences of users on platforms — especially when users come into conflict with platform operators, and seek to ameliorate that conflict. Through the usage of case study methodology, this research will focus in particular on a select few cases of platformed conflict, with the intention of presenting an analysis based on thorough and detailed inspection that focuses on the opportunities for learning afforded by the cases events analysed (Stake, 1995 4).

Thesis Outline

This thesis loosely adopts the ‘compromise model’ suggested by Patrick Dunleavy (2003) as a way of grounding the research work in appropriate theoretical context and positioning the case studies in a manner that allows for critical investigation and analysis. This thesis is neither exclusively fine detail nor high theory — it compromises and presents a detailed analysis of the core issues, and a theory-focused exploration guided by both the literature and the case studies.

This brief, introductory chapter explains the underlying imperative for this research — why it is necessary, and the fields in which the research is relevant. This chapter also provides a short guide as to the rest of the thesis.

In Chapter 2, I present a literature review that grounds this thesis across disciplines and establishes key issues that arise in the investigation of power and governance on social media platforms. With this, I demonstrate how the growing number of users on social platforms represents an impetus for more thorough

engagement with the political economies of these platforms, and the political conflicts users are involved with due to their participation on these platforms. I draw on literature from communications research, governance research, as well as the works of writers, popular thinkers and people involved in the development or criticism of these platforms.

In the following Chapter 3, I present my research design — elaborating on the decision to use Case Study methods for this research, and the decision process involved in selecting cases. Beyond this, I detail the desk-based research that relates for the lion's share of the research presented in this thesis, and the interview methods that were used in discussing platformed conflicts with a handful of users who were so gracious as to talk to me about their involvement in social media platforms, and their efforts to change platform policies they saw as unjust.

Chapter 4 marks the beginning of the case studies in this thesis, an investigation of Facebook and its policies as they relate to breastfeeding photographs, and photographs of mastectomy scars. This is perhaps one of the oldest and most well-known conflicts or controversies that the platform has faced, and one it continues to face to this day (albeit with greater prowess and media savvy). In this chapter, I discuss the context of the platform's governance decisions, and the public-facing development of this policy over time.

Chapter 5 presents the second case study, an examination of Google+ and the 'nymwars', a comparatively short-lived conflict, characterised by the platform seeming to give in after several years of inaction on the issue. This case, like the one before it, grounds this issue in the significant social and political context that is often underlying discussions that focus on issues of identity, or issues of anonymity and pseudonymity on online platforms.

Chapter 6 is the third and final case study, an investigation into Twitter, and the safety issues that have plagued its users. This case in particular, looks at two prominent periods of protracted abuse and harassment on the platform — abuse and harassment which was largely directed towards women, people of colour, and people who identify as Lesbian, Gay, Bisexual, Transgender, Intersex or Queer (LGBTIQ). These conflicts, a harassment campaign against prominent women in the United Kingdom’s media sphere after a successful campaign to see Jane Austen featured on English currency, and a harassment campaign that coalesced under the name ‘Gamergate’, targeting women, people of colour and LGBTIQ people in the gaming industry and press, are instructive, as they reveal the tensions between safety and free expression, and how the platform has struggled the balance these in a manner amenable to some of its users.

The analysis chapter, Chapter 7, serves as an opportunity to engage more broadly with themes that emerged in the cases, and the findings this research presented. To that end, Chapter 7 discusses the themes underlying the strategies of platform operators and tactics of dissident users, and the opportunities these present for reform, or further research.

Finally, the concluding chapter, Chapter 8, revisits the conflicts of each case, asking “What happened?” — these case addenda provide more context to each case, as a follow-up outside the formal bounds of the case investigations. These illustrate not only the pace of change on these platforms, but the urgent need for continuation of this sort of research, interrogating power and conflict in these spaces.

Chapter 2: Literature Review

Understanding Unequal Power Conflicts

Foucault, de Certeau and rationalising power

Michel Foucault is perhaps one of the most influential critics of societal manifestations of power. In *Discipline and Punish* (1995) he utilised the metaphor that had previously been considered as part of an architectural concept for the design of observation-centric institutions, such as schools, hospitals and prisons — a ‘panopticon’ — in critical analysis of societal tendencies for observation, normalisation and discipline. In doing this, de Certeau (2002 46) explains that he distinguishes two heterogeneous systems: the elaboration of a body of doctrine; and the imposition of political technology to govern the body. Both of these systems deploy institutionalised power against an ‘Other’ — typically a citizen or a user — by regulating the relationship within which the Other can engage with the institution and their systems. By deploying outright power or regulating relationships against the comparative weakness of the Other, the institutions that wield this power seek to establish, normalise and thus entrench their powers through strategies — structures of governance.

Strategies are about the calculation of power, which becomes possible when a subject of power (a proprietor, enterprise, government or institution) can be conceptually isolated from the environment they govern. The calculation serves to

become the basic rationalisation of relationships with that environment, and those are subject to it — the governed subjects (Ibid. xix). As a structure of power, strategies are an attempt at the formalisation and regulation of practices of everyday life by the calculation of a governing power. To provide an example, a business proprietor might develop strategies to regulate the participation of employees with their workplace environment. Strategies like surveillance can serve to regulate the practices of employees by establishing expectations of work habits as well as of how those habits are perceived and then surveilling the workers to ensure participation as regulated. Indeed, even the awareness of surveillance acts as a form of discipline, as the surveilled adjust their behaviours to in response to it.

An art of the weak

Tactics, conversely, are formulations of power that are constructed through the means available to the less powerful. These are often created in reaction or in opposition to strategic deployments of power. Both strategies and tactics are manifestations of power, the key difference is that understanding ‘tactics’ in opposition to strategies serves to identify and understand the wielder’s logic and the rationalisation of their action from a different perspective — typically that of the weaker participant (Colebrook, 2001 546-7). Tactics serve to calculate power, but from a perspective other than the traditionally powerful — as tactical actions do not derive their power from spatial or institutional privileges and instead are powered by manipulation of everyday practices to create and secure different circumstances (de Certeau, 2002 xix). We can surmise as de Certeau (Ibid. 37) has, that “in short, a tactic is an art of the weak”, involving trickery and opportunities afforded by particular structures of power. Effectively used tactics allow the weak to “manipulate events to

create opportunities” and continually turn forces that are alien to them to suit their own ends (Ibid. xix). The elements of their subversive discourse combine to “produce a flash shedding a different light” in their conflict. The clearest example is of what de Certeau refers to as *la perruque* (which translates literally to ‘the wig’) (Ibid. 25), a tactic whereby workers exerted their power by undermining the strategies of surveillance that employers use to attempt to control the productivity of workers, a method where workers steal their own time and effort back from employers by working deliberately slowly. In the factory-based context in which de Certeau explained ‘tactics’, workers could game or co-opt the means of production through their actions. In essence, this demonstrates the subversive power that could be wielded by individual or collectively organised workers.

Of course, the tactics and implied power of a company’s workers are now well-acknowledged. The notion was the central conceit of the cult-hit 90’s film *Office Space* where the protagonist, Peter, proudly confesses to his employer’s consultants, Bob and Bob, that “I just stare at my desk — but it looks like I’m working” and that “It’s not that I’m lazy. It’s that I just don’t care!” (Livingston, McGinley, & Wilson, 1999). Though *Office Space* is obviously a work of fiction, its success is emblematic of the way in which large and powerful organisations are vulnerable to subversion.

Platforming de Certeau

In *The Practice of Everyday Life*’s most memorable chapter ‘Walking in the City’, de Certeau contrasted the perspectives and powers of strategies and tactics in the context of a modern-day city. From the viewpoint of the 110th floor of the World Trade Center building, de Certeau posits that the cityscape offers a voyeuristic view

of the city shaped by authors and observed by spectators (2002 92-93), acknowledging the power with which institutions of business and government have shaped and strategised the city and the complexity of the city as a system. This is contrasted with the vantage of the city's practitioners, the 'walkers' on the streets below, who engage the city as an urban 'text' and shape the discourse of the city in their actions. The city of de Certeau's (Ibid. 94-95) discussion is an utopian, urbanistic discourse that operates under three key parameters: the production of its own space; the existence of a linguistic system of "indeterminable and stubborn resistances offered by tradition"; and finally, the creation of a universal, anonymising subject — the city itself — that will gradually become an attributable 'actor' in its own right — so that the 'actor' becomes intellectual shorthand for the actions of the city's inhabitants. The conceptual 'city' of de Certeau's description is a landmark for the technological power exerted in socio-economic and political strategies, but the streets of said city are where the individuals thrive, and the strategies of governance are most ripe for subversion through the use of tactics.

In the information age, the theoretical conceit of de Certeau's conceptual city can be applied to structures that exist outside of the concrete and steel of urban environments; structures that primarily exist in a computer-mediated digital form — software platforms that mimic the conceptual city wherein the operators are the authors, the strategists who shape the digital 'city' by governing with linguistic systems vaguely detailed as rules (Twitter, nd) or standards (Facebook, nd), and the users are the urban walkers, "unrecognised producers" and "discoverers of their own paths" (de Certeau, 2002 xviii), armed with the unpredictability and adaptability to influence change in their environment.

Contextualising Platforms and their Politics

The Technological Apparatuses of Online Political Spaces

Some of the most noticeable changes wrought by new systems of communication serve to disrupt existing norms and change the ways in which existing organisations must operate. Driven by the emergence of the internet and changing practices of communication, the familiar intermediaries to political, cultural and industrial participation of decades past have undergone dramatic change. Theorists like Katz (1988), Chircu and Kauffman (1999) tend to call this disintermediation, a term used in economics to describe the removal of intermediaries — often termed middle-men — such as wholesalers or retailers. This description is perhaps problematic, as although internet access, services and related technologies have resulted in a lessening of importance or relevance in some intermediaries, it has not altogether removed intermediaries. In fact, because the transaction between consumer and supplier has become more complex, the technological landscape of the internet has given rise to a broader field of intermediaries than before. In some cases, this is merely remediation (Bolter & Grusin, 2000), introducing digital intermediaries that mimic the tasks of analogue counterparts, but as noted earlier, some of the newly introduced intermediaries may have unfamiliar purposes (such as filtering mail messages or webpage comments for certain terms to identify spam) and their operation is entirely invisible to a consumer or other intermediaries — for example, data analytics software, payment gateway services and anti-spam filters. These new intermediaries can have

divergent purposes, that do not necessarily relate to or facilitate the needs of the user. This new, hyper intermediated (Morozov, 2012) environment increases the political complexity of user interactions by introducing the influences and imperatives of new and unfamiliar participants. This is also noted by the Electronic Frontier Foundation (n.d.) in a campaign entitled ‘Free Speech is Only As Strong As the Weakest Link’ that analyses the most fundamental intermediaries between writers and readers, including web hosting providers, upstream providers, domain name system (DNS) servers, internet service providers (ISPs), search engines, third-party platforms, payment gateways and other service providers. A key example of this is the role that payment service PayPal played in the blockading of donations to the WikiLeaks organisation after the 2010 release of U.S. State Department Diplomatic cables (Poulsen, 2010).

Another complicating factor is the political orientations and goals of the owners and operators of social networking platforms. Consider, for example, Facebook founder Mark Zuckerberg’s public statements about privacy and Facebook’s social norms; “We view it as our role in the system to constantly be innovating and be updating what our system is to reflect what the current social norms are. [...] We decided that these would be the social norms now and we just went for it.” (Kirkpatrick, 2010). Or, alternatively Google Chairman Eric Schmidt who has previously dismissed users concerns regarding data collection, retention and privacy by announcing in a televised interview that “If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place” (Esguerra, 2009). Both of these statements are interesting, as they reveal a certain urge to shape social norms — either explicitly, by determining how users should be able to control their

privacy, or implicitly, by questioning the necessity of privacy as a social need. The operators of social media platforms have explicitly political motivations and their roles as operators allow them a range of instruments — like platform design and policy — that can foster norms in the behaviour of their user-publics to suit their political ends in ways that are not obvious to the participants.

Lingua Platform

In a widely-recognised article entitled “The politics of ‘platforms’”, Tarleton Gillespie (2010) explored the language of the term ‘platform’ and illustrated the ways in which its many metaphoric qualities are used by platform operators to position and influence the ways in which we understand and discuss the platforms. ‘Platform’, he argued, had become a popular term in the public discourse and self-assessments of online intermediaries, and functioned partially as a way of describing the delivery of their services. As platform operators encourage the use of the term ‘platform’ it served to become a trope — “a discursive resting point” — of the language that serves to “sanction and sanctify” a certain method of understanding the companies, software and their effects and influences (Bazerman 1999 and Bordieu 1991 in Gillespie, 2010 348-9) which in turn influences discussions in the commercial and political interests of platform operators. To do this, ‘platform’ serves as a semantically-rich, structural, ‘root metaphor’ (Mio, 1997 125) — Gillespie highlights this with reference to other uses of the term.

‘Platform’ serves: as architectural metaphor, to impart the concept of a human-built structure; as computational metaphor, to distinguish (often through market-style competition) both technological and territorial boundaries like programming languages and software architecture or device and software ownership and control; a figurative metaphor, to assert the ‘platform’ as an arena that conceptualises opportunity for experience and industry; and as a political metaphor, conceptualising an open space for political positions to be announced and enacted (Gillespie, 2010 349-50). The diverse metaphors and signified meanings within the term ‘platform’ highlight why it has become problematically ubiquitous in its use in popular and academic discourse. With a single word, the operators of a service can imply: that users have a figurative platform for free expression; that advertisers have a structural platform for commerce; that platform operators have a level of political and legal abstraction from the positions and actions users take. Moreover, the term ‘platform’ allows a figurative pose that operators are politically motivated to use — a method of absolution through abstraction, where operators can neatly dismiss their roles and divert the focus of unwanted scrutiny or responsibility to Others: content, advertisers and users.

Gillespie’s concern that ‘platform’ could become a ‘discursive resting point’ for discussion is warranted given the term’s use as broad media-style categorisation. ‘Platform’, like ‘radio’ and ‘television’ has become a shorthand for discussing a number of internet-based companies with widely disparate content, policies and politics. Just as that ‘radio’ categorisation is broadly used to include talkback shows like The Alan Jones Breakfast Show and The Howard Stern Show, the documentary reporting of the BBC’s World Service and the light entertainment of NPR’s Fresh Air

program, terms like ‘platform’ are practically devoid of context and detail. A problem perhaps is that the promise of ‘platform’ as a generic metaphor was a compelling enough meme to influence broad adoption of the term. As Gillespie notes, the fact that the term has moved into common parlance suggests that the notion of ‘platform’ is a compelling one to many people (Gillespie, 2010 355), in spite of its potential to be misleading. As a result of this, the commonly referenced notion of ‘platform’ reflects the ideological paradigm implied by its creators and frames public discourse, and in turn, influences the discourse of participants and policymakers on the terms of ‘platform’ operators (H. Galperin, 2004 161) and allows them to trumpet or downplay their role as operators to suit their needs (Gillespie, 2010 357-8), allowing operators to slip into the guise of neutrality by positioning the platform as an apolitical conduit. Indeed, this guise and the presumption of neutrality it often affords is likely a compelling promise for platform operators. As Pasquale (2015 282) argues, the power to regulate is as much a function of architecture as it is of ideology — and in this regard, platform operators are often able to deploy a remarkably flexible regulatory environment towards users.

‘User’ is a similarly common term in both popular and academic discussion. Its prevalence is best explained by its utility: ‘user’ serves as a descriptive term to describe the people who ‘use’ something, typically a platform or software, and serves to describe them by referring to a past and present relationship with that something. ‘Users’ are often stakeholders of platforms because of their continued use and thus continued relationship with the platform. The term however is interesting, given the relatively nondescript choice of language. Other stakeholders are often described in more descriptive or evocative language: ‘advertisers’, ‘artists’, ‘content owners’,

‘producers’ (Geller, 2013). In some cases, this serves to stratify different classes of stakeholders, as Youtube does with its distinctions between ‘advertisers’, ‘artists’ and ‘partners’. More broadly, this serves to constitute users as the Other, distinctly distancing the relationship from other stakeholders and classifying it as a consumptive relationship. This is often, but not invariably, an incorrect perspective that mischaracterizes the relationship that ‘users’ have with the platforms by ignoring the extent to which users are creatively interacting and engaging in diverse and responsive acts, such as that of production and usage (Bruns, 2008 15-21) or dissent and rebellion.

The use of metaphor does not stop at the explanation of the medium, or the role of the people who use them. Platforms are also rife by design with metaphors used to describe technological functions and affordances. In the languages that develop around these platforms ‘tweets’ refer to text and image broadcast rather than birdsong, ‘favourites’ refer to saved ‘tweets’, ‘posts’ have nothing to do with mail and ‘friend’ is suddenly a verb, ‘follow’ implies but does not necessitate movement and you might ‘like’ something that you do not find pleasing at all. These are not each necessarily problematic, but they do highlight a tendency to appropriate existing social and societal constructs as metaphors to describe functionality. This in turn embeds an existing form of meaning into the platform functionality and can serve to politicise the technological action. The functionality of a ‘favourite’ on Twitter is most akin to saving or ‘bookmarking’ a selected tweet for later reference. Similarly, the term ‘friend’ implies an intimacy not necessitated by the platform or its function. In both cases the metaphor used to describe the function conveys a different meaning and politicises the it in a manner not implied by its functionality.

“Does my intermediation look big in this?”

In 2011 a brief period of rioting and civil unrest in Britain was unceremoniously referred to by reporters and media sources as 'The Blackberry Riots' (The Economist, 2011). The characterisation, 'Blackberry Riots', made reference to the usage of mobile technology, social media and peer-to-peer communication platforms in organising rioters and looters, in notifying bystanders, and in organising the volunteering effort for the clean-up. The use of social technology by the rioters, in particular, catalysed a range of media commentary about the uses, consequences and potential for regulation of social media platforms. As politicians, journalists and citizens searched for meaning in the unrest, they also began to voice concerns about technology's role in the riots. “It is a mistake to look into the mirror and try to break the mirror [...], Whatever the problem was [that caused the riots] the internet is a reflection of that problem. If you have a problem, use the internet to understand what the problem is.” Google chairman Eric Schmidt argued (Robinson, 2011). This choice of metaphor, a mirror, to describe the role of technology in the unrest is interesting. It casts 'the internet' and its constitutive parts (including software platforms and service providers like Schmidt's Google) as a neutral factor in society, like a conduit mediating the social and epistemological world it inhabits. This 'technocratic pose' (Alder, 1999 87; Morozov, 2013 151) is useful for Schmidt and allows him to easily deflect reactionary criticism or assertions of technological determinism. However, the metaphor is also a deeply reductive one — it flatly ignores the role and effects of internet technologies.

The “technocratic pose” as identified by Alder referred to an epistemological framework employed by scientists and manufacturers during the French Revolution

that served to shelter themselves from retribution by conceptualising a neutral ground which could obfuscate the extent to which they influenced and engineered the revolution (Alder, 1999 296-301). In the modern context the technocratic pose is demonstrably still in use — and perhaps best evidenced by its use in the language deployed by platform operators and companies that provide software-based services over the internet. Google, like many other online service providers, are not mere conduits or mirrors for human behaviour — the ways in which they mediate human behaviour and personal interaction is a great deal more complex. The most regularly used metaphor is that of a 'platform', a term frequently used by internet-based service providers, academics and technology journalists. As noted earlier, 'platform' is still a problematic metaphor (Gillespie, 2010 348), because of its multiple meanings and implications — but it remains a more useful and helpful term to interrogate the social discourses around internet and technological mediation with than Schmidt's 'mirror'.

On inspection, the evidence that platforms are anything but neutral is quite obvious. We can decipher more of the political influence of platforms by analysing their policies of rules and enforcement, their user-facing design and support decisions and their corporate cultures. In recent years, we have seen Facebook's decisions to censor certain types of nudity and certain types of violence (Ingram, 2013), Google's crackdown on pseudonymous users (E. Galperin, 2011b), and Twitter's commitment to be the "free speech wing of the free speech party" (Barnett, 2011) — to name a few issues. Each of these are indeed political acts. They constitute cultural, organisational or ethical presumptions about how users or platforms should interact. As a result, users, governments and other platforms come into conflict with one another (for instance, boyd, 2011; Cameron, 2013; and Panzarino, 2012 respectively) when these

presumptions do not align with their own cultural or ethical beliefs or organisational goals. Although it would be churlish for me to suggest that Schmidt's intentions were devious, it is undeniable that his use of the technocratic pose to misrepresent his company's role and effect serves to mislead further discussion and deeper understanding of the Google company and its associated platform(s).

There is perhaps no clearer contradiction to Schmidt's pose than the fact that Google, like Twitter, Facebook and many other platforms, have written and enforced policies to attempt to govern and regulate the behaviours of users interacting with their platform. Simply put: Mirrors do not moderate; platforms operators do. The key players in this field are well aware of the importance of their role as moderators in part because of their continued efforts at collaboration in determining policy. Jeffrey Rosen (2013) described the efforts and evolution of content policies at Facebook, Google and Twitter as a consortium of 'deciders'. These 'deciders' weigh the expectations that they host and operate platforms that respect or encourage notions of democracy and free expression against the expectation that they regulate and guard a "well-ordered society" (Waldron in Rosen, 2013). Whether well-acknowledged by policy-makers and users or not, the people working at Facebook, Google and Twitter have an unambiguous role in influencing the use of their platforms, and as a result their government of the platform can foster conflict where it fails to meet the expectations of the platform's users. This isn't a typical style of governance. Rather, this intermediated governance presents a new way for platform operators to wield power over their domain.

Modulating Control through Distance

The affordances of modern internet technologies like JavaScript, and cookies create a process of algorithmic inference what forms a new structure for potential regulation and control. As Cheney-Lipold (2011 167), this moves users into a framework in which all aspects of their identities might be measured and quantified, and critically, a framework in which identifiers and identities are inferred by code and algorithm, rather than supplied by a user. As Lessig (2006, 1) argued, the rules of digital spaces are intentionally and unintentionally written into internet technologies that structure the affordances that guide how users can act in digital spaces.

Cheney-Lipold argues that these frameworks enshrined in the code of digital platforms form a regulative approach to online governance that “allows for a distance between power and subject”. This is both a literal and figurative ‘distance’, an abstraction of governance that serves to undermine the user’s perception of being governed, and offers platform operators the ability to remain oblivious and indifferent to the effects of their governance. By restricting, customising, and otherwise ‘personalising’ the conditions of possibility, platforms can configure particularly limited affordances for users. This style of platform governances regulates and predicts users lives by tethering the potential for alternatives to the previous actions of users based on an observed history of other users, or countless other potential algorithms designed in the platform’s development. This style of governance works to fundamentally limit the possibility space of political engagement in these spaces, as technological affordances for engagement are shaped in response to the existing behaviours and discourses of the users being governed. Cheney-Lipold situates this

argument within discussions of biopolitics and bio power — framing this with a Foucauldian discussion of population and control. The move towards digital platforms using systems like those described here results in a loss in control for people's aspects of identity as digital shareholders shape and design the categories of behaviour and identity they wish to infer for users.

Common Areas of Conflict

A useful way of taxonomising conflicts on social media platforms is to identify the participants and the stakeholders affected by a conflict. By doing this, we can determine whether the conflict is primarily between users and platform operators — a vertical conflict — or primarily between users and other users — a horizontal conflict. Determining the orientation of the conflict allows researchers the opportunity to tease out the complex relationships of the stakeholders in a conflict. Conflicts tend to be reactionary in nature; a tension between policies, practices or participants is catalysed by a particular event — such as a policy change, or an undesired user action — and users or operators react to the conflict according to their motivations. Vertical conflicts arise from tensions caused through regulation and enforcement: platform operators regulate behaviour and enforce regulation in reaction to the actions of users in an attempt to govern the space that they operate; and users react to strategies of regulation or enforcement that affects them or to their own perception that the platform operators are poor custodians of the platform they operate — such as concerns that platforms poorly govern themselves, or that the platforms prioritise their vested interests at the expense of ordinary users. Horizontal conflicts are more difficult to predict — they

arise from conflicts between users, and as such their causes can be as varied as the users themselves but contentious political issues are often a factor. Because of this, horizontal conflicts are often multi-axial: The users involved often seek arbitration or vindication from platform operators as the custodians of the space.

A key area of conflict on social networking platforms is related to formalised change to the platform itself. These are vertical conflicts that are spurred by user reactions to the alteration of existing platform infrastructure (like a design or feature change) or the alteration of previous platform policy. Where platforms deploy strategies to elicit self-regulation, they encourage users to take responsibility and ‘own’ the spaces they help to govern. This becomes problematic when the platform ‘owners’, the operators make a unilateral decision to change the platform without the consent or deliberation of users. Users on Facebook have rebelled against features as mild as the change of profile design in ‘Timeline’ to changes as intrusive as ‘Graph Search’ (an advanced boolean metadata search of all publicly available Facebook content) (Garside, 2013). Feature and design changes in particular are an obvious demonstration of the ease with which the operators can upset the environment of users by disrupting their familiarity and their routine with the systems that they navigate. Beyond disruption, features like Graph Search can prompt conflict by demonstrating a usage of power that users might find troubling. Indeed, some of the initial criticisms of Graph Search noted that the ability to search user-supplied metadata with advanced boolean queries could allow users to be found by searching for attributes. This was perhaps most notably criticised by blogger Tom Scott (2013), whose Tumblr blog *Actual Facebook Graph Searches* garnered significant media attention by juxtaposing search queries to find politically complex or devious results, searching for: ‘Mothers

of Catholics from Italy who like Durex’, ‘Family Members of people who live in China and like Falun Gong’, ‘Islamic men who are interested in men who live in Tehran, Iran’ and ‘Single women who live nearby and who are interested in men and like Getting Drunk!’. The feature highlighted disparities between user expectations of privacy as a concept, and Facebook’s implementation of ‘privacy’ as a feature.

The vertical conflicts between users and platform operators can manifest in a number of guises. In some cases, the platform operators use of enforcement methods exacerbate tensions by revealing a problematic policy or enforcement strategy that is not typically visible to users. For example, poorly-designed algorithms enforcing anti-spam policies have resulted in false-positive identification of user content as spam and the subsequent prevention of posting that content on Facebook. The algorithms identify certain comments as spam because the comments do not conform to the expectations of the algorithm (and by extension the expectations of the platform operators who use the algorithm as a method of enforcement). When the algorithm identifies ‘spam’ it provides an error message that reads “This comment seems irrelevant or inappropriate and can’t be posted. To avoid having comments blocked, please make sure they contribute to the post in a positive way”. Taylor (2012) noted that whilst Facebook describes the algorithm as an anti-spam mechanism, its effect is to institute normative standards of “appropriateness” by using algorithmic measure to quantify hidden linguistic metrics that can preclude users from participation in public conversation. This similarly occurs when algorithms are designed with the intention of censoring topics that might be politically controversial. York (2010) highlighted in 2010 that Facebook’s algorithms stopped certain words such as ‘Al-Qaida’, ‘Nazi’ and ‘Palestinian’ from being used in the names of Facebook Pages. Furthermore, O’Neill (2016) demonstrates that the assumptions imparted in the development of systems for

algorithmic policing can be flawed for a range of reasons, including biases inherent in previously collected data, biases inherent in data collection mechanisms, and biases imparted by creator's beliefs and assumptions about human life and behaviour. This is in-turn complicated by the secrecy often afforded to platform designs and decisions.

Vertical conflicts can also arise through capricious or wrongful use of enforcement methods. In this case we can look to the exploits of twitter-satirist Rurik Bradbury, who is perhaps better known on Twitter as 'Fake Jeff Jarvis' (using the handle @ProfJeffJarvis) (Love, 2013). The account is a satirical parody of "buzzwords, empty ideas, and the titanic egos that rule [the] tech and media [industries]" (Love, 2013). An update to the Twitter platform and its associated apps started replacing the hyperlinks to images shared on Twitter with embedded inline images that automatically loaded, removing the need for a user to click on a link to see an attached image. A number of users questioned the wisdom of this feature, including @ProfJeffJarvis, who tweeted an image of a penis before remarking "I recently lost 99 followers", "Puzzling increase in unfollowers tonight. I'll need to look into this carefully using Big Data Metrics" before including an @mention to notify Twitter's CEO Dick Costolo (@dicke): "Totally stumped by my #UnfollowerMystery. @dicke, could you have an engineer take a look at my raw data?" (@FakeJeffJarvis, 2013). Costolo responded and seemed to take the matter seriously, perhaps confusing 'Fake Jeff Jarvis' with the real Jeff Jarvis, and shortly after their interaction the satirical account was suspended. This was particularly peculiar given the permissiveness of Twitter's content policies as neither nudity nor pornography are forbidden on the platform. The suspension was short-lived as Twitter's operators re-enabled the account later in that week after the episode made its way into the tech press, where Bradbury

noted that the saga played off some interesting tensions in Twitter's operation as a platform, including the tension between Twitter's roles as both a platform and a publisher (in Hill, 2013). The seemingly erroneous censorship also highlights the precarity of editorial censorship, where the judgements of the individuals empowered to censor users can be capricious and politically-motivated rather than a stoic interpretation of platform policies.

The resolution of conflicts is often not favourable to the users involved in conflicts, especially where the platform operators have considered that the issue embroiled in the conflict is an issue relating to their commercial viability as a business. For this reason, conflicts relating to issues of advertising (such as the redesign of a user interface that facilitates larger advertisements) are seldom resolved to the benefit of users. For some issues of user upset, platform operators seem prepared to take a wait-and-see approach, preferring to gauge the scale and sentiment of user discord before taking any actions — and potentially waiting out users, in the hope that they give up difficult campaigns.

The policies of Facebook are, in contrast to Twitter, markedly less permissive for users — and as such are a hotbed for vertically and horizontally oriented conflicts. The platform maintains a list of standards and is known to police the types of content it allows, and their strategy of encouraging social enforcement involves both horizontal and vertical tensions into the conflicts relating to content that is verboten on the platform. Facebook's 'Community Standards' and 'Statement of Rights and Responsibilities' variously deems the infringement of laws, intellectual properties, sensitive financial information, non-consensual tagging, threats of violence or self-

harm, 'hateful speech', 'graphic content', 'pornographic content' and 'explicitly sexual content where a minor is involved', whilst making reference to undisclosed 'limitations on the display of nudity', as unacceptable on the platform and subject to removal and flags the possibility of punitive action to be taken against the publishing user. For Facebook to act on content that breaks these rules, they rely on their users using the platform's 'report' functionality to flag it for review by one of their User Support Teams in Menlo Park (California), Austin (Texas), Dublin (Ireland) or Hyderabad (India). After a judgement is made, the user is notified of the outcome of their report — and if a removal is made, so is the original publisher of the content. As a result, the distinctions that the platform operators draw in terms of what is allowed and what is not can become particularly obvious to the affected users. As such, some of the most controversial instances of platform intervention on Facebook to date have been editorial judgements of what is and is not worthy of censorship.

Many conflicts on social media platforms are multi-axial and demonstrate both user-user and user-operator tensions. One of the best known examples of these multi-axial conflicts came early in 2013, when two well-known feminist activist groups called Women Action and the Media (2013) and the Everyday Sexism Project noted a trend of misogynistic content being posted onto Facebook. They identified captioned images in the style of memes that were, to be blunt, grossly misogynistic — including unsettlingly graphic images of women who had been bloodied and bruised with offensive slogans overlaid reading “1/3 of women are physically abused, 2/3 of men aren’t doing their job” and “if you can’t handle him at his worst, you don’t deserve him at his best”, as well as Facebook user groups and pages with offensive names like “Raping a pregnant bitch and telling your friends you had a threesome” (Bates, 2013).

The content they identified seemed to quite clearly transgress upon Facebook's community standards, with reference to the 'hate speech' and 'graphic content' sections. The policy expressly forbids "individuals or groups to attack others based on their race, ethnicity, national origin, religion, sex, gender, sexual orientation, disability or medical condition" and declares that "graphic images shared for sadistic effect or to celebrate or glorify violence have no place on our site" (Facebook, nd). When some of this content was determined by Facebook moderators to not warrant removal, the involved users shared images of their rejection report and campaigned against both the platform and the advertisers who used the platform. After significant media attention and the withdrawal of a number of advertisers, Facebook's Vice President of Global Public Policy Marne Levine (2013) pledged to reform the existing training for content moderators and alluded to the potential for changing platform policy in the future.

Of course, Issues of sexism and misogyny are not exclusive to technological spaces. It bears discussion that attitudes of sexism and misogyny in society are reproduced in, amplified by, and adapted to the technological platforms observed in this research. This particularly apparent in the cases analysed in this research project, a number of which relate quite strongly to issues of sexism, misogyny and gender. This will also be discussed in more detail later in this chapter.

The fact that platforms like Facebook, Twitter, Youtube and more have 'report' tools is indicative of social methods of policing and enforcement for platform policies, a practice that advocates that users assist in policing horizontal conflicts in accordance with the platform operator's practices. In spite of the possibilities proffered by advancing computing technology and conditional algorithmic programming, the

policing arm of a platform cannot ‘be everywhere’ and ‘see everything’, so to speak — and as a result platform operators tend to elicit the help of users through content reporting tools. The underpinning logic is similar to that used by the Department of Homeland Security (2010) in their “If You See Something, Say Something” campaigns and the Thai Government’s “Protect The King” website for reporting anti-monarchy sentiment (BBC News, 2009). The platforms operators, much in the same way as the Thai or United States governments, use a style of governmental authority to establish a user-led policing effort whereby users acting as investigators and informants can police themselves. The end game of such self-policing policies is that users could ultimately provide the benefits of a policed society or social network through self-regulation and social regulation without necessitating shows of force. This is a particularly practical manifestation of the effects of surveillance in disciplining citizens as theorised by Foucault (1995), and similarly reflects the governmentality (discussed later in this chapter) which platform operators seemingly hope to inspire in users. The ‘report’ systems currently deployed by platforms elicit the help of users to socially enforce policy like Facebook’s community standards or Twitter’s rules but can also be gamed and intentionally misused by users seeking to inconvenience platforms, advertisers or other users in horizontal conflicts. Aja Romano (2013a) highlights the case of video games critic Anita Sarkeesian’s “Tropes vs Women” video series which was temporarily removed from Youtube after a number of critics used the reporting tools available on Youtube to highlight it as a transgressor of Youtube’s policies — in spite of its lack of transgression. Though the video was restored after Sarkeesian appealed its removal, the fact that it was removed in the first place signifies either an automated response to the reporting (a practice that Youtube disavows in its Community Guidelines asserting “flagged videos are not automatically taken down by

the system” (Youtube, nd)) or that a moderator reviewing the flagged video applied the policy injudiciously, with little regard to either the video content or the policy itself.

Politically-motivated Twitter users deployed a similar process in attempts to silence, or at the very least inconvenience conservative American political activists, by using Twitter’s “Report for Spam” tool en masse to trigger automated account suspensions. The victims and their supporters perhaps slightly over-state their tribulations by referring to the suspension as a ‘Twitter Gulag’ in reference to the horrific Stalin-era Soviet labour camps that often incarcerated political prisoners, but the cases warrant attention nonetheless. Pseudonymous conservative blogger ‘SemperBanU’ (2012) detailed a process that seemed to show proof-of-concept, and then evidence of a method of exploiting Twitter’s “Report for Spam” function that would automatically suspend user accounts after certain algorithmic markers had been met. The blogger contextualises this in discussions of information warfare and alleges that the use of this tactic is widespread amongst his liberal political opponents. Without access to Twitter’s systems or official commentary from the platform operators on this matter it is difficult to consider the veracity of these allegations, but the more important issue that remains overlooked by SemperBanU’s analysis is that it seems to reveal some of the assumptions that Twitter engineers have made in the design of the Twitter platform’s anti-spam systems. Namely, that users will only deploy the report function in good faith reports of ‘spam’-related misconduct, and that veracity can be presumed on the basis of a multitude of user-submitted reports.

It seems likely that the functionality of the report and flagging systems is prejudicial towards the interests of the platform-operator and their commercial partners

and stakeholders or their perceptions towards what problems are and how to fix them. We can see this clearly in Twitter's (Twitter, n.d.) 'Rules', which highlight a disparity between the platforms policies regarding bad user behaviour (especially relating to abuse, or harassment of other users), and regarding content that allegedly infringed upon the copyright of a claimant. This is mirrored in the way that reporting functionality seems to prioritise the policing of issues of intellectual property and copyright. In a similar vein, videos on Youtube flagged for copyright claims are automatically barred from advertising revenue during the course of the copyright claim dispute, regardless of the veracity of the claim (Totilo, 2013), which can allow bogus claimants to interfere with the revenue stream and copyright infringement records of Youtube users and partners alike (Gutelle, 2013). Similarly, the anti-spam algorithms that Twitter seems to use result in a crowd sourced definition of 'spam' that can be gamed for political ends. In cases like that of the 'Twitter Gulag' and the abrupt removal of Anita Sarkeesian's videos, users subvert the expectations and systems of platform operators to wage their own horizontal conflict amongst each other.

Horizontal conflicts are, by nature, common. Users disagree with each other, individually or en masse, about a wide variety of subjects from the banal ('Mac or PC?', 'Superman vs. Batman?') to the more conventionally political (government policy, current affairs). In many cases horizontal conflicts occur without particular oversight or intervention from platform operators — but in large, sustained conflicts like 'flame wars' (N. A. Johnson, Cooper, & Chin, 2008) or in conflicts where users seek the intervention of operators, the dynamics of the conflict often shift from horizontal, inter-user, conflict to a multi-axial dynamic of tension between differing groups of users and platform operators. As users and platform operators continue to

stake claims to social media platforms as political spaces, conflicts of political opinions and conflicting expectations of governance have become more regular occurrences. Whilst platform operators attempt to govern the space through the imposition of strategic regulation and enforcement, users attempt to exercise their own sovereignty over the spaces that they use by using tactical methods to subvert platform operators and other users alike.

Platforms and Strategies

Governmentality and Platformed Ideology

Given the presence of users (a population or citizenry), software platforms (a jurisdiction) and a group of owners and operators (a government) it could be helpful to consider social media platforms as a type of non-electoral, governed space. Government, insofar as it is afforded authority and power, is a structure that attempts to shape through deliberation aspects of the behaviour of its citizens according to certain types of norms for a variety of purposes (Dean, 1999 10). In any attempts to then analyse the ways in which social media platform operators govern, we must be concerned with the forms of knowledge, techniques and means used by governments, the types of authority and agency they exhibit, the entities they attempt to govern as well as the ends they seek, the outcomes of their attempts and their broader consequences (Dean, 1999 11). These constitute frames of research for what Michel Foucault terms ‘Governmentality’. As Blau (2013) notes, Foucault struggled to consistently define Governmentality as a concept and area of knowledge, cycling through a number of concepts, theories and descriptors of governance and governmentalisation in lectures before his death. Blau concedes that Foucault’s descriptions were formative and not “intended as the finished product” and cautions researchers against being too deferential to Foucault. With that in mind, this research has used O’Farrell’s (2005 107) synthesis of the term; “the rationalisation and systemisation of a particular way of exercising political sovereignty through the government of people’s conduct”.

Dean (1999 15) argues that the very notion of government and governing entails the possibility that the governed are “to some extent capable of acting and thinking otherwise” — that is, capable of dissent and disagreement. As seen in de Certeau (2002), this is evident in most forms of ‘government’ from the traditional country or locality based-government, to the types of government attempted by powerful economic and cultural forces. As highlighted earlier, governments and dissidents often have unequal power and unequal opportunities to affect change to the systems they operate or operate within. Governmentality, in turn, is concerned most with strategies as a rationalisation of governmental power.

Iván Torres notes that “The paradox and contradiction of our contemporary governmentality is: on the one hand, it is open and tolerant, while on the other hand it deploys much more flexible, penetrating and exhaustive forms of control” (translated in Gehl, 2013). Perhaps herein lies an interesting distinction regarding the strategies used by platform operators — Not only are they ‘not as they seem’, so to speak, but they are often intentionally designed in that manner. This is an insidious manifestation of strategies of governmentality, where the technological structures afford opportunities for invisibility through duplicity and thus the ability to govern and police without public demonstrations of force. By offering an avenue for users to express themselves, such as a text input that asks, “What’s on your mind?”, the platform operators seem to invite user’s expression, although their rules consider that certain types of expression; expressions of nudity or expressions that feature unlicensed copyrighted material for example, are forbidden as a matter of policy. Similarly, expressions of identities like names or genders may be forbidden as a result of platform design. As a result, the understandings we can gain from the uses of social media

platforms can be politically muddled by ideologies that are not immediately obvious. By hiding the will and the power of platform operators behind notionally inviting language and design, the platform becomes invested in the creation of self-regulating behaviours whereby its users are mindful of the possibility for censorship and punishment and moderate their public conversation accordingly. These elements of language and interface design show that the ideology of governmentality can be encoded into the technological structure of social media platforms.

Media theorist Fredrich Kittler argued that as media became digital, the ‘noise’ of politics and ideology in communications would cease to exist (Packer, 2013 2). Kittler’s prediction did not come to fruition and the ideology embedded in media and communications has not been filtered out by the advent of digital media. As Packer (2013 3) argues, although internet services like Google might not present as politically, or ideologically motivated — but that is not to say that Google (or any other software, for that matter) is not ideological. Ideology may not be necessarily expressed as an algorithmic function in digital media, but algorithms are designed by platform operators and software engineers — to favour, for example, novelty over difficulty, or celebrity over obscurity for example — and process data for usage for use in a variety of means. Platforms, like the structures of code and algorithm that underpin them, are rife with the subtext of political ideology. As Ullman (Ullman, 1995) notes, software embeds the “engineer's assumptions and presumptions” in the code, and “reproduces and re-enacts life as engineers know it”.

Tarleton Gillespie (2011) proposes a number of concerns that researchers should note;

“We do not have a sufficient vocabulary for assessing the intervention of these algorithms. We’re not adept at appreciating what it takes to design a tool like [Twitter’s] Trends [list] – one that appears to effortlessly identify what’s going on, yet also makes distinct and motivated choices. We don’t have a language for the unexpected associations algorithms make, beyond the intention (or even comprehension) of their designers. Most importantly, we have not fully recognized how these algorithms attempt to produce representations of the wants or concerns of the public, and as such, run into the classic problem of political representation: who claims to know the mind of the public, and how do they claim to know it?”

Gillespie’s concerns are all well-founded and serve to highlight the inherently political nature of designed software and technology. Though data may be politically-charged or politically neutral, the ways in which social media platforms use data can be transformative, ideological and invisible. When a platform is designed with functions or designs that are ideologically motivated and politically charged, as say, Youtube’s ‘Content ID’ algorithm (Youtube, n.d.) it can be used to enforce a punishment and to influence the behaviour of users in the future. This serves to provide another example of what a platform can be — an instrument or structure of power, to be shaped and exercised by its operators.

Another example of this effect are the so-called ‘Trending’ algorithms deployed in most popular platforms. ‘Trending’ features are often communicated to users as an at-a-glance summary of what other users on the platform are discussing. But the metrics and qualities that make topics eligible for consideration by these algorithms are hidden or de-emphasised from users. Some aspects of these algorithms can be

gleaned from close observation — others from the way platforms explain and discuss the functions, features and affordances of their own platform. As Gillespie (2011) argues, an aspect of this may be a preference for ‘novelty’ in news and topics highlighted. Thus, Facebook and Twitter’s algorithms may partially favour ‘new’ news, over news that relates to a ‘slower’, sustained issue. The opacity of these algorithms, and of myriad other functions and affordances of social networking platforms, makes them obvious examples of what Franklin (1999 17) describes as ‘prescriptive technologies’ — technologies that are exceedingly effective and efficient in their purpose, but also technologies that come with an “enormous social mortgage”, a conditioning to accept technologically imposed orthodoxy as “normal” and necessary.

Social networking platforms and underlying software can be understood in two senses: in the first, as mentioned earlier, as a jurisdiction for this form of governmental power — a public with a user/citizenry that can be governed; and in the second, as a technique of this power — an instrument, akin to Foucault’s conceptualisation of prisons or clinics, with which government dominion over citizenry is achieved. Technology, in the view of Foucault was both castigated and criticised as a means of control — and celebrated as a project for self-emancipation (Behrent, 2013 95). Both of these understandings are cause for further investigation, so that we can better understand the motivations and rationalisations of attempts at governing users and understand the ways in which platforms are used not only as jurisdictions but as strategic instruments of enforcement for the government of platform operators.

With an Iron Fist or Maslow's Hammer?

Politicians, newspaper editorials and media moguls alike have proclaimed the democratising powers of internet technologies for decades. Reagan prophesised that “the Goliath of totalitarianism” would be struck down by “the David of the microchip” (Reagan 1989 in Konstantinidis, 2003). Rupert Murdoch heralded innovation in telecommunications as “an unambiguous threat to totalitarian regimes everywhere” (Murdoch 1993 in *The Observer*, 2003). Journalists and editorials alike have celebrated “Twittered” revolutions (Sullivan, 2009). What is perhaps overlooked in the discussions, debates and disputations surrounding the purported democratising effect of social media is the very nature of the systems themselves. Whilst some claims could be made as to the effects that uses of technology have had in instigating democratic practice, the technologies could very seldom be described as fundamentally democratic, and the freedoms enjoyed by users are conditionally based on the rules and designs of the platform operators and designers. Our understandings of a user’s digital-democratic freedoms are, frankly, in their formative stages and the concept is fraught with disagreement. Whilst few would contest the rights of the operators of Facebook or Twitter to censorship of their platform as their own choices for political or corporate speech and expression, courts have established legal precedent to also protect the actions, such as ‘likes’, of users on these platforms as expressions of their ‘freedom of speech’ and as such afford them constitutional protections in the United States of America (Kelly, 2013). These judgements, of course, are different from jurisdiction to jurisdiction, and the rights and protections afforded to users and platform operators will vary accordingly — which, in time, may lead to a regulatory patchwork of rights and regulations for users and platform

operators. In the absence of a set of clearly understood rights and responsibilities, it is perhaps most useful to look at the practice of democratic governance on these platforms. Whilst Facebook has flirted with the ability for users to vote on their policies (Kelly, 2012), the function attracted few users, and seemed impractical to operate at scale — with most users either not knowing of its existence or not caring to participate. The system eventually required approximately one third of the Facebook’s large user base to cast a vote (over 300,000,000 voting users — and growing) for the result to be judged as binding to their policy development.

In essence, this neutralised Facebook’s experiment with user-deliberative democracy and shared governance, and Facebook eventually terminated the project in 2012, replacing the system with a mechanism in which users could e-mail the platform with suggestions. This process was heralded by a public vote (Schrage, 2012), held by Facebook to determine the future of the shared governance project and the creation of a new Statement of Rights and Responsibilities. Whilst the vote’s results were largely in favour of retaining user’s ability to vote on site governance, not enough users voted for the results to be held binding to the platform, and as such the platform considered the results ‘Advisory’, and moved ahead with their proposed alternative (Schrage, 2012).

In analysing the governance and governmental strategies of social networking platforms, it is useful to draw analogies with the governmental strategies and practices of governments around the world where they bear similarity. Platforms operators and governments alike are disproportionately powerful when compared to the users and citizens that they govern, and whilst platform operators cannot incarcerate dissidents,

they can censor dissenting users and they can impose punishments similar to deportation or exile by use of temporary or permanent bans. Consider, for example, the Singaporean government — a government well known for its coercive strategies in dissuading insurgency and dissent. George (2013 390-2) noted a tendency for dissidents to use what Sharp (in George, 2013 387) describes as “political jiu-jitsu” — a tactic of deliberately nonviolent protest designed to contrast with the violent action of their oppressors, throwing their oppressor off-balance in a political sense and weakening their perceived authority. After several cases where citizens had successfully used the internet as a platform to magnify the effect of their civil disobedience and the harsh responses of the Singaporean government, the government began to learn from experience and introduced different strategies to mitigate opportunities for activist citizens to cause damage to their reputation and regime.

Where the Singaporean government learned to save face by mandating passivity and politeness in law enforcement officers in engaging protestors and restricting public videography to state-approved broadcasters, Facebook learned to save face by introducing new (and potentially controversial) features like ‘Timeline’ and ‘Graph’ slowly and over a staged rollout instead of to all users at once (Gannes, 2011). Of course, neither the agents of an authoritarian government nor the operators of popular social media platforms can be everywhere at once. For many social networking platforms, the enforcement of rules relies on users reporting other users for their transgressions. As noted earlier, Facebook’s policing of content that is forbidden by their community standards largely relies on the reporting of users who work as de facto informants to flag public or semi-public content that may breach the standards. As this is a reliance on supplied information, it results in a system that does not seem to

particularly police users who furtively breach their standards in ways (such as with closed or private groups) that do not seek to attract attention. Conversely, the users who have political objections to Facebook's standards are more likely to breach them publicly to seek attention and more likely to be susceptible to reporting and subsequent enforcement. In this manner, Facebook, like the Singaporean government, is seemingly more concerned with the perception of order than the enforcement of it. To this end, users can subvert the loudly proclaimed policies and rules of the platform, so long as they can agree to not report each other for misbehaviour. Unlike Bentham's panopticon, participants in Facebook's platform can alter their surveillance with simple collaboration and cooperation.

In a similar vein to the 'political jiu-jitsu' discussed earlier, Ananny and Gillespie (2016) helpfully explore public shocks and platform exceptions. To borrow Ananny and Gillespie's description, "The shocks are public moments that interrupt the functioning and governance of these ostensibly private platforms, by suddenly highlighting a platform's infrastructural qualities and call it to account for its public implications." Through the workings of what Ananny and Gillespie describe as a 'Moral Economy', stakeholders in platformed conflicts like users, or the press tend to focus attention on problematic or undesirable parts of platform governance. Ananny and Gillespie give the example of the ride-sharing company Uber, and its application of "surge" pricing — a pricing scheme that exponentially raises prices during periods of high demand. Ananny and Gillespie argue that, by highlighting platform infrastructure, these shocks generate "public indignation and regulatory pushback that produces critical— but often unsatisfying and insufficient— exceptions made by the platform". These exceptions are often reactive announcements, salves intended to ease user anger or investor discomfort, like an apology from platform operator's, and a

promise to do better (often unaccompanied by any plan for improvement). At the introduction of this thesis, I gave the example of Facebook's relatively zealous policing of the term 'dyke', and the frustration that Facebook's context-less policing of this term caused for lesbian women who had worked to reclaim and re-appropriate the 'dyke' epithet into a positive identity, followed by Facebook's apology. This may as well be a textbook example of 'Shock and Exception'. The shock, in this example is the censorship and policing of a term originally used as a pejorative epithet, but now commonly used in a 'reclaimed' manner by lesbians and members of the LGBTIQ communities. The exception of this example is the Facebook organisation's apologies and their platitudinous offering to improve their systems.

Though the examples of Timeline and Graph relate to technical implementations of design-related policy, it is possible for the platform operators to use the powers mentioned to enforce policies relating to social and cultural issues, and governing interaction, allowing a web-enabled form of cultural imperialism (Rosen, 2013) and to use popular web platforms as vessels to promote social, cultural or political norms. Facebook, Twitter and similarly widely used platforms can become vessels for normalising the ideologies that are programmed into them and enforced through policies. In a very real sense platform operators create a system whereby they can indicate and then curate the types of expression that they allow their users to publish, and control the technological affordances of platforms to shape and enforce public conceptions and norms of concepts such as 'privacy' (Zuckerberg, 2010).

Morozov cautions users against the wholesale adoption of a platform operator's political perspectives (or 'Silicon Valley Values') in reference to what he terms as

‘solutionism’ — an urge to use technology to solve perceived problems (such as the public/private dichotomy) without critical analysis. It is a more nuanced theoretical distinction than Maslow’s Hammer and the famous phrase “I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail” (Maslow, 1966 15). With this, Morozov prods the some of the problems that technologies have aimed to solve. As he phrases it, “not all bugs are bugs; some bugs are features” (2013). Whilst a technology may strive for efficiency, it may cost users some of the benefits of inefficiency. Simply put, a ‘problem’ in the eyes of a platform operator may not need fixing from the perspective of a user — or may not actually be a problem at all. Perhaps a more topical and practical example of the potential for disconnect between user values and ‘Silicon Valley Values’ can be most easily shown in Facebook’s policies towards hate speech, and Lesbian, Gay, Bisexual, Transgender, Intersex and Queer (LGBTIQ) users. Facebook is well known as a reasonably progressive technology company. Hundreds of Facebook employees, including founder Mark Zuckerberg, have recently participated in the 2013 ‘San Francisco Pride Parade’ and the Facebook policies against hateful speech specifically note content that attacks people on the basis of their sexual orientation as hateful content. These policies, expectedly, cause consternation amongst Facebook user’s with particularly conservative or regressive views about sexuality and gender. A recent example of this is given by Vadim Gorshenin, the president of the Russian daily newspaper Pravda (Gorshenin, 2013), who wrote an editorial about being punished under Facebook’s hate-speech policies and decried it as a “gay social network” propagandising “totalitarian tolerance”, by silencing his views and ostracising him from the platform. This disconnect between user and platform perspective quite clearly highlights that the views and expectations of platform operators and some of their users may not always

correspond neatly to the views and expectations of other users. Facebook’s operators have made their policy clear on the types of content that it terms “hateful speech or harmful content” (Facebook, 2013a) through blog posts, media releases, and through the information on their Community Standards page — the rules that they expect users to abide by — and in doing so has indicated that it will use the technology and power available to enforce these rules to punish users and remove content that is they consider undesired for the platform. Though this policy has obviously irked Gorshenin, and likely continues to irk Facebook users who share his views, it is also likely to be popular amongst some of Facebook’s other users, given that it does bear some similarities to laws relating to discriminatory or hate speech in western democracies like Australia and the United Kingdom. Nonetheless, the decisions made in this context are not politically neutral.

Indeed, the concept of ‘hate speech’ and the debates contesting methods of fighting hate speech are politically vexed; some consider the antidote to be ‘more speech’ or ‘counter-speech’ – to engage aggressors in debate and to show their opinions as wrong-headed (American Civil Liberties Union, 1994), others consider that the harmful nature of the language and attitudes warrants punishment or censorship (Morris, 2013; see May in both Rigby, 2013). In a recent book on the subject, *Viral Hate* Foxman and Wolf (2013) argued that “It’s nice to imagine a purely technological solution [... but] the reality of the problem makes it clear that, while technology can play a helpful role, it offers nothing remotely like a complete solution”. The authors are closely associated with the Anti-Defamation League, a non-government organisation known for activism against anti-semitism (Anti-Defamation League, 2012), and cite their experience with HateFilter — a program designed by the

League in a seemingly Sisyphean attempt to help parents protect their children from websites with hateful content. The technological fixes falter because human communication is complex. A number 88 may be a reference to the element Radium, or the eighth letters of the alphabet ‘HH’ — a signifier for ‘Heil Hitler’. Similarly, the term ‘lolita’ can be used in reference to a gothic fashion style and subculture, literature in reference to Vladimir Nabokov’s novel or Stanley Kubrick’s film, or in reference to pornographic material that is exploitative of children. As Jonathan Rauch (1995) argued, “the vocabulary of hate is potentially as rich as your dictionary” and as a result differentiating these uses and the many more obfuscations involved in hateful or other undesirable speech will inevitably involve the subjective, political judgement of human moderators. In essence, politics (organisational, personal and otherwise) are fundamental to understanding and interpreting the architecture, operation and use of social networking platforms. Understanding the conflicts of these politics —whether in relation to appropriate or acceptable speech, or a platform’s standards regarding censorship and nudity— is necessary to critically engage with these ideologies in research, and to recognise the political and ideological influences that drive and shape the development of these platforms. In this regard, it is worth noting that both the formulation and enforcement of policies on platforms like Facebook are arguably influenced by the everyday politics of Facebook’s operators, at the executive, developer, and community moderator level. Whilst this is ostensibly most apparent at the level of the executive, where the guiding philosophies and principles of the platform are authoritatively laid out, it is just as noteworthy at the level of developers, whereby the platform’s programming gives shape to technological affordances for users and operators. Similarly, the employees tasked with the moderation and enforcement of the platform’s policies may be faced with judgments that require their

own perspective on issues of policing and censorship. Whilst this process of moderation is often entirely opaque to users and observers, a handful of investigative journalists have conducted interviews (Buni & Chemaly, 2016) with some of the people who have worked in facilities as moderators for platforms like Facebook and Twitter (Chen, 2014). As Chen (Ibid.) notes, platform operators often require that the moderators working for these services sign strict non-disclosure agreements, preventing them from even discussing their work with colleagues and co-workers. Because of this, research and theory relating to practices and policies of moderation and governance on social networking platforms are essentially constrained, and must rely on the data gathered from ‘non-official’ sources, like leaks, whistle-blowers, and anonymous interviews published in popular media as a first resource, and build on this with carefully considered discussion and assumptions about the possibilities and affordances of moderation in the abstract.

Ownership, Control and Responsibility on the platform

Platform operators are afforded controlling powers by their ownership and their development of the platform infrastructure in ways that serve to circumvent or restrict uses of their platform that they find undesirable. Recent examples of these include Facebook and Twitter designing their websites to restrict the extent to which they can be searched, indexed and archived by Google and other web-crawlers. Twitter’s operators have notably changed the rules governing and the functions available in their application programming interface (API) to restrict the ways in which Twitter data is used by non-Twitter applications (Ingram, 2012). By doing this, Twitter’s operators can restrict the portability of the data on the platform. Users who had formerly used

IFTTT — a web automation service named in reference to conditional computer programming syntax: ‘if this, then that’ — to automatically e-mail or archive their Tweets or their Favourites found that the platform’s programming interface no longer allowed them to do this. Similarly, it has restricted the functionality of third-party interfaces for the platform. Dash (2013) describes these limitations of functionality as “bullshit turf battles” with platform operators privileging the goals and needs their own platform ahead of their users. Actions like these from platform operators are explicitly user-hostile as they impinge on a user’s agency over the data they create or use on the platform. This is because of the perception of value on their platform. Interoperability advantages users, but it also advantages a platform’s competitors. To Twitter, tweets are the reason that users use the service – whether they are reading them, writing them, or saving them for later.

Yet even if the ways in which operators control platform affordances influence users to leave, the design of platforms can still manage to privilege the powerful. As such, a number of platforms have a number of strategies in place that serve to undermine the potential and the ability of users to leave. In Facebook’s case, users have noted the persuasive strategies attempted to keep dissatisfied users from leaving. Kessler (2013) chronicled the strategies — from the pragmatic warning “you won’t be able to reactivate your account or retrieve anything you’ve added”, to the needy question “why are you leaving” automatically answered with suggested actions users can take instead of leaving, to the emotionally compelling “your n friends will no longer be able to keep in touch with you”, along with pictures of friends’ photos and suggestions that “[friend’s name] will miss you”. Finally, if users manage to resist the platform’s attempts to persuade them to stay, they will likely find that they have left

something. Data. Both the data created by users (posts, likes, photos, social graph relationships) and the data harvested by the platform (profiling, analytics data, observed information) stays with the network. There is a limited amount of portability afforded to users through Facebook's Download Your Information (DYI) tool — which provides a printout of their profile/timeline page and post history data only, in a relatively poorly-coded and not particularly interoperable Hypertext Mark-up Language (HTML) file. The richest quality of data is kept after users leave by the platform operators — as it is still a potentially valuable commodity for their purposes. Many users who have requested their accounts be deleted have failed to achieve any such result from their contact with platform operators. One user, Nipon Das, who succeeded after several months of communication and legal threats towards Facebook remarked in an interview with the New York Times that “It's like the ‘Hotel California’ — You can check out any time you like, but you can never leave” (Das in Aspan, 2008).

Intriguingly, a person may not even need to ‘check in’ for a platform to start collecting data on past and prospective users. A recent security bug discovered by independent security researchers (Blue, 2013a) resulted in Facebook disclosing a practice in which it attempted to create profiles of users who were not yet using the service, and who were not providing all of their data to the service by associating data that it had observed or been supplied through other users — for example, data collected using the tracking web cookies required by the Facebook ‘like’ buttons that popular websites used, or data collected from users that allowed the platform to scan or crawl their email account or address book to find their other friends on the platform — with a supposed profile invisible to users, notionally termed a ‘Shadow Profile’. Facebook's

Policy Communications department told reporters “that concern about collection, storage and shadow profiling of contact data is the sole fault of users who failed to read (or remember) the Facebook policies they agreed to when they were getting started on Facebook” (Facebook Policy Communications in Blue, 2013a). Blue (2013b) noted the backlash from users and commentators cautioned people from using the platform, but these responses are of little help to the non-users whom the platform has identified. As non-participants, people whose data and contact information is volunteered by their friends and families are given no opportunity to opt-out of inclusion in the platform’s data-mongering. To paraphrase Blue, from the platform’s perspective the choice to participate is irrelevant.

The existence of more clandestine methods of data collection on users has been alleged against platforms like Facebook for some time, particularly by aggrieved users like then Austrian Law student Max Schrems. Schrems is one of the 10 Facebook user-activists with an activist group called ‘Europe Versus Facebook’ agitating for policy change on the platform through European legal processes filed with the Irish Data Protection Commissioner against Facebook’s Dublin-based offices. They allege twenty-two complaints about Facebook’s practices — including the aforementioned ‘Shadow Profiles’, non-consensual data gathering, non-deletion of the user data that users’ mark for deletion, unclear languages and terms of privacy policies, non-consensual inclusion in facial recognition software, the use of tracking cookies and the absence of a guarantee to securely protect user-related data (Europe Versus Facebook, n.d.).

Platform operators, it must be said, do not exclusively hold power in negotiating these conflicts with users. Both the platform operators and the users of a platform have strategies and tactics of final resort — a form of ultimate power over each other. Platform operators if faced with a user base they no longer desire can close their platform, to shut it entirely and close or change their model of business. Similarly, users can leave the platform altogether, closing their accounts to neuter their interaction with the platform and their value to the operators. These are both ‘nuclear’ options, so to speak, and we can accordingly draw some parallels between this power tension and John von Neumann’s famously well-acronymed doctrine of military strategy ‘Mutually Assured Destruction’ (MAD). The prospect of user-less social networking platforms is obviously not as earth-rending, devastating or depressing as the Cold War’s prospect of global thermonuclear war — but the point of the analogy is to note that the tension between users and operators contributes to a platform’s continued existence.

And yet, in spite of the great power that platform operators and platform users have in conflict with each other, there seem to be no notable examples of platforms closed or abandoned as a result of political conflict between users and operators. There are no missiles fired in anger, so to speak. Platforms have failed — but as yet, these have not been overtly political conflicts over policy. Users have left platforms they formerly used for newer ones that better suited their needs and desires. Garcia et al. (Garcia et al., 2013 1-2) note that whilst the growth of online social networking platforms is well documented and well analysed by researchers, the decline of online communities and online social networking platforms is less analysed, leaving open questions of “the role of the social network in keeping user engagement, or in the

spreading of user dissatisfaction” and the types of “network structures that lead to higher risks of massive user departures”. Their research, unfortunately, did not answer either of these questions and instead analysed and quantified the resilience of interconnected user communities using graph theory to determine the ability of communities using platforms to withstand changes, focusing on the effects of changes but not the causes. This is interesting, but it only tells a fraction of the story — telling how users stayed on or how users left their chosen platforms, but not why they stayed on or why they left (Garcia, Mavrodiev, & Schweitzer, 2013). There are some analogies worth exploring in relation to the interdependence of platform operator and users. The most notable examples are of ‘traditional’ media, where media producers depend on the continued presence of readers, listeners and/or viewers as an audience, and the audience depends on the producers to produce texts to suit their desires. These relationships are interdependent because of the ways in which media producers attune their creation to react and adapt to their audience (Ball-Rokeach & DeFleur, 1976 4-7), and the ways in which the audience is influenced by and reacts to the media they choose to consume. When this interdependence is interrupted by a perception of non-reactiveness and thus non-dependence from audience or producer, the impetus for producers to keep producing or consumers to keep consuming those media content becomes questionable.

Of course, the infrastructure and systems of Silicon Valley almost certainly play a role here — a key part of the development of many popular social platforms is the involvement of venture capitalists as stakeholders, either through direct funding and investment rounds, or through the involvement of ‘incubators’ like Y Combinator. These circumstances work to essentially preclude platforms from taking the ‘nuclear

option' of shutdown in conflict with users, as financial constraints (such as expected returns on stakeholder investment) become a leading concern in the continuing existence of the platform — meaning that in general, platform operators will only terminate the operation of their platform once it is no longer financially or legally tenable.

Given that they only exist in a digital, seemingly ephemeral, form, platforms can only stake a tenuous claim towards their existence. Then, when that claim's validity is lost — the users or the operators leave — what happens? We can walk through the ruins of ancient cities and find the artefacts of technologies of decades and centuries past. How will we fossilize Facebook, and what will the ruins of Twitter's be? Where, for that matter, are the ruins of the platforms that came before them? We can consider MySpace and Geocities as examples. MySpace, in its earlier incarnation, was a burgeoning social networking website — before Facebook or Twitter became household names — it sold to Rupert Murdoch's News Corporation before fading into obscurity as users left to join other social networks. News Corp sold it again, years later and at massive loss to marketing company Specific Media and pop-music star Justin Timberlake (Halliday, 2011). The digital artefacts of MySpace's history were purged from web access as it updated in June 2013 with a new design that focused on music and photography. 'Classic Users' were contacted in September and given the ability to download some (but not all) of the data they had created over that time (Ask Myspace, 2013). MySpace's operators had perceived a commercial impetus to 'reset' the platform and create a fresh, new design with new a new focus on multimedia, in part because the platform had been in decline, but what is not particularly well understood or identified is 'why' MySpace was in decline. Common knowledge

suggests that once Facebook came to prominence users left MySpace and joined Facebook. This is true — but a very simplistic perspective. Industry analysts have since insisted a number of causes: ‘anonymous’ users (Chan in Lee, 2013); a lack of innovation (Arango, 2011); a lack of utility (M. Jones, 2011); bad design (Parker and Fallon in Tsotsis, 2011); a lack of ‘visionary’ leaders (Goulston, 2011); Rupert Murdoch (Taibi, 2013); Microsoft’s ASP.net server programming framework (Scoble, 2011); and so on, but these causes are almost invariably imparted as anecdotal wisdom, and without qualifying evidence.

Referencing her own interviews with teens using social platforms, boyd (2009) argued that teenage users had in many cases determined their network uses either as a matter of choices relating to taste considerations or perceived prestige, and other teenage users chose platforms by following their pre-existing ‘analogue social network’ of friends. boyd’s research underlines the likelihood that user choices made in relation to joining platforms are not necessarily always platform-dependent. It follows logically that the decisions made by users to leave a platform may not be contingent on that platform either, but this is less well evidenced by research. Again, this is emblematic of Garcia et al.’s (2013 1-2) complaint that the documentation and research of platforms occurs more during their growth than during their decline.

Geocities provides a different chain of events. Beginning as a free web-host that provided its users with a (then generous) 15 Megabytes of storage space, Geocities was an internet juggernaut in the late 1990’s. After users moved on to new and different services or hobbies, Geocities was eventually acquired by Yahoo! but continued to flounder with a lack of interest from existing users and prospective users

alike. Yahoo! eventually shut down Geocities, but a group of users known as ‘ArchiveTeam’ sought to preserve it. As they described it, Geocities was both “the largest self-created folk-art collection in the history of the world” (Scott in Fletcher, 2009) and “for millions of people the first experience dealing with the low-cost, full-color, world-accessible website and all the possibilities this contained” (ArchiveTeam, 2009). The Archive Team eventually released a torrent of the Archive, for those interested in preserving Geocities to share and mirror multiple copies of the archive. The Ruins of Geocities are no less ephemeral in their digital form but have been preserved by users long since the interest of the platform operators was lost. This is interesting, as it illustrates the formation of a power vacuum in platform operation, and the response that users and former users can take to fill that vacuum. Though users could not (or did not want to) continue operating the platform, they saw value in its preservation.

A broader lack of concern for platform preservation and web archaeology might contribute to some of the disconnect between the operators of platforms and the expectations of their users. Dash argues that today’s platform giants operate without a well-informed historical context (Dash in Funnell, 2013), that platforms like Facebook, Twitter and Google Plus have epochal perspectives that perceive their platforms of social networking as fundamentally new and without precursor. Some conflict, particularly in Dash’s perspective, stems from the development of these platforms with two notions in mind: that complexity is harmful to platform growth, and that extreme control over users maximises profitability. According to Dash, these were debunked earlier in the ‘social web’ / ‘web 2.0’ era, with the failings of earlier platforms like Friendster and Myspace. His criticisms bear some resemblance to the

oft-paraphrased reflection of philosopher George Santayana (1980); “When change is absolute there remains no being to improve and no direction is set for possible improvement: and when experience is not retained, as among savages, infancy is perpetual. Those who cannot remember the past are condemned to repeat it”.

The Challenges of Archives, and Learning from our Past

Our understanding of conflicts between users and platform operators of this is hampered , by a lack of archival preservation by former operators and users. As boyd (2006) explains, “social technologies succeed when they fit into the social lives and practices of those who engage with the technology”. This turns out to be an ongoing, two-way bargain between a platform’s users and a platform’s operators. The bargain requires the platform to serve the needs or desires of a user, and in turn the user needs to provide value — either as prospective revenue or in payment — to the platform. The neglect of this bargain is detrimental to the continued use of a platform and the impetus for its operation, but further research must strive to add detail and complexity to our understanding of these processes of platform neglect and failure. The history and failures of older platforms can provide context and insight into the contemporary challenges faced by platform operators and users alike.

Archives present a necessary avenue for the preservation of knowledge — not just for institutions, but for the activities and politics of everyday life. Unfortunately, archival systems managed without institutional systems like user- or citizen-led projects can often be of limited quality and resources — and must struggle with many

problematic issues regarding ‘what’ is archived, ‘who’ makes archival decisions, and ‘who’ has custodial responsibilities for the information preserved.

Perhaps one of the most gruelling, unpleasant, but fundamentally important area for researchers to focus on in research of histories of violence, aggression and hostile online spaces is the task of archival and recording work. As Consalvo (2012) argues “Despite the seeming persistence of online documents and artifacts, much is disappearing from the internet or becoming increasingly harder to find [...] it can still be difficult to locate or identify materials from even six months ago”. This is particularly relevant to this case, as well as any research that reckons with online platforms and networks. Even during the course of writing up this doctoral research, dozens of primary sources cited have ceased to exist — some removed due to users being banned or removed from platforms, others no longer exist for financial or legal reasons (notionally archival social storytelling platform Storify was shut down in 2018, and Archive.is (formerly Archive.today) purged its historical archives in 2017 because users had invoked its archive functionality to preserve child abuse material), and others like Snapchat, 4Chan and other *chan-style image-boards and message-boards are ephemeral by design.

Whilst it’s perhaps tempting or easy to simply characterise abusive, belligerent and anti-feminist as simply as they appear, Consalvo argues that understanding the logics promulgating online abuse may yet be an important factor in combatting and reforming these spaces in the future — but this will require the labour of researchers performing this essential work in digital spaces that have expressed outright hostility to their work and presence. As I highlighted in an earlier chapter — this is a fraught

task, not only because of the hostility towards researchers and the mental health toll that regular, continuous exposure to abuse and harassment can take. There's no easy solution here — the work Consalvo urges is depressing, dangerous, and essential.

That said, there are many further risks associated with the archival process. Firstly, the tension between availability for research (especially as 'open-access' data and publications continue to grow in popularity) and availability to abusers — for better and worse, research and archival projects provided under Open Access arrangements lower the barrier for entry to research and academic knowledge for those seeking it. Whilst we might often presume this to be a universal positive — it's important to consider the potential for bad faith interaction and usage of archival data and publicly-available research. This includes the potential for notionally anonymous sources to be re-identified, or for non-sources to be mis-identified because of vague similarities (or even the wildly conspiratorial 'logic' of a harasser). Researchers must also contend with the politics of archival itself — who will 'own' the archives, who is a custodian of the archives, who can contribute or correct the data within? Maintaining archival data presents its own array of difficult governance challenges for researchers to engage with.

Following on from this, it's important that as researchers, we don't simply archive for the sake of archival. Researchers should look to archival sources like those Consalvo encourages, the cases documented like those in this thesis with a view to shaping and influencing reforms that address the issues that users are concerned about — especially where these issues relate to the safety and freedom of expression of people who have been harassed, abused, and otherwise marginalised.

Finally, researchers should consider the effects of archival data on the subjects recorded. Quinn (2015), whose case is investigated in much closer detail in the third case study of this thesis, speaks of her own experiences grappling with hostile, antagonistic ‘trolling’ behaviours in her youth, and her journey towards ‘growing up’ and eschewing these behaviours in later life. While it’s important to catalogue key issues and actors in cases like these, it’s also important to weigh the influences that archival and attention may have for prospects of reform and personal growth — will further attention catalyse a change in views, or see a hardening and calcification of prejudices? There is no one-size-fits-all solution to this issue. The prospect of personal reform and interpersonal forgiveness is particularly complex — and warrants significant thought from any who delve into this practice of conflict archival.

Adjacent Industries, Adjacent Research

Whilst poor archival practices within industry and a lack of industry interest in historical learnings from platformed communications presents issues like those outlined above, there is also the potential for important, relevant research and findings to be siloed within different and adjacent industries and fields of research. As such, interdisciplinary approaches help to draw parallels and uncover useful theories and findings with which to explore platform issues. In particular, I draw attention to the fields of computer game development and ‘game studies’.

Digital ‘social media’ platforms, much like computer games, have the ability to “quantify, measure and compare” users. The design of games, characterised by Humphreys (2005, 51) as “cybernetic feedback loops”, can allow similar structures of surveillance, assessment, performance and reward as social networking platforms — where we might consider likes, hearts, favourites and shares as a different kind of ‘score’, and a user’s social graph as a different type of ‘multiplayer’. As such, looking at the design of systems for policing and shaping behaviour in computer games, provides analogues for these systems in digital social platforms like Facebook or Twitter. Just as a Player-vs-Player arena might constrain the duel of players in an online Role-Playing Game, the Privacy settings of a user’s post might similarly constrain the bounds of an argument to a subgroup of the user’s social graph.

Whilst the design of these digital spaces can favour certain behaviours or impose subliminal rules, it’s worth noting, as Humphreys (2015) does, that “not all rules can be coded” — given this, both spaces of social platforms and spaces of multiplayer games have layer of external, user-facing policy to their governance — ‘laws of the digital land’ that can, like laws in traditional governmental contexts, sometimes be capricious, arbitrary, ineptly deployed, or overzealously enforced. Where these laws ostensibly differ —perhaps mildly— to traditional governmental laws is the axis at which civic goals (like the provision of social spaces and congregations) give way to commercial imperatives (like subscription costs, or advertiser interests).

Given this, we can note that although spaces like Facebook, Twitter, Google+ and other social platforms might appear to be new, the issues they face are often not new at all. Researchers should look, as I have in this thesis, to draw interdisciplinary

connections across fields of research to highlight the histories and tensions of platform issues across the history of computer-mediated communication.

Looking to Games Studies, I want to highlight ‘Co-creation’, as explored by Banks and Humphreys (2008, 401-18), which provides a useful, but imperfect framework for analysis of governance issues in these spaces. In particular, platform affordances serve as the liminal, hybridised zone between the networks of users and the creative potential of users. The relationship between ‘gamers’ and ‘developers’ is shaped, in part, by the desires and the efforts of each participant. In the field of online games, co-creation practices like those explored by Banks (2010) have facilitated messy but beneficial negotiations regarding governance, community management, as well as notions of labour and ownership in gaming platforms and gaming spaces.

These are parallel relationships — as developers have come to terms with the investment and involvement of users in their gaming space, they recognise the extent to which these spaces have been co-created by their players. Whilst social platforms don’t acknowledge or engage with their users in the co-creative manner Banks has illustrated, it is evident from their continued existence that users have, and continue to, co-create spaces and works on social platforms.

That being said, it bears reflection that many gaming platforms and spaces have significantly smaller communities than popular digital social platforms like Facebook, Instagram, Twitter and the like. What has worked for the smaller, perhaps

comparatively like-minded or like-interested communities of gaming spaces, may not necessarily map as clearly to larger, more complicated networks like Facebook.

Users and Tactics

The Economics of Ownership and Platform-Intermediated Data

If we are considering social web platforms like Facebook, Twitter and Google Plus in the conceptual frame of de Certeau's theories then we must consider that platforms, like the cities and workplaces which de Certeau used as exemplars, are not just political spaces but public spaces as well. Who then, are the owners of platforms? In a legal sense — the owners are the shareholders of the publicly traded companies that run platforms (Facebook, Twitter, Google), the creators or in some cases venture-capitalists that have funded the platform's development. In cases where users directly pay for use of a platform (as on microblogging platform 'App.net', for example) they are considered customers, not owners. These are relatively typical understandings of ownership, where the owner(s) have the control of a platform because they are most directly responsible for both the costs and the operation. This is subsequently obscured somewhat by the language that the operators of the platforms use in communicating with users. On Facebook, parts of the platform are distinctly acknowledged as 'yours'— 'your post', 'your timeline', 'your activity', 'your likes'. Facebook's Statement of Rights and Responsibilities (Facebook, 2013b) grants users 'rights' and demands 'responsibilities' from users and even disavows ownership of 'your data', an action echoed in the terms of their Data Use Policy explicitly states that "while you are allowing us to use the information we receive about you, you always own all of your information" (Facebook, n.d.).

What a user owns is a little ambiguous, so it is useful to try disambiguating what can be owned on a platform. There seem to be two key areas of ownership in this context — the ownership of the data and the ownership of the platform itself. It is also worth disambiguating ‘data’ into a few key types, as the concept is all-encompassing, and as a result somewhat vague. We can consider ‘user-contributed data’ as data that a user knowingly and willingly supplies to a platform with the intent of its reproduction on the platform (tweets, posts, photos, tags and so on). We can consider ‘user-facing data’ as data that is not necessarily created by a user, but data that purports to relate a user and is displayed to the user as their data — it is the platform infrastructure that most immediately surrounds user-contributed data (a user’s timeline, a user’s activity log, a user’s privacy settings). Finally, we can consider ‘user-related data’ as the data that is collected and created about a user, their data, and what they do on the platform that is gathered through the analytics and surveillance functions of platforms and typically not available to the user. ‘User-related data’ bears similarity to ‘metadata’ and would ordinarily be an adequate descriptor for the type of data I describe here — but I distinguish between the two as ‘metadata’ has been particularly framed by the stakeholders of post-Snowden and post-WikiLeaks surveillance debates to narrowly focus on certain types of user-related data. For an example of this — consider the relatively “vague” and “nontechnical” (Stilgherrian, 2012) definition of metadata tabled in an Australian Senate Estimates committee hearing by the Australian Federal Police (Ludlam, 2012): “Information that allows a communication to occur” and “Information about the parties to the communications” relating to “telephones - fixed and mobile” and “Internet”. Australian politicians have simplistically referred to metadata as “essentially billing data” and “a contestable concept” (Abbott and Brandis in Duckett, 2013). As a result, the conventionally useful term ‘metadata’ has become

problematic through its politicisation at the hands of “ignorant” politicians, their partisan point-scoring and “boilerplate cybercrime-is-scary squawk[s]” (Stilgherrian, 2013).

The ownership of data on platforms is a well-travelled area of debate. Data is both the end-product and the by-product of labour on a platform — both the labour of users and the labour of the platform operators. Both Hegel (2012 73) and Locke (Vaughn, 1978 311) agreed that ownership of property was conferred by the output of labour — and this view seems no less applicable to the ownership of data. Both platform operators and users could stake claims to the ownership of different sorts of data by virtue of their labour. The question of ‘who owns what fragments of data and why?’ is fraught with the potential for dispute — which perhaps explains why platforms are hesitant to claim ownership of any user-facing or user-contributed data. As noted earlier, platform operators make distinct and overt attempts to imply ownership of user-facing and user-related data: “*your* timeline”, “*your* posts”, “*your* Tweets”, “*your* favourites” and the like. This is also expounded upon in their policies. Instead of making any claim to ownership of user-contributed and user-facing data, most platform operators will explicitly deny claim to the copyright to these kinds of data, instead declaring that user-contributed and user-facing data belongs to the user and that the user’s contribution of data or participation in the platform to facilitate user-facing data creation grants the platform operators impressively broad licenses: Twitter stipulates that users grant it a “a worldwide, non-exclusive, royalty-free license (with the right to sublicense)”, whilst Facebook’s conditions are similarly broad and allow it to declare ownership of any data that users no longer associate with on the platform — which gives them potential claim to the data that users have marked to

delete from view, as well as the data that users leave behind when they leave the platform (O. Smith, 2013). However, it seems that if these policies are explained at all to users, they are done so quite poorly as evidenced by the user reactions to the policy changes that alter the terms of the platform operator's license. In recent memory Facebook (Walters, 2009), Twitter (Newman, 2009) and Instagram (McCullagh, 2012) have all had relatively well-known conflicts where users have reacted negatively to the news of copyright and license policy changes — including a reported exodus from Instagram (Brustein, 2012), and the copy-pasting of notionally legally-worded (but practically ineffectual) objections to the use of their copyrighted materials as Facebook posts (Read, 2012). This is perhaps part because users of the platforms see themselves as stakeholders in both a participatory sense and an economic sense. In either sense the umbrage at change is, according to Ford (2011), derived of a sense of entitlement to consultation. As economic owners or participatory stakeholders in the data of a platform, users are often directly concerned by the ways in which their data is used. In many cases, users will feel they have the agency over their data because the platform provides affordances for managing data — creation, deletion and distribution. When platform operators change the terms on what users can control, it is a threat to the user's perception of agency over 'their data'.

After a late-2013 controversy about Facebook's moderation policies in the United Kingdom, a political commentator, Hugo Rifkind (2013), argued in an opinion column that “we carry these companies around in our pockets all day and stare at them on screens, and we begin insidiously to think that they are ours. [...] These are our new public spaces and they are not, in fact, public spaces at all.” This is pessimistic and needlessly dichotomic, but it also shows a reasonably prevalent perception of

platform ownership. The ownership of platforms is a debate less-often had. To users, academics, observers and operators alike, the common perception is that the platforms are owned by the companies that operate them. Like Rifkind's contribution this conservative view of ownership is reductive. On user-pays platforms like App.net or Vimeo, the usage habits of users can be considered as a form of occupancy within the platform, especially where the platform operators have used possessive language to imply a user's ownership over parts of the platform. Although adverse possession rights (or 'squatter's rights', as they are colloquially known) might not be entirely applicable to a currently operated platform like Facebook, they seem particularly applicable to operator-abandoned platforms like Geocities. The Archivers who attempted to preserve Geocities only attempted to capture user-facing data, but they could also have staked a claim to ownership and operation of the platform had they tried to run it on their own after Yahoo! shut it down, particularly if they had been long-term users and occupants of that platform space.

A similar practice can be seen in certain Massively-Multiplayer Online Role-Playing Game communities. The cult-hit online video game 'Star Wars Galaxies' was operated by Sony Online Entertainment (SOE) for eight years, between 2003 and 2011 when SOE shut down the online servers for the game (perhaps due to diminishing interest and profits, or changes to the Star Wars licensing). Not content to let a beloved game and digital space vanish into the ether, hundreds of fans and players continue to play Star Wars Galaxies through 'SWGEMU', a home-brewed server software that emulates the server functions provided by SOE, five years after the cessation of service (Alexandra, 2016). Similarly, an effort called 'Nostalrius' (L. Johnson, 2016) has been launched by players of Blizzard Entertainment's long-running "World of Warcraft"

(WoW) game, keen to revisit the game as it was twelve years ago in what players term as ‘Vanilla’ or ‘Classic’ WoW.

Through efforts like the Geocities Archiving, SWGEMU and Nostalrius, we can see that players and users of these digital spaces sometimes have stronger relationships with online spaces than might have been anticipated by the creators and operators of these spaces. They may even seek a custodial duty over the space if they feel the space is being poorly served (or not served at all) by the operators and creators of the space.

Platform ownership and the specific questions of ‘who owns what’ on the social networking platforms is a particularly complex area of study and practice. This is partly because platforms are often complex systems. Systems that, whether using computing power supplied and directed by platform operators or enticing production by promising entertainment, utility and the rewards of ‘affective labour’ (Myerscough, 2013 25) to users, serve to obscure the labour that produces the all-important data that creates value in the system. Our understanding of ownership on platforms is hamstrung by both the complexity of systems and the shallowness and historicism of contemporary economic debate in this field. To the extent that this debate is happening, it is often by nature abstracted from practical context by theorists like Richard Stallman (2010) and Jaron Lanier (2013) — both of whom seem more excited to focus on utopian notions of ownership or warnings of an information superhighway to serfdom than on any practical and pragmatic contributions to economic discussion of property, labour and ownership in the age of platforms. The conflicts of these areas are largely conflicts over what is currently understood as simple property rights — and more

insightful and robust debate could serve to help users and platforms alike to better understand the owners, the labour and the economics of their efforts online.

Conflicts between Users

For most of the conflicts highlighted earlier in this chapter, the tensions of governance are decidedly vertical — that is, they are hierarchal conflicts of power, between the owners and operators of the platform space, and the users of the platform. These are akin to the conflicts you might expect in a government or a workforce, when new laws or policies are introduced that citizens or workers find onerous or disagreeable. This, of course, is not the only kind of conflict that can exist in these spaces. Just as citizens can come into conflict with each other for social or political grievances, or workers can come into conflict with each other over union membership or pay rates, users can come into conflict with each other over their uses or expectations of internet platforms. This is exemplified, time and time again, through the arguments and ‘flame wars’ associated with platforms like chat rooms, bulletin boards and blog comment threads. These horizontal — user-user — conflicts serve as a reminder that the policies in these spaces serve not only to regulate the interactions of users with platforms, but also of users with other users. In this manner, the metaphor of platforms as ‘states’ is particularly appropriate, as laws in civil society also serve to both regulate citizens interactions with the state (through taxes, for example), as well

as with each other (through laws relating to work, to civil disputes and to criminal harms, for example).

In most user-user conflicts, there still remain tensions with the governing bodies of the digital spaces these conflicts occur in. Users can and do appeal to the wielders of power in digital spaces as the arbiters of rules in conflicts. These conflicts can be sparked by a range of diverse issues, such as conflicts about political ideologically, or about everyday politics. Some conflicts are decidedly hostile — with participants on one or more sides acting with a particular belligerence. Belligerent conflicts on internet platforms have become a significant area of research interest, given their elevated coverage in news and mainstream media. Belligerent conflicts have been perhaps too-readily categorised as garden-variety ‘trolling’ or bullying behaviours and dismissed with the trite and ineffectual “don’t pay attention to them” advice that elders so often give to childhood victims of bullying. This attitude both discounts the complex political motivations that drive some belligerent behaviours and dismisses the belligerent behaviours as a kind of norm that should be expected in these spaces.

Users engaged in belligerent, adversarial behaviour often constitute their own behaviour as something of a noble pursuit — the pursuit of ‘rationality’ over what belligerent users see as “feminized and othered” (Phillips 2015, 126) ways of thinking that they deem to be inferior. This is particularly exemplified by notorious and self-identified ‘troll’ Andrew ‘weev’ Auernheimer, who argued in an interview with the Huffington Post, that the methods of ‘trolling’ and other belligerent inter-user abuse tactics were, Socratic by nature (Auernheimer in Phillips 2015, 127). As Phillips (2015, 127) notes, Auernheimer’s assertion is particularly focused around *Meno*, as

the rhetorical onslaught that Socrates focuses towards his instructor is particularly focused on ‘winning’ their argument.

It is perhaps useful here to make a distinction between generalised, ‘drive-by’ harassment efforts, and systemic ‘organised’ harassment efforts. Whilst users who participate in drive-by harassments might simply be responding to a gut-instinct dislike of a person or their ideas, the repeat offenders who persist in abuse and harassment efforts may seem to have the additional motivation of an underlying of justifying ideology. This is particularly exemplified in *Down Girl*, in which Manne (2017) interrogates the logics and rationalisations presented in misogynistic justifications for misogynistic actions.

Gender is often a key factor in these hostile conflicts, especially from belligerent actors who construct their behaviour as a reactionary response to the growing popularity of modern feminist movements (Broomfield, 2016; Brown, 2015), but it is not the only factor. The ‘Kiwifarms’ bulletin board platform, for example, hosts users recording and discussing their antics against what Kiwifarms users describe as ‘lolcows’ — callous term used loosely to refer to people who are socially awkward, people with autism, people with learning disabilities, and people who can be manipulated for the purposes of their harasser’s amusement (Todd, 2016).

Gender in Platform Spaces

A particularly notable trend in the course of this research has been the prevalence of conflicts that have been prompted by gender-related issues and the regularity in which women have become both the targets and the most prominent activists in platform policy conflicts. With this in mind, we should consider the idea that platforms have gendered biases in design and operation, or that platforms have provided functionality that serves to empower existing gender biases in society as well as the people who hold those biases. Whilst the underpinning ideologies might be consistent with conventional, offline misogyny, the methods with which sexist behaviours can be promulgated have changed through computer-mediation. “The Internet made misogyny routine and sexual bullying easy” argued British writer Laurie Penny (Penny, 2013a) in a book that explored web-enabled sexism as “the disinhibition provided by time-delay and anonymity” on platforms that empowers misogynistic internet users.

There are two key ways in which misogyny can affect women using social platforms like Facebook and Twitter: from the platform; and from other users. As broad concepts user-platform and user-user conflicts have both been explored earlier in this document, but they also warrant examination in this specific context of sexism. Platform-based sexism is arguably the most obvious, and the most wide-spread manifestation of sexist practices. We see it in the policies of more conservative platforms like Facebook, where images of nudity are notionally verboten. Of course, there is a gulf of ambiguity in attempting to discern which kinds of nudity are forbidden in policy and in practice. In theory, Facebook bans pornographic and sexually explicit content, but in practice they have censored publicly-posted images of breastfeeding

mothers, breast-cancer survivors and acts of political expression, whilst ignoring privately-posted content in private groups or user messages.

This dissonance is partly understandable as a matter of resource management and pragmatic enforcement — it is both easier and more pressing for platform operators to police public-facing content than to investigate more clandestine content. According to the team responsible for adjudicating these reports, Facebook’s “hate and harassment team” receives approximately two million reports of hateful content or harassment weekly but the team remains concerned that the most deserving and offensive content goes largely un-reported by users (MacKinnon, 2013 154). Nonetheless, the team must make moral and political judgements about what constitutes obscenity on the platform. The platform’s policies are firstly a method of understanding the expectations of the platform operators, but they may also be reflection of broader societal expectations of femininity. Of course, those broader expectations can be challenged or upheld by platforms operators and their policies: the fact that Twitter does not police expressions of sexuality or nudity in the way Facebook does underscores the political decision that platforms can take to police gender norms.

Inter-user sexism, on the other hand, is the manifestation of a user or group of users’ sexism against another user or group of users. At its most *bien pensant* it manifests as a sort of conflict of opinion between users on comparatively equal footing, but at its most concerning it features users familiar with the affordances of a platform flouting the platform’s policies to make threats or hateful comments. For sufficiently motivated users, the obscuration of their identity to assume a functionally anonymous guise — even on notionally ‘real name’ (anthroponymous) platforms like Facebook

— is relatively trivial. Registering a free or disposable e-mail address, devising a pseudonym or username and password, and accessing platforms through proxy servers or anonymising services like TOR to obscure their original IP address serves to remove the consequences of disobeying a platform’s rules. After all, why would a user care if a deliberately disposable account is banned from the service? The ways in which motivated users can subvert the notion of punishment provide a clear example of how users can use the affordances of a platform to express themselves in ways not condoned but not easily policed by the platform operators. The catch-22 of platforms enabling users’ political expression is that the more freedom and convenience is afforded to users, the more difficult those users are to manage and police.

According to a report from advocacy group Working to Halt Online Abuse (2012) female internet users are significantly more likely to report online abuse. This is explained best by their increased likelihood to be selected as victims — as researchers have found that in online chat-rooms “accounts with feminine usernames incurred an average of 100 sexually explicit or threatening messages a day [, while] Masculine names received 3.7” (Hess, 2014). Given the prominence of sexist behaviour on social networking platforms and the internet more broadly, there is a lot of discussion amongst users about how they would like to see platforms respond to inter-user sexism. These vary, from encouragements of censorship by way of “private enterprise or government intervention” (West, 2014), through to declarations that “private companies that aim to be the new public sphere should not be in the business of censorship” (York, 2014).

Subcultural Harassment with Platformed Spaces

Although online gendered harassment is perhaps most often associated with Twitter in news and current affairs reporting, it has been a recurring issue in online communities and platforms more broadly, and especially in gaming and anonymous spaces. It tends to manifest as a sort of in-group/out-group bias against women and minorities who are either seen as meddling interlopers, or who are perceived as capitalising on the social capital of facets of their identity (such as race or gender) rather than subsuming themselves to the norms associated with identity groups like ‘gamer’ or ‘anonymous’ (Cross, 2015) (in relation to 4chan-style imageboards). This is particularly evident in the ‘Gamergate’ campaign, which I analyse in detail with a case study in this thesis.

The incidences of harassment, abuse and intimidation in online spaces, that are isolated for the purposes of case study research later in this thesis are indicative of greater issues of sexist behaviour and beliefs, both in subcultural spaces and in society at large. Shannon Sun-Higginson, the documentary filmmaker behind ‘GTFO’¹ — a chronicle of misogyny and harassment in the gaming industry, worries that “the general public will focus too much on Gamergate and say, ‘Look at this crazy thing that happened,’ [...] It was a terrible, terrible thing, but it’s actually symptomatic of a wider, cultural, systemic problem.”

¹ GTFO is an often-used acronym for the phrase “Get The Fuck Out.”

Whilst we know that men and women play games in comparatively similar number, the field of gaming and the cultural spaces relating to games have been perceived as a stereotypically masculine space since the emergence of video games as a popular medium in the 1990's (Behm-Morawitz 2014). Indeed, these spaces have become somewhat well known as a 'macho' or braggadocios masculine spaces — known well for “trash talking” and aggressive language and mannerisms — in-part due to the behaviours of the dominant personalities and players in these spaces. Whilst there is a spectrum of behaviour that can be associated with the 'macho' discourse in these spaces — Fox and Tang (2016) focus on analysing hostile, and gender-adversarial (misogynistic) behaviour, and particularly look to anonymity and the deindividuation it provides as a key element of hostile behaviours. This is a particularly useful distinction, as it is a distinct reasoning from many other criticisms of anonymity in these spaces.

Many critics of anonymity in digital spaces describe it as providing disinhibition — thus letting anonymous users more freely express themselves without the inhibitions provided by culpability and authorial presence. This notion of disinhibition takes a particularly cynical view of all users — positioning belligerence and bad behaviour as something of a base motivation that can only be guarded against with social constructs like persistent identity. Fox and Tang's discussion of deindividuation is a useful alternative approach, highlighting instead the key element of anonymity that is leveraged by belligerent users. Deindividuation allows users to blend into a crowd and avoid culpability — this is a key element of empowerment that a belligerent user might seek from anonymity in their social interactions, as it leads to a perceived freedom from consequence.

As Fox and Tang (2016) highlight, a regular element of behaviour in gaming spaces is adversarial belligerence — whether through overt sexual harassment, ‘trash talk’, or various practices of ‘trolling’. A corollary effect of these behaviours is reciprocal withdrawal by the targets of these behaviours. Fox and Tang (ibid.) outline that these behaviours (and especially, severe manifestations of belligerence) can have an effect of policing participation by discouraging participation from users who will ruminate on unpleasant experiences and form unhappy and unenjoyable associations with the spaces and practices associated with their experience.

This leads into what Noelle-Neumann (1974 44) terms as the “Spiral of Silence”, a phenomenon in which people who might otherwise express dissenting or critical perspectives choose to remain silent for fear of isolation, neglect, or even hostility from people who hold a dominant opinion and some degree of influence or ‘power’ within that space. This is particularly notable in relation to the dual-purpose advice-admonishment “don’t feed the trolls!” often given to people who have experienced some degree of harassment for expressing their views. Whilst “don’t feed the trolls” is a token of advice often earnestly given, it unfortunately serves to affirm the spiral of silence, and often condemns the people who follow this advice to suffer in silence, without discussing or bringing attention to their harassment, let alone their perspective.

Nakamura (2013) offers a counter-argument to popular and academic discourse surrounding ‘trolling’, and the sexist, racist, and otherwise belligerent behaviour associated with it, asking “What if, in the spirit of media archaeology, we understood

online racism not as a glitch but as part of the signal?”. Nakamura (ibid) illustrates the behaviour of people using racist language and behaviour by using their own justifications and arguments as context. Whilst belligerent users do not necessarily see their abusive behaviour in the same light as the people they target (often minimising themselves as ‘annoyances’, rather than abusers) these belligerent users are aware of their behaviour alongside platform affordances, and the participatory nature of platforms. Given this, Nakamura argues the essential need to move beyond simple perceptions of ‘trolling’ and abusive behaviours as an unexpected ‘glitch’, — like noise accompanying a radio signal — and treat the behaviour as the signal itself — more of the ‘content’ of the platform.

Shaw (2014) builds on Nakamura’s thesis, outlining that abuse is not a failure of the “utopic view of online communication”, rather that it is simply another common element of that communication — and that like all racism and sexism, their online expressions are similarly emergent from the same positions of relative privilege. This is the case both in general online communication, as well as subcultural spaces like gaming culture. To borrow Shaw’s apt summation: “People are jerks not only when they are in anonymous Internet spaces, but also when they are in spaces where they can get away with being jerks.” (Ibid.)

This problem manifests in physical spaces (like conventions and tournaments), in digital spaces (like the chat-enabled lobbies of competitive multiplayer video games) and is enabled by a range of stakeholders that have either endorsed these behaviours or given them support in turning a blind eye. A key example of this is *Cross Assault*. *Cross Assault* was a 2012 reality show promoting the upcoming *Street Fighter*

x *Tekken* video game. The show was co-produced by Capcom, the Japanese development studio behind the well-known *Street Fighter* series of video games, and IGN a popular video games and popular culture website, and featured teams of professional gamers representing ‘Street Fighter’ and ‘Tekken’ teams who competed for a prize of USD\$25,000. The show was broadcast live, 10 hours a day for a week through the ‘Twitch’ video streaming platform and was published in edited ‘webisode’ form on the IGN website.

During the event Aris Bakhtanians, the coach of the ‘Tekken’ team, to focus his attentions on team-member Miranda ‘Super Yan’ Pakozdi. This begins on the first day of the tournament and is not addressed by the production teams or by other players until the fifth day after Pakozdi decides to forfeit a match.

Pakozdi: *I’m trying to play Aris, you’re messing me up.*

Bakhtanians: *You need to be able to focus when people are heckling you!*

Pakozdi: *That’s fine, but like... this is just creepy.*

Bakhtanians: *You need to be able to play when people are harassing you.*

Pakozdi: *Thanks for that, Aris.*

Bakhtanians: *Take off your shirt.*

(Ibid.)

Games culture scholar Todd Harper noted that whilst Bakhtanian’s behaviour was reprehensible, it was also reflective of the whims of the Twitch streaming

audience, and that he responded to their prompts when pestering Pakozdi (Ibid.). On the fifth day of the tournament, Twitch's community manager, Jared Rea interviewed Bakhtanians, asking "Can I get my Street Fighter without sexual harassment?", to which Bakhtanians responded "You can't. You can't because they're one and the same thing. This is a community that's, you know, 15 or 20 years old, and the sexual harassment is part of a culture, and if you remove that from the fighting game community, it's not the fighting game community." (Klepek, 2012)

Though Cross-Assault shows a particularly confronting example of harassment, it's important to note that this style of harassment is not exclusive to the gaming subculture and is not necessarily representative of the full gamut of abusive, harassing and intimidating tactics that have been wielded by users in hostile user-user conflicts. The tactics deployed by belligerent users in other conflicts range from simple insults, to the publication of private documents and information, to the creation of hoax reports to police agencies in effort to use an armed-response team to assault a target by proxy. Threats have been made — of bombs, of rape, of murder, and of grievous harm.

In 2007, well-known game developer, programming instructor and author, Kathy Sierra leaves the cancelled her public speaking engagements at a conference held by O'Reilly publishing. "I have cancelled all speaking engagements. I am afraid to leave my yard, I will never feel the same. I will never be the same" (Walsh, 2007). For weeks, visitors to her website had called an 'open season' taunting, harassing, and threatening her. Eventually, Andrew 'weev' Auernheimer, a notorious "troll" (Schwartz, 2008), is said to have circulated a fabricated narrative of her career that falsified a history of sex work and domestic violence, as well as her home address and

social security number (Sandoval, 2013). This was understandably alarming — Now all of the people who had threatened to rape and kill Sierra had access to her home address.

Sierra left the tech blogosphere soon after bringing her harassment to public attention, rejecting the culture that enabled it: “I do not want to be part of a culture — the Blogosphere— where this is considered acceptable. Where the price for being a blogger is kevlar-coated skin and daughters who are tough enough to not have their ‘widdy biddy sensibilities offended’ when they see their own mother Photoshopped into nothing more than an objectified sexual orifice, possibly suffocated as part of some sexual fetish” (Mantilla, 2015 109).

These user-user conflicts are not necessarily fundamentally concerned with the platforms on which they occur — in practicality, user-user conflicts are often multi-platform affairs. The targets of abuse and harassment are often followed and targeted across multiple platforms, and those who wield belligerent and abusive tactics against them seem to do with an understanding of the affordances of the platforms they use — a site on the open web, like a self-hosted website for example, offers very little in the way of oversight or governance, at the expense of hosting costs, platforms like Twitter —where users do not opt-in to social contact— are more susceptible to direct harassment tactics, platforms that seem to automate moderation processes also present opportunities for harassers to subvert the power wielded by platform operators, and use it against their targets. Certain platforms provide different affordances and different systems of governance that can be beneficial to harassers, which makes the

platforms and both their technological or policy frameworks important to analyse in the context of ‘horizontal’ user-user conflict.

Whilst it might seem tempting to simplify these styles of conflict to simple ‘horizontal’ user-user conflicts, eschewing the platformed dimensions of internet-enabled harassment risks losing a particularly valuable perspective on how users can tactically wield power (albeit belligerently against each other) thanks to the design of platformed spaces. Massanari (2015) writes of the dominance of toxic technocultures in a handful of sociotechnical networks including Reddit, 4Chan, Twitter and online gaming platforms. Gamergate, which is analysed in further detail later in this thesis, is a key example of toxic technocultures in action — a public that coalesced around a particular issue or event, like networked and affective publics of boyd (2011) and Papacharissi’s (2015). Where toxic technocultures contrast however because of the deployment tactics used within these cultures often rely heavily on implicit or explicit harassment of others.

Platforms like 4Chan, and Reddit have provided formative spaces for behaviours in part because of the flexible affordances these platforms offer to users, and these affordances (including anonymity, and/or the easy creation of new accounts, as well as account-less access) can be easily used for abusive behaviour. Indeed, chief amongst distinguishing features for the participants in toxic technocultures, are the technical proficiency with online platforms, the penchant exploiting platform policies that often value aggregating large audiences, and a value for both the networked labour and plausible deniability that large communities provide. In turn, certain platforms become tactical spaces for belligerent users.

Practices of Protest

Whilst conflict can often take the form of belligerence and displays of inter-user aggression, it's also important to distinguish between belligerent acts, and acts of civil disobedience, wherein behaviours that may well be seen as hostile, aggressive, or otherwise 'not typical' are leveraged in pursuit of change or reform from an authoritative power. The philosopher Bertrand Russell made a case for civil disobedience of laws in a magazine essay in early 1961. Shortly before his arrest (for participation in Nuclear Disarmament protests) later that year, he wrote in the *New Statesman* magazine arguing that two kinds of conscientious civil disobedience exist: The first, is "the disobedience to a law specifically commanding an action which some people profoundly believe to be wicked" — such as the conscientious objectors to war efforts; and the second, an employment of civil disobedience "with a view to causing a change in the law or public policy" (Russell, 1961). The latter was the focus of his essay, which argued against a widespread perception that dissenting individuals were impotent in conflicts with government and ineffective in influencing change. Russell argued that this form of conscientious law-breaking, driven by an urge for the reform of public policy is a powerful form of communication to governments that seek to represent their constituents. Russell's tactical proposition of disobedience in pursuit of reform still has a good deal of relevance in contemporary activism, especially in the context of user/platform conflict where platforms fulfil the role of a notionally-liberal, though not democratic, government and users represent the citizen-dissidents. In platform/user conflicts, user tactics often require a form of activism that flouts a rule

or policy of the platform — often with the express intention of having that rule or policy changed.

The ways that users have responded to the sexism both of platform operators and of the other users has been to demand change from the platform operators to the platform's policies and practices. In the case of sexism enacted by platform operators, users have campaigned for the platform's policy reform to change the practices of platform operators. The platform operators are also the primary focus of campaigning in the case of user-based sexism, because the operators are recognised as arbiters of conflicts governed by platform policies. In both cases of operator- and user-based sexism, platform operators are recognised as political figures by virtue of their ability to create and change policy, and their power to enforce it.

The nature of the protests made by users against platforms can be particularly revealing — for some users, changing the platform through participation and conscientious citizenship is an important part of their activism. This bears a great similarity to Bertrand Russell's (1961) notion of civil disobedience by breaking the rules that they were politically protesting. Off-platform protesters often fielded similar arguments, but by escaping the jurisdiction of the platform they escaped the punishments that platforms can impose — like bans and censorship — in a similar way to political dissidents leaving authoritarian countries to criticise tyrannical governments from the comparative safety of another country's jurisdiction. In many cases, off-platform protests go to another platform: whether it is a competing social platform, or a blogging platform like WordPress or LiveJournal, to make political commentary against the platform where it has no influence or ability to police. Cherian

George noted a type of ‘off-platform’ protest in the context of Singaporean activism as online and publicly accessible platforms became the benefactors of counter-hegemonic alternative media (George, 2013 389). This allowed “the Internet”, as George nebulously and inclusively described it, to be the ‘off-platform’ staging ground to the ‘platform’ of authoritarian Singapore. Much as “the Internet” put Singaporean dissidents ‘out of reach’ of their government (Ibid.), alternative publishing platforms like WordPress and competing social networking platforms can provide staging grounds that are ‘out of reach’ of the platform operators who are being protested against. In the case of vertical, user-platform conflict, staging grounds provide an important opportunity to attempt to speak and affect change outside the boundaries of the platform’s jurisdiction. For better or worse, this is also a key factor in the use of multiple platforms in horizontal user-user conflicts. Users who spread their hostile, belligerent behaviours across a number of platforms present a challenge for the policing and enforcement of platform-specific policies regarding abuse, harassment and user safety.

Because of this, our scope must be broad in looking at the tactics of users. Some will use tactics, to flout policies on-platform with the intention of entrapment and exposition of policies for their perceived short-fallings. Other users can wield influence outside of the platforms, by trading on high-profile positions or the exposure of mainstream media coverage to deliver a political argument to a wider audience. On- and Off-platform tactics each have different potential for effects — but both are important in understanding the ways in which users attempt to affect change in these types of political conflicts.

Platform Governance

Governance & Constitutionalism

As social networking platforms are developed, and grow with their users, the use of a contractual approach, like a ‘Terms of Service’ document to govern a user’s relationship with a platform can become particularly challenging as the limits of an old-fashioned, conservative legal approach become apparent to users. A common reaction to criticism of social networks often portrays the issues of governance as a failing of users to understand the contractual circumstances set out in the terms of service — a document of governance in which users have very little stake and very little say (Wilson in Centre for Independent Studies, 2011). In practice, this perspective offers a simplistic, deferential and defeatist attitude towards conflict with platforms — a ‘love it or leave it’ approach wherein users either abandon platforms in disgust, or stay users, and implicitly consent and concede conflicts.

Posing governance as a strictly contractual issue is reductive, as social networking platforms present issues in a semi-private/semi-public platform environment that is significantly more complex the scenarios typically served by contractual agreements and laws. For example, a consumer-style contract offers little understanding of the networked dimensions of Facebook. Whilst a consumer could choose to abandon a brand or a service provider by taking their custom elsewhere, they can often still substitute the brand or service with alternatives that are essentially interchangeable and interoperable (changing bank accounts, for example does not tend to necessitate a change of currency — nor does changing laundromats tend to

necessitate changes of linens) still play the sport with their friends or loved-ones in a different shoe — Facebook, as a networked service, has a significant lock-in component, that is not portable for users: their colleagues and families. If a user decides to leave Facebook, they cannot easily take their social connections with them. It is challenging to take ‘your data’ — that is either the data you have created and have legal ownership of, or the data created in relation to your usage of the platform (at time of writing there has been no legal precedent set as to the ownership of user metadata on social networking platforms) — away from the platform. Facebook and other social platforms are ill-suited to the product-consumer metaphor, as you cannot substitute Facebook for other brands or services in the way that you could substitute conventional consumer brands. Whilst clothing or foodstuffs are easily enough replaced or discarded, Platforms like Facebook often work exhaustively to ‘lock-in’ users to their digital ecosystem, and attempt to make the transferral of user data to other (non-subsidiary) services arduous and unpleasant (Constantine, 2013). Moreover, as platforms like Facebook and Twitter grow in social importance, the external social and cultural, and professional pressures to use these platforms play an increasing role in user’s decision-making. These are effects and influences of the network, which has become intertwined with the social and professional relationships of its users (Light, 2014 123). In ‘Code 2.0’ Lessig (2006 289) noted that the costs of action and sovereignty were more significant in ‘real space’, as opposed to the costs in ‘cyberspace’, remarking that a decision to leave a massively multiplayer online game in response to conflict was akin to choosing a competing product in a competitive market. This would reportedly minimise the exertion of power by ‘cyberspace’ governors, by making them compete for citizens. This would make digital

governments similar to Coca-Cola and McDonalds, Lessig (Ibid.) argued, “eager to please, fearful of revolt.”

These decisions about participation are complex, because the decisions to participate — and in what way to participate are complex. We can and should argue that use and inhabitation of the digital space governed by digital platforms is not an implicit provision of consent to their governance (or lack thereof). Whilst common amongst technologists, these notions of a passive, implied, and uninformed consent are not solely used by platform operators.

As Suzor (2018) notes, this is something of a fraught area where “the disconnect between the social values at stake and the hard, legal realities” (Ibid.) becomes a challenge for platforms and users alike to navigate. On most platforms, documents like the ‘Terms of Service’ or ‘Terms of Use’ arrangements cast the relationship between users and platform operators as a relatively simple commercial endeavour (Ibid.) — in which users are allowed access to the platform and its affordances, under the assumption that they will behave in a prescribed manner (or, will not misbehave in a proscribed manner).

These ‘Terms of Service’ are something of a pragmatic non-solution to the issue of governance, serving primarily to legally affirm the power of platform operators and espouse a contractual expectation of the users on the platform. Given that, as argued earlier, platforms are digital spaces that intermediate a range of inter user social and civic communications, ‘Terms of Service’ systems represent a woefully inadequate

framework for governance, as they “are almost universally designed to maximize [Platform] discretionary power and minimize their accountability” (Suzor 2018).

Platform Operators give ‘Terms of Service’ a power simply by virtue of their deployment as an incontestable set of rules — whereas a meaningful system for governance would be afforded legitimacy by participation, interaction, and ultimately, consensus with its stakeholders (Black, 2008)— in this case, the users of each platform. As Black (2008, 144) notes, “Legitimacy may be an objective fact, but it is also socially constructed”. To this end, the idea that a platform’s ‘Terms of Service’ can have meaningful legitimacy means that platform operators must be perceived both as having a right to govern the platform, and as have the support of those it intends to govern on behalf of (Ibid.).

The abject failure of contractual systems like ‘Terms of Service’ to provide a meaningful framework for legitimate governance, in turn provides ample possibilities for future reforms that might more meaningfully engage with users and platform operators in a more democratic or collaborative approach to negotiating these digital spaces. As Suzor (2018) argues, “This is an opportunity to set out the constitutional principles that we collectively believe ought to underpin our shared social spaces in the digital age”.

The complexity of these environments is something platform operators seem ready to admit, as Zuckerberg noted in 2009;

“Our terms aren't just a document that protect our rights; it's the governing document for how the service is used by everyone across the world. Given its

importance, we need to make sure the terms reflect the principles and values of the people using the service.”

This view is a radical departure from the way that law views both the governance of platforms and their terms of use. In analysing governance in digital platforms, Suzor (2010 66) provides an understanding of “digital constitutionalism” and proposes a “rule of law” framework as a method of informing and understanding governance in digital spaces that responds to the difficult application of contractual doctrine in these spaces.

Platforming Power Relationships

Whilst this research project is influenced by Certeau’s framing of Strategies and Tactics — and uses this framing as a lens for this analysis, it bears discussion that this framing of platform environments is, in part, reductive. It is an attempt to frame a new and complex social, political, technological and cultural environment through a familiar lens and familiar metaphors. Platforms pose challenges that are similar to those posed by other unequal political relationships. In this manner, we will see that many aspects of the conflicts interrogated in this research project have tensions that are uncommon in the conceptualisation of strategies and tactics that is encapsulated in the metaphoric pedestrians and cities or workers and factories proposed in de Certeau’s theory.

The political relationship between platform and user has similarities and differences from many of our typical theoretical dichotomies, between city and pedestrian, factory and worker, state and citizen, producer and consumer — and while all of these are partially applicable and can provide insightful lenses for analysis — none of these metaphors for power relationships map perfectly to the complexities of the relationships between users and platforms. As both Humphreys (2005) and Suzor (2010) convey, this complex relationship is not a fixture solely of social networking platforms — the changing environments of video games, especially networked, multiplayer video games encounter similar issues in understanding the relationship between developer (as moderator and governor) and the player (as labourer in a digital space).

These complications reveal the tensions implicit in user-user and user-platform relationships and highlight the need for theory to further conceptualise these relationships, both with reference to the technological and political relationships of power struggles past, and an eye to the way platformed power relationships are becoming a fixture of everyday life.

Whilst the relationship of player and online game developer bears major similarities to the relationship of user and social media platform operator, it's also worth reflecting on this user/platform relationship in light of other digital intermediation platforms — including those that eschew 'social networking' for intermediations of labour or property — Uber, and Airbnb for example.

At face value, Uber's "ride-sharing" platform might seem like a simple crowd-based replacement for taxi services, but the Uber platform and its operators are actually providing an intricate and multifaceted system which harnesses the labour, capital, and reputations of its users and intermediates their relationship for commercial gain. Both Uber and Airbnb have been resistant to external regulation, and has encouraged practices of legal nullification to subvert laws that they do not care to abide by (Pasquale & Vaidhyanathan, 2016).

Regulation and Justice

Law sometimes takes a reactive stance towards new technologies (Suzor, 2010 69-71) — often informed by the immediacy of struggles posed by privately designed technologies in contemporary discourse. Indeed, as of the writing of this thesis, governments around the world are beginning to grapple with platforms and their complexities under newly-designed laws that have been constructed in reaction to the tensions and conflicts of social networking platforms. A recent example of this is the bevy of legislation aimed at so-called 'Revenge Porn' (or, to use a less sensationalist and more accurate description, the non-consensual publication of intimate photographs) passed at local, state and national level of a handful of Western nations (Jeong 2013). The spectre of platform regulation by conventional governmental actors is both a concern and an impending certainty for popular social networking platforms — especially those in which the failures of platforms self-governance have become evident to users and to conventional political actors. Strictly speaking, platforms have been subject and compliant to external government regulation for quite some time — as Gillespie (2017) and Jeong (2015) demonstrate — and are subject to a range of

laws governing things relating to copyright infringement, hate-speech, taxation and the like. The kind of regulatory change that is would most concerning for platform operators would be a closer examination of their role as a ‘platform’. In many legal jurisdictions, platforms like Facebook are treated similarly to telecommunications providers — as ‘common carriers’. The similarities exist but the comparison is tenuous, as the role of platforms is significantly more comprehensive and complex than that of phone and internet service providers. This categorisation, as a ‘common carrier’ allows some degree of a hands-off approach to governance, as it (legally) absolves the platform operators of some responsibility for the actions of users that are published and propagated on their platforms — so long as they act reasonably when informed to quell, censor, or remove illegal content.

A compelling aspect of platform-based and platform-related governance is that it can seek ameliorate to the issues faced by users situated in wildly disparate locations and governed by wildly divergent laws. As Lastowka (2010 83-90) notes, issues of legal jurisdiction become complicated, when a platform, and the users of a platform are each subject to different laws. If the platform’s governance can ameliorate the challenges in seeking inter-jurisdictional justice for users, the specifics of the laws become a less pressing concern for users, but where this fails, the failure of laws to provide frameworks for inter-jurisdictional disputes and conflict resolution becomes immediately apparent.

To the extent that technologists seek regulation, they tend to seek it to suit their own ends (Morozov, 2013) — indeed, they often seek to define the very terms by which they are regulated. Consider, for example, discussions around the ‘crypto-

currency' Bitcoin and the prospect of its regulation. As Golumbia (2016) notes, the positioning of Bitcoin as a reaction to government regulation of the internet implies an “expansive cyberlibertarianism”, which encompasses a belief that freedom will inherently emerge from digital technology, and that regulation (particularly by government) is therefore antithetical to freedom.

This argument is disingenuous and obscures the reality of regulation. The prospect of Bitcoin and alternative/crypto-currency does not inherently provide for an unregulated currency, it rather changes only the relationship to the regulator (Ibid.). Whilst traditional currencies are typically regulated in relation to governments and central banks — crypto-currencies like Bitcoin instead enshrine regulatory power in the hands of the software engineers who designed the currency. This is a clear example of what Ullman (1995) describes as the cult of the engineer — presumptions, assumptions, politics and a highly specific manner of doing things are inscribed in the code — whether of social platforms or of crypto-currencies. These systems inevitably seek to “reproduces and re-enact life as engineers know it: alone, out of time, disdainful of anyone far from the machine”.

I noted earlier in this chapter the use by technologists of what Morozov (2013) and Alder (1999) term the ‘Technocratic Pose’ — a rhetorical flourish that attempts to obscure the power, roles, and influences of technology in society. It has become clear from the rhetoric of platform operators, that whilst social and technological platforms clearly have a present and growing role in the social and political lives of users, the prospect of possible changes to the regulatory environment for platforms entices operators to conceptualise their roles and influences in ways that may minimising of

their roles or the expectations of their responsibilities. This is likely borne part of the ‘safe harbor’-style protections that platform operators have been afforded, and their legal treatment as a platform being considered more akin to a telephone service provider than a newspaper or content publisher.

The systems and affordances of governance on social networking platforms are numerous and complex. These complexities underscore the public good that these platforms can serve, and the public desire for these platforms to serve users well. Policymakers and lawmakers must be prepared respond to these complexities. These systems are clearly of importance to the daily lives of people around the world and eschewing these complexities in favour of self-regulation would leave these important systems to be governed by the whims and social mores of the technologists involved in their operation. For laws and external regulation to serve users and platform operators justly, these laws must seek to understand the relationship between platform operator and user not as one of producer and consumer, or service provider and service user — but recognise and negotiate with the labour of users, and the governance tensions of platform operators.

The Role of this Research Project

This research is an interrogation of power and conflict on social media platforms, by becoming familiar with the rationalisations of strategic and tactical power and the political issues that affect a platform’s operators and users alike. The social media platforms at the centre of these conflicts are political spaces, shaped by policy and

enforcement and inhabited by users whose participation and expression on the platform can be reflective of political values. The conflicts are often multifaceted and include a range of horizontal (inter-user) and vertical (user-platform) political tensions.

These spaces are governed by platform operators who have commercial imperatives and political desires to operate their platforms in a certain way without significantly alienating their users. This requires operators to wield their power strategically — to influence through design and social norm, rather than shows of force. Conversely, users tend to have less power in the platform environment they use — and wield their power through tactically ambitious subversions of policy, regulation and governmentality in attempts to affect change on the platforms they use. The growth of availability, the increase in political discourse and the stakes of potential regulation in this field warrant the investigation into social media platforms as political arenas as well as the conflicts that shape their use and governance.

This research provides evidence for the ways in which the governance strategies of platform operators are deployed, and the ways in which users tactically respond to or subvert these strategies when they come into conflict with the platform, or each other. This evidence demonstrates the importance of detailed direct observation and analysis in developing understanding of the operations of social networking platforms and the ways in which people use them.

Chapter 3: Research Design

This thesis investigates practices of governance and subversion on online social platforms. As an area of research, the study of social platforms and their governance structures is relatively new. This makes case study an ideal methodology for exploratory research — as the cases are constructed from key events, but remain attentive to inferences that can be drawn from similar events outside the scope of formal investigation (Gerring, 2004 353). As I noted in the introduction, this thesis uses Dunleavy’s ‘compromise model’ — which is ideal for humanities and social science research, as it allows for a focused literature review, rather than the expansive literature review of Dunleavy’s ‘focus down’ model, and allows for a more contextualised discussion and analysis of research findings than Dunleavy’s ‘opening out’ model (Dunleavy, 2003).

Research Methodology

This research applies case study methodology, informed by Stake (1995) and Yin (1994), to analyse user and platform-operator conflict on social networking platforms. As a methodology, case studies represent a disciplined and qualitative form of inquiry into the events studied. Given the relative youth of social networking platforms, the use of case study provides an opportunity for deep research and a relatively holistic analysis of the phenomena observed. This research explores in detail several examples of platform-related conflict to provide a foundational and extensively considered understanding of these cases to guide future research and researchers in this field. As Stake (1995, 2, 7, 17) argues, each case is “a specific, a complex,

functioning thing” and by seeking greater understandings of each case we can move towards observing the problems, conflicts and complexities of human concern with each case — perhaps to the extent of refining informed generalisations from observation of the cases. This research largely involved desk-based research practices — including close readings of the policies, rules, formal proclamations and actions of platform operators and an analysis of the development of these policies, rules and strategies of governance as they have developed over time. These analyses are considered alongside readings of user-contributions and analysis of the tactics employed by users in their conflicts with platform operators, including activism on- and off-platform: in mainstream media and on other, competitor-operated platforms. This research is supplemented by semi-structured interviews conducted with stakeholders in the studied conflicts — typically users or close observers who have had experience in the conflict. Finally, the conflicts are contextualised in a political and social history — by tracing popular and academic literature and reporting that references similar political conflicts on other platforms or in broader public discourse. By analysing both the strategies of platform operators and the tactics of users this research provides a reasonably holistic view of the conflicts that it analyses in the case studies chosen.

Case Selection

The key platforms focused on in this research were amongst the most prominent at the outset of the research project: Facebook, Twitter and Google Plus were each amongst the most used platforms in both the English and non-English use worldwide. They also own and/or operate very popular subsidiary social networks like Instagram,

Vine and Youtube respectively. These are ideal candidates for examination because of their widespread usage, their prominence in modern popular culture and their enormous user bases. Of course, this is not to imply that the conflicts on smaller platforms do not warrant attention, but rather a reflection of where the most thorough research can be done by virtue of the greater amounts of participants in the space.

The limited sample size of only a few case studies limits the extent to which any of the cases could be argued as typical or representative of other cases on the field as a whole. Because of this limitation, case studies have been selected not on the basis of representativeness or universality-of-application, but rather by focusing on what learning opportunities are afforded by the cases (Stake, 1995 4). In choosing the three conflicts analysed in the following case studies, I sought conflicts that involve group of users and the operators of a platforms and disputes over the policies or governance of these online spaces. Where possible the case studies were chosen to exemplify certain types of conflicts: in particular, conflicts were chosen that relate to a dissonance between the political ideologies of platform operators and users; conflicts relating to policies implemented by platform operators; and conflicts relating to the dissonance of expectations of either platform operator or user and their expected actions — for example the policing of inter-user conflict, or the disobedience of platform rules.

Case Studies Selected

Three cases were chosen for this research, each observing a different political conflict on a different social network.

The first case study analyses the conflicts relating to Facebook's policies regarding photographs of mothers breastfeeding and mastectomy scars. This case analyses a period from 1st of July 2007 to 31st of July 2016, wherein there was significant conflict between users of the Facebook platform and the platform's changing policy towards the publication of photos of mothers breastfeeding, or of mastectomy scars.

The second case study analyses Google+'s policies regarding user identities, and in particular, their so-called 'Real Names' policy, which mandated that users display their identity using their 'real' or 'legal' name, and not use the system pseudonymously. This case analyses a period from the beginning of Google+'s private beta (in which some members of the public were invited to test Google+ before a widespread launch), 01st of June, 2011 to 15th of July, 2014, during which time there was a significant conflict between Google+ users and Google+ platform operators as to the ways in which users could present their identity on the platform. Google formally abandoned the 'real name' policy on the 15th of July 2014 — this case includes analysis of a short time after this point to facilitate observation of the consequences of this change.

The third, and final, case study analyses Twitter's responses and policies relating to interpersonal abuse and harassment on their social networking platform. This is widely understood to be a major problem for Twitter. In particular, this case focuses on the misogynistic harassment of female users of Twitter in relation to a campaign to feature Jane Austen on a banknote in the United Kingdom, and additionally 'Gamergate' — the campaign surrounding harassment and abuse of women and

gender-diverse people, people of colour, and people of diverse sexual orientations and identities in subcultural video gaming spaces. This case analyses a period from 01st of June 2013 to 31st of July 2016. The termination of this case is particularly problematic to define — as although the inter-user conflict (harassment and abuse) ended relatively early for the banknote-related campaign, ‘Gamergate’-associated harassment and abuse continued for months-leading into years, and the conflict between user and platform has continued beyond the bounds of this research. Indeed, many developments relating to this case study have arisen after the end of the research period and during the time in which this thesis was written.

Case definition and boundaries.

In selecting cases and scoping cases for this research, it is necessary to define boundaries for each case. In general, the opening boundary — the beginning of the case — is not particularly easy to define. Although we could loosely define a case’s beginning by describing the first instance of a case’s conflict on a social platform, this could ignore key precursory events — earlier conflicts that would influence policy or user perspectives. The opening boundary of the cases studied in this research is, by necessity, imprecise — to allow for retrospective analysis of the precursors to conflict in each case study. Defining the closing boundary - the ending of the case - is similarly complex. Whilst some cases may have seemingly clear and delineated endings, like the public abandonment of a controversial policy, many of these conflicts continue in different ways long beyond the initial conflict as platform strategies and user tactics change over time. Cases are long enough to accommodate years of conflict but must have a functional end to their scope to help analysis. For cases where endings have not

been significantly found, I use the ending boundary of the **31st of July 2016** as a functional ending to facilitate analysis. That said, it is worthwhile contextualising the case studies themselves with the broad context of these conflicts on internet platforms, and as such, I follow-up on each of the cases with addenda in the conclusion of this thesis.

Research Process

This research has been created in order to develop better understandings of; structures of governance in online spaces, dissident users and their online tactics, and the interplay of platform operator and dissident users in platform-intermediated conflict. As such, this research has been designed to answer three key research questions;

1. What strategies do platforms use to influence and control user behaviours on their platform and the development of their governance?
2. What tactics do users deploy when in political conflict with a social media platform?
3. How do these strategies and tactics influence the interplay between users and platforms in political conflict?

Research Question #1: What strategies do platforms use to influence and control user behaviours on their platform and the development of their governance?

This research question is used to attempt to explore online governance conflicts from the perspective of the governmental actors — the platform operators. Unfortunately, it has been challenging to engage platform operators in interviews for this research. Accordingly, I have compensated by focusing on close readings of media statements, policies and announcements by platform operators to provide operator perspectives as best as can be afforded in this circumstance.

The primary research for this question — interviews and close readings — is supplemented by the analysis of the historical development of the conflict, and its contextual relevance in broader social discourse or reoccurrence in other platform jurisdictions. Beyond governmental policies and rationalisations, this research question has been answered with a particular focus on the methods of policing and enforcement utilised by platform operators to achieve their governance over their platform.

Research Question #2: What tactics do users deploy when in political conflict with a social media platform, or other users on a social media platform?

In political conflict, tactics are by nature a method of action that is by nature reactionary — an attempt at political subversion or rebellion from a position of comparative weakness. In this, tactics are often at their most effective when they have not been predicted or understood by platform operators. Answering this research question is as a result, a slightly more complex affair than Research Question 1. Whilst

analysing conflicts, I attempt to identify key activists — the users whose tactics were either notable or effective. By using semi-structured interviews, I intend to elicit perceptions of conflict from the involved users, as well as rationalisations for their choice of tactics. These interviews are supplemented by attempts to faithfully reproduce a timeline of key events in each case of user/operator conflict.

Research Question #3: How do these strategies and tactics influence the interplay between users and platforms in political conflict?

Understanding the interplay of strategies and tactics is the focus of the third research question, and expectably, it is best answered after analysis of each case and reflection on both preceding research questions. By reflecting on the development of policies and policing explained in response to Research Question 1 and the timelines established in response to Research Question 2 a clearer insight is gained into the way that certain strategies elicit subversion or rebellion, and the ways in which these inspire the actions of users. Similarly, by analysing the development of policies after a user/operator conflict necessitates the introduction of tactics, we can better understand how, if at all, this affects the organisation of power and governance on social networking platforms like Facebook, Twitter and Google Plus.

Details on Methods

Potential Interview subjects were discovered through early desk-based research, with a particular focus on engaging participants who were involved in the conflicts studied. Participants were approached either through the social networks they used or

email, if they had a publicly listed email address. Failing that, if initially-sought participants were unavailable to participate, or did not respond to requests to participate, I attempted to identify participants who would likely have had similar experiences.

For the Facebook case study, this involved seeking out participants of Facebook Groups dedicated to breastfeeding, or anti-breastfeeding censorship. For Google+ and Twitter case studies, this involved finding users who had difficulty either using their identity on Google+ or in facing harassment on Twitter.

Desk Based Research

For each case study, a process of desk-based research was undertaken to source documents for close reading and contextual analysis. I used a loose ‘snowball sampling process’ to begin sourcing data, and following the data founded to refine and adapt data gathering methods. These methods included web search tools (e.g. Google), and on-platform search functionality through the websites and associated applications of Facebook, Twitter and Google +. In each case study, search terms were used appropriate to the subject matter of each case and refined based on the ensuing data gathered.

For the Facebook case study, search terms started with “Facebook” “breastfeeding” “mastectomy” “censorship”. From this research, more terms were

developed, including terms that investigated feminist campaigns relating to the case, and the temporary bans that users reported receiving during this conflict.

In the Google+ identity study, search terms began with “google plus” “real name” “identity” and “nymwars”, a common way of referring to the conflict. As with the previous case, these terms developed over time — and were influenced by developments in research from my analysis or from the interviews I conducted with conflict participants.

Finally, in the Twitter case study, search terms began with “Twitter”, “abuse”, and “harassment”, and were quickly expanded based on initial findings to include more pointed searches about types of harassment “doxxing”, “bomb”, “rape”, “threat”. On-platform searches also the platform’s API to search public ‘mentions’ — tweets sent at victims — for misogynistic slurs or terms that were likely to be interpreted as threatening or extremely hostile.

To follow up-and contextualise each case’s key conflict, I also use desk-based research to investigate the social, technical and political environments of each of the conflicts researched, so that I can contextualise the conflicts researched, where relevant, with broader conflicts, trends and issues on other platforms or in similar conflicts.

Whilst this desk-based research is not a comprehensive bibliometric analysis, it has served to provide a strong contextual base for the primary research of this project

— the interviews with participants, and the close reading of conflict-related documents and publications.

Interview Process

Interviewing conflict stakeholders allows for this research to examine and analyse more diverse perspectives in each case study than just the researchers. As such, it is an invaluable part of this project's research methods.

I approached potential research participants with the intent of conducting at least five interviews for each case study. Participants were selected for their position in relation to each case's conflict. Because of this, I attempted to engage with users and platform operators for a range of perspectives on each conflict. In seeking users, I sought users who had personal experience of the conflict, as well as users who wrote about, documented, or spoke publicly about the conflict. Platform operators were also sought, but interviews could not be obtained with representatives for the social platforms analysed in this research.

Interview Questions

The questions were developed in order to facilitate discussion around the participant's expression of their views or experiences in each conflict. A general set of Interview questions guided each interview.

To begin with, participants were asked some general questions about their usage of social media platforms. These questions were decided upon to help provide insight into the user experiences, especially as those experiences may be divergent from other users, or even my own experiences as a researcher.

Generalised questions for *users* included:

- What attracts you to **[social media platform]**? Why do you spend time there?
- What do you do on **[the platform]**? How do you use it?
- How much time do you spend on **[the platform]**?
- Do you interact mostly with other users that you have already known, or met online?
- Do you use any other social media platforms?

Following on from these questions, I asked questions relating to the conflict I was researching for the case related to the participant. These questions were designed to be sensitive to the participant's experiences. Follow up questions prompted by the participant's discussion were similarly sensitive to the personal nature of the types of conflict discussed, particularly as conflicts related to the participant's identity, freedom of expression, or safety.

In general, conflict-related questions for users included;

- What can you tell me about **[platform]**'s initial reaction to **[your content / your identity / your concerns]**?

- What was the outcome of this conflict?
- Were you expecting this outcome?
- How did you react?
- Did you document your interactions with [the platform]? If so, why?

Though some of these questions could have been answered with reference to reporting or published documents relating to the conflict, this research and interview presented an opportunity for participants to express their experience in their own words, rather than having their experiences explained from the interpretative standpoint of a reporter or researcher (Charmaz, 2014 73).

Finally, I asked participants to reflect on the structures of power that influence their uses of platforms. With these questions, I sought to invite users to consider the complexities and political dimensions of the conflict they were involved with. These questions facilitated discussions of users desires and understandings of governance and government in these spaces.

The questions asked included;

- Why do you think that **[platform]** took this action against **[you / your content]**?
- Have you noticed any change in the policies of **[platform]** since your conflict with them? If so, have they addressed your concerns?
- How did you feel about that?
- How did others respond?

- Have others that you know experienced this?
- Has this experience changed the way you use **[platform]**?
- Has this experience altered your perception of **[platform]**?
- Would you (still) recommend **[platform]** to friends or associates?
- Are you considering leaving (or have you left) **[platform]**?

Similarly, questions were created for interviews with the operators of social networking platforms. Some basic questions were formulated to contextualise the participant's answers with their role(s) at the platform;

- What is your role at [social media platform organisation]?
- What does this role entail? (What do you do?)
- Describe a typical interaction with a platform user.
- How are platform policies formulated?

In focusing the discussion on the platforms' perspective and handling of conflict, especially at an organisational level, questions were designed to discern how policies are formulated at an organisation level;

- How have platform policies been designed in relation to **[area of conflict]**?
- Is this policy development consultative? If so, is consultation also sought outside of your organisation?
- How does **[the organisation]** react to users dissenting against policies?

Following this discussion, platform operators were to be asked practical questions about the methods of policing for the platform's policies. These questions were designed in order to reveal information not often made public about the operations of these systems.

Questions about policing included;

- How does **[platform]** police this policy? Are algorithmic tools used, or perhaps reports from users?
- Has user feedback or backlash resulted in a notable change to any of **[platform]**'s policies?

As is typical for semi-structured interviews, these questions formed a basic structure for open-ended discussion with research participants. For users, the questions were designed to obtain some foundational details about their use of online social platforms, to gain some insight into their experience in platform-related conflict, and to invite them to critically reflect on that conflict and the role and effect of platforms in online governance. For platform operators, the questions were designed to gain a fundamental understanding of their role as an operator, to gain some understanding of policy development, and to gain insight into the structures and responses of policing users on online platforms.

Interview Difficulties

Unfortunately, interviews were quite challenging to organise, and much fewer were recorded than anticipated. There are a number of issues that contribute to this,

including technological challenges with contacting participants, as well participant availability. I did not anticipate how social networking platforms like Facebook, deploy simple but effective anti-spam countermeasures that create additional challenges for researchers attempting to contact potential research participants that they have no pre-existing social connection to. This could technologically be ameliorated by sending ‘friend requests’ rather than ‘messages’, however this would grant the researcher access to a significantly larger cache of data about the participant’s use of the platform than would typically be associated with interview data collection. I considered this approach but determined that I was uncomfortable with any attempt to make this kind of intrusion into the personal affairs of any research participants. Future researchers seeking to use this approach should note that gaining research data through access to a research participant’s semi-private social networking profiles is a significant responsibility that should be acknowledged in Human Research Ethics proceedings for any prospective research projects.

To this end, I have interviewed less participants than I would have considered ideal for this research. A total of five users were interviewed, and no platform operators could be interviewed.

To compensate for this, I have conducted desk-based research with a strong focus on publicly reported materials, opinions, and blog posts from users who have been involved in each of the case studies. I have used these sources and provided lengthy quotes where appropriate to attempt to provide the perspectives that interview participants can offer to research, and to also allow the users quoted to have their case represented in their own words.

Finding Meaning from Methods

To develop meaning and theory from this research, I have adopted several approaches often used in qualitative research. Grounded Theory, to borrow from Charmaz (2014, 1), provides “systematic, yet flexible guidelines for collecting and analysing qualitative data to construct theories from the data themselves”. It is applied in this research to generate theory that may be relevant beyond the chosen case studies, and more broadly applicable to studies of social networking platforms. Whilst this research has not strictly followed Charmaz’s (2014, 18, 51) suggested methodology for grounded theory construction, it uses several of practices that Charmaz (2014, 43-45, 290-299) outlines for coding and developing categories for analysis. Critically, this research has three factors required by Hood (2007, 70) for Grounded Theory research — “theoretical sampling”, in the form of interview and close reading data, “constant comparison of data to theoretical categories”, with the practice of coding data gathered and the reflexive practices used in conducting semi-structured interviews, and “and focus on the development of theory [...] rather than substantive verifiable findings”.

Hood’s last requisite — the development of theory, rather than a focus on verifiable findings — is worth stressing for this research project, as studies of governance in online platforms are few and far between. Even more broadly, the field of research involving social networking platforms and their political dimensions is relatively young. For this reason, it is important to look to develop theory in this field in attempts to more clearly explain the occurring phenomena and their significance to research.

A key part of grounded theory methodology involves the use of interviews. In particular, gathering rich data from interviews has been facilitated by engaging participants in ways that ask them to reflect on their experience, the underlying social processes of their experience, and the meanings and consequences of these processes (Charmaz 2014, 43). These goals have informed the construction of interview questions (discussed later in this chapter), as well as the analysis of the interview data.

Similarly, in performing close reading-based analysis of documents relating to each case study, I have reflected on Charmaz's (2014, 45) discussion of 'Documents as Data' sources for grounded theory-style research. Charmaz notes; "Documents comprise one type of text whose form, content, purpose, accessibility, visibility, utility, legitimacy, and consequences can raise intriguing questions", and extolls the virtue of analysing not just the content of a document, but the audience, authors, production and presentation of said document. These concerns have informed my practices of close reading and are addressed in each case study.

In this research, I have applied coding practices based on Charmaz's (2014, 113-137) discussion of coding practices. In particular, I have adopted the practice of initial coding (using loosely described themes in a first-pass analysis of data) on an incident by incident basis (applying coding at key points) for interview participants and timeline-based analysis and on a line-by-line basis (applying coding based on sentences, sequentially) for close readings of documents. By applying and reflecting

on codes for data, grounded theory practices allow for unifying analytically similar data.

In conjunction with these grounded theory practices, I constructed case-conflict timelines based on data gathered as part of desk-based research and provide a contextual analysis of the case's conflict — to facilitate analysis of the case's conflict not simply as a socially constructed phenomenon, but as a part of a complex socio-political environment.

Ethical Considerations

This research has been approved by the QUT Human Research Ethics Committee, with approval number 1400000446. The application sought clearance for semi-structured interviews with users and platform operators relevant to the conflicts studied in cases. The risks involved to participations were considered to be negligible to low risk. The identified risks include: the possibility that publication of errors of fact or interpretation and analysis could be harmful to the professional reputation of participants and, even where individuals are not explicitly identified, it may be remotely possible for readers to deduce individual identities and information sources from contextual information. Participants were given the option of having their identity masked by being referred to pseudonymously in this research, as well as the ability to drop out of the research at any time by indicating their wish to withdraw from the research project.

Summary

This research provides greater insight and clarity into increasingly regular conflicts between users and platform operators on popular social networking platforms, with particular reference to understanding the way in which operators attempt to govern, and the way in which users react to being governed by platform operators. It does this by using case-study methodology, combined with close reading, semi-structured interviews and traditional desk-based research to provide a more thorough understanding of the cases of user/operator conflict selected.

Ethics Addendum: Online Harassment, and new dangers of research

At the outset of this research project in early 2013, it was difficult to know just how severe online harassment would become. Although both anecdotal and documentary evidence can show that online abuse tactics (ranging from sustained and severe interpersonal conflicts in bulletin-board flame-wars (Pfaffenberger, 2012 28-38) to more distressing, invasive attacks like simulated digital ‘rapes’ (Dibbell, 1993)) have existed for quite some time, it became particularly clear during the course of this research project that both technological advances and affordances were able to be leveraged in ways that could amplify these tactics, and more significantly affect harms for the people targeted for abuse.

The conflicts analysed in the Twitter-related chapter of this thesis particularly engage with these issues, but it should be noted here that campaigns of abuse, harassment and harm online can be prompted by a cluster of users who are disproportionately incensed at the existence or prominence of conflicting or dissenting

opinions to their own. This is particularly evident in the ‘Gamergate’ conflict, discussed more thoroughly later in this thesis, where relatively *bien pensant* desires were being voiced for more thoughtful, and less stereotypical representations of women and people of colour in video games and the games press — and these desires were met with allegations of corruption, collusion and malfeasance.

In particular, the ‘Gamergate’ conflict brought into light the need for particular care and concern for the safety and representation of research participants and interviewees in research projects. A number of times throughout that conflict, pseudonymous and functionally anonymous users of Gamergate related message boards and Reddit forums made reference and discussion of Open Access academic papers (especially for academics who either identified as or were identified by people in the conflict as ‘feminists’, or for academics who were associated with the Digital Games Research Association (DiGRA)) as sources for identifying targets for abuse and harassment under the veneer of investigating “corrupt journalism”, and the influence of “staunch feminists” (Otton, 2014) in manifests as attempts from subcultures in gaming and internet-culture spaces that seek to tame and establish hegemony in games criticism, ostracising critical voices by creating social environments in which they are either socially constructed as illegitimate.

With this in mind, consideration and care for participant privacy and safety have been essential to the research conducted in this thesis — both for participants who requested pseudonymity and for participants who chose to participate with their own names. The risks of re-identification can be particularly severe, and require not only considering of their wishes, but the consideration of their spoken or written

contributions — to ensure that these are free from information which could be later used by a reader of this thesis to uncover their identity.

Similarly, data breaches on the end of the researcher (or the research institution responsible for their data storage or email account provision) can threaten the safety and privacy of participants. For this reason, much of the information used in this thesis that directly relates to research participants has been kept on encrypted, password-protected drives and only accessed on a non-network connected computers. Similarly, I worked to locally and securely archive necessary digital communications (including interview recordings, transcripts, or emails) — and will delete these immediately after the minimum retention period for research data and primary materials, according to the Queensland State Archives University Sector Retention and Disposal Schedule, to protect the privacy of participants.

Whilst the usage of technology to facilitate abuse and harassment is concerning, it is worth noting that these risks and problems have effectively always existed for academic researchers. A reactionary internet blogger who takes umbrage with an academic publication and attempts to draw fervent critical attention to it is not meaningfully different from a similarly minded newspaper columnist, talkback radio host or politician. In these cases, the key concern of a diligent academic should not be the attention of the blogger, but rather, the attention of their audience. A criticism of reactionary bloggers and the more muck-racking of tabloid columnists concerns an often *laissez-faire* indifference to the misbehaviours or outright belligerences of their audiences. I highlight this not to argue that it is necessary or appropriate to hold authors responsible for the actions of their readers, but rather to highlight that this is a recurring

pattern, and that researchers ought to take particular care when researching topics that might be seen as controversial, heterodox, or worse by potential readers. The advent of Open Access publishing and the availability of many PhD theses on University ePrints sources that are indexed by popular search engines means that academics are no longer afforded the protections of obscurity. Whilst the advent of more open publishing practices is a boon for academia at large, it demands a thorough consideration for the dangers that publication can present for research participants.

Whilst these issues may not strictly be ‘new’ for researchers, both the technological advancements and technological affordances of our current digitally inter-connected age have served to amplify these risks to a level that institutional ethics boards at Universities have likely not yet caught up to. This is exemplified by the risks that even the simplest of data breaches can pose — a participant’s name or e-mail address could easily be used in conjunction with dozens of data-brokering services in attempts to trawl publicly-available databases for other information about them. Combine the availability of data with the poorly secured, often easily discovered security questions that many services require: “Where did you grow up?”, “What was your mother’s maiden name?”; and a sufficiently dedicated harasser might soon have the wherewithal to wreak more serious havoc on a research participant’s life. For these reasons, it is imperative that researchers investigating issues related to online abuse, conduct their research with a thorough understanding of the risks associated with these projects, not only to themselves, but for their research participants. Moreover, it is critical that researchers have develop a strong awareness of information security practices to protect the privacy and safety of their participants as best as possible.

Chapter 4: Case Study One: Facebook and Breastfeeding

Photographs

Platform–Political Tensions

Contemporary Breastfeeding & Toplessness activism in society

Public Breastfeeding, laws and social stigma.

Facebook, the social networking platform that began as a social hub for college students and grew to a daily user-base that dwarves most nations within a decade, has long had challenges in governing the platform in accordance with the expectations and desires of its users. This chapter addresses the long-running controversy within the Facebook community, and between Facebook itself and some of its users, about the legitimacy of posting images of breastfeeding mothers.

Whether in relation to motherhood or to sexual agency, the toplessness or nudity of women has been a cause of controversy in Western societies for much of recent history. This controversy has historically manifested in both legal and social pressures on women who expose their breasts, whether for personal, socio-political or maternal reasons. Breasts, to borrow from feminist discourse, are both personal and political — as a part of a person's anatomy they are both in the personal domain of that person, but subject to regulatory pressures from the state and many other non-state actors.

Whilst this study looks particularly at nudity in the context of breastfeeding, there are some contextual similarities between the pressures and stigma faced by breastfeeding women and the social pressures being fought by women who seek to desexualise female nudity and campaign for the same topless ('topfree')² rights that are enjoyed by other male citizens.

Although breastfeeding is generally encouraged by parents and the medical community, there are perceived social pressures that act as a disincentive for mothers to breastfeed their children in both public and private situations. These include “hypersexualisation of the female breasts” in western discourse, benevolent and hostile sexist attitudes —both as expressed by others or as internalised by a mother (Johnston-Robledo, Wares, Fricker, & Pasek, 2007)— and general societal unfamiliarity with breastfeeding as a publicly performed practice (Acker, 2009). As a result of these problems and the poor societal support for breastfeeding, many mothers abandon the practice (Gertz, 2013). Although medical scholars, practitioner and policymakers almost universally encourage the practice (Gallegos, Vicca, & Streiner, 2013; U.S. Department of Health and Human Services, 2011) mothers who breastfeed in public still face stigma and attempts at shaming or admonishment from others.

These issues are complicated by local ordinances and laws relating to female nudity and/or breastfeeding. An example of these laws includes the now-defunct §

² The terms ‘topless’ and ‘topfree’ are often used interchangeably, however — ‘topfree’ is used by some advocates as a way of distancing the freedom of dress from the term ‘topless’, which is considered more of a sexualised term given its relationship to phrases like ‘topless dancers’ and other uses of the word ‘topless’ in the context of sexualised entertainment or titillation.

245.01 of the New York Penal Law, which imposed criminal sanctions upon women who bared their chests in public, but not on men who did the same (Glazer, 1993 113). Under this section of the Penal Law, it was women could be found guilty of criminal exposure if any area of the breast below the top of the areolae was shown in public. Whilst this section of the law provided specific exemptions for breastfeeding infants and performance art, it was challenged in 1986 by seven women who bared their chests and were arrested in protest at the section's unequal policing of male and female bodies — a six-year-long legal challenge that garnered a public following and significant media attention. In discussing this legal challenge, Glazer (Ibid. 114) indicates that the law influenced both the behaviour of the citizens it governs, but also how those citizens feel about themselves. This is apparent in the results of studies like Mulready-Ward and Hacket (2014) who found that in their survey of New York residents' attitudes towards public breastfeeding, over half of the 2,075 respondents were felt that breastfeeding should be done in private, and that approximately one third of respondents expressed discomfort with the idea of women breastfeeding near them. Mulready-Ward and Hacket (Ibid. 2014 198) considered that the poor support for public breastfeeding might be ameliorated by laws and ordinances that more comprehensively protect a mother's right to nurse children in public. They note that whilst New York City residents have this right protected under New York State law, interactions with figures of authority challenge this right — citing the example of a police officer who charged a mother with a misdemeanour offence for nursing her infant on a public commuter train. Whilst these laws are typically established by conservatively-minded lawmakers with a view to reducing 'the risk of harm' caused by 'obscenity or indecency' (Valverde, 1999 181-3). It's important to clarify, as Valverde (Ibid.) notes, that the notion of a 'risk of harm' established in this context

serves as both a patriarchal and heterosexist form of moral regulation, that would not consider ‘harm’ from a feminist perspective, or the perspectives of sexual minorities.

These laws and their effects on both the abilities of mothers to safely and legally feed their child, and the abilities of women to ‘go topless’ in the same fashion as shirtless men, provide context for much of the contemporary activism aimed at promoting breastfeeding awareness or extending equal ‘topless or topfree’ rights to women.

Toplessness activism

As Judge Vito Titone (Wachtler et al., 1992) notes in his judgement of *THE PEOPLE & C., RESPONDENT, v. RAMONA SANTORELLI AND MARY LOU SCHLOSS, APPELLANTS, ET AL., DEFENDANTS* (the aforementioned case relating to the seven women who bared their breasts in protest of § 245.01’s classification of areolae, nipple and lower breast nudity as a form of criminally sanction-able obscenity obscenity); “expert testimony at appellants’ trial suggested that the enforced concealment of women’s breasts reinforces cultural obsession with them, contributes toward unhealthy attitudes about breasts by both sexes and even discourages women from breastfeeding their children.”

Topless and Topfree activism is a relatively modern form of rebellion, a revolt against the inequality of policing female bodies and expression (in a manner that is inconsistent with the regulations that govern male bodies), and against the ‘unhealthy

attitudes' and 'cultural obsessions' noted in Titone's judgement (Ibid.), as exemplified by popular interpretations of female nudity as inherently sexualised (associated perhaps with a lack of exposure to non-sexual nudity (Herold, Corbesi, & Collins, 1994 140-141), or by exposure to 'pop porn' and 'raunch culture' (Oyler, 2015)).

Whilst organisations like the Topfree Equal Rights Association (*TERA*) of Canada have campaigned to raise awareness and understanding of topfreedom for over a decade (Rapoport, n.d.), their advocacy has been eclipsed in popularity and awareness, by the recent, celebrity-adorned 'Free The Nipple' campaigns in major western cities and against social platforms with policies perceived as censorious in relation to images of women's breasts.

Through their actions in support of equal public topfree rights for women in both physical public spaces and digital social networks, celebrities including Scout Willis, Miley Cyrus, Robyn 'Rihanna' Fenty, Chelsea Handler and Chrissy Teigen have bolstered the 'Free the Nipple' movement (Zeilinger, 2015), a campaign that takes it's from Lina Esco's 2012 documentary film of the same name. As Esco (2013) notes the film became emblematic of a wider movement "in a case of life imitating art — or [...] art catalysing civil, civic action," and in turn encouraged activist action from topless women to take action in both online, and in the film's focal location New York City. Esco noted that the campaign for nipple freedom, equal topless rights for women needed not only to be fought in cities and public spaces, but also online — on platforms where the policies platform operators can drastically affect the type of content that other users are exposed to. Esco responds to online censorship with a commonly-used Eisenhower quote that expresses an absolutist condemnation of

censorship, whilst also criticising the priorities of the platform's moderation — complaining that “you [can] show public beheadings from Saudi Arabia on Facebook, but not a nipple? [or] sell guns on Instagram, but yet they will suspend your account for posting the most natural part of a woman's body? As President Eisenhower said, in the era of McCarthy: "The most dangerous weapons of any tyrant, are not weapons and guns, but censorship."

Whilst Facebook is a platform developed in, and seemingly influenced by, the cultural and social practices of people in the United States of America, it bears considering that the globalised nature of the platform provides avenues means that the values and cultural practices that have informed its operators can be significantly different to those of its users. Consider, for example, that although attitudes towards nudity have been historically quite conservative in the United States (Komodiki et al. 2014), attitudes in other countries, like Australia, Canada, and some western European countries have been comparatively permissive. Additionally, there is an even greater diversity of behaviour and attitudes worldwide. Given this, we should consider the complexity of platform governance in this space — a ‘one size fits all’ approach may be desirable for platform operators looking to eschew the costs of complex moderation practices but may also alienate users who feel that the platform operators do not share their values.

The globalised usage of platforms like Facebook exists somewhat in contrast to the infrastructures of its governance — offices and bureaus operated in a handful of key countries, with a significant amount moderation labour outsourced to developing countries, and countries with lower labour costs than the United States and other G20 nations. Because of this, the owner/operators of Facebook, the moderators and the users often use the platform in radically different cultural contexts, Further research

should examine the contingency of cultural norms and values in globalised digital platforms and moderation practices on these platforms.

Nudity, Breastfeeding & Mastectomy on other platforms

When analysing the policies of social network platforms with regards to nudity, we can broadly categorise platforms as either ‘permissive’ or ‘non-permissive’.

Permissive platforms, broadly speaking, allow most forms of nudity in user-submitted content regardless of whether the purpose of the nudity is educational, journalistic, artistic or titillating. Whilst permissive platforms often have some caveats — like omitting nudity and other 'NSFW' (Not Safe for Work) content from search functions, or forbidding accounts associated with 'adult content' from using any of the platform’s advertising functions.

Non-permissive platforms, on the other hand, are often quite particular as to which forms of nudity are 'appropriate' (generally male torsos, occasionally male and/or female buttocks), which forms of nudity are 'inappropriate' (generally female torsos — more often the nipple and areolae than the breast or midriff — as well as reproductive organs like the penis, testes, labia (etc.)), and the contexts in which the appropriateness of nudity is negotiable (journalism, advertising, education, etc.).

Permissive Platforms

Twitter is ostensibly one of the most permissive platforms in regard to nudity in user content. Their policies allow for users to post content featuring nudity. Twitter's *Rules* instruct that "You may not use pornographic or excessively violent media in your profile image, header image, or background image", but aside from that regulation, most other instances of nudity are freely allowed on the platform. The Twitter 'support centre' also instructs Twitter Users that if they "upload media that might be considered sensitive content such as nudity, violence, or medical procedures, you should consider applying the account setting 'Mark my media as containing sensitive content'." This setting enables a function on the Twitter Web interface and official Twitter apps that shows other users a 'Sensitive Content' warning that they must click on to access the posted content. Although this is offered as an option for self-regulation for users, it can also be applied post-hoc by Twitter's operators — who might act proactively, or reactively — in response to complaints and reports about a user's content.

These responses can also be gamed by users — who might use flags in the hope of having rule-abiding content censored by the platform during an interpersonal (user-user) conflict (Lapowsky, 2015). To borrow from Crawford and Gillespie, these 'flags' that Twitter (and many other platforms) rely on as a part of their regulation framework represent a flawed system of governance, as they provide no public record or discursive space for contesting, and as such "they leave us little chance to defend the right for those things to be said, or to rally against something so egregious that it warrants more than a quiet deletion" (2014).

Twitter also makes exemptions to the rules for nudity for reasons including privacy and harassment—including the posting of intimate photographs without the consent of the photo’s subject (often termed ‘Revenge Porn’)— and for the posting of child abuse and exploitation material. Finally, Twitter also excludes ‘adult’ products from using their ‘sponsored tweets’ advertising functions, where ‘adult’ is shorthand for “internet pornography and magazines, books and films,[...] Prostitution and intimate companionship services,[...] Full and partial nudity, as well as sexual content, in images, video or language, Penis enlargement and breast enhancement products and service, Modelled clothing (such as lingerie) that is sexual in nature, Dating sites with sexually explicit themes, [...] Intimate and erotic massage services, Sex toys, International marriage brokerage services (e.g., mail-order bride services), [and] Dating and marriage services.” However, specific exemptions to this policy are made for “Safer sex education, HIV/STD awareness campaigns, and non-prescription contraceptives, [...] Modelled clothing (such as lingerie) that is not sexual in nature, Nude fine art, Content about breastfeeding, Medical drawings, [and] News and information related to sex and sexuality that is not sexual in nature”.

The social blogging platform, Tumblr, has similarly expansive and permissive framework for nudity in user generated content. To wit, Tumblr’s policy for ‘Adult Content’ is explained in the form of a frequently asked questions document (Tumblr, n.d.);

Is adult-oriented content allowed on Tumblr?

Sure. We have no problem with that kind of stuff. Go nuts. Show nuts. Whatever.

Tumblr deploys similar practices by asking users to self-regulate their content, by either marking their own uploaded content as 'NSFW' and flagging the unmarked NSFW content that other users may have uploaded. Tumblr largely promotes this practice by extolling the feature that flagging provides users — a 'work-safe' dashboard (the Tumblr analogy for a 'news feed' of followed blogs and sponsored content, so that users can have NSFW content filtered from their view when they use Tumblr at a workplace or institution with strict guidelines about pornography and internet usage. Whilst Tumblr does not clearly delineate a policy for 'revenge porn', intimate photos posted without consent are notionally forbidden based on the community guidelines against 'Privacy Violations', which reads;

“Don't post content that violates anyone's privacy, especially personally identifying or confidential information [...] or private photos of your ex's junk (no matter how remarkable).”

Tumblr's policies regarding advertising are much more permissive than Twitter's. It reads; “We take an open-minded approach to the use of sexual content in advertising. Tumblr allows ads that contain a sexual theme to a certain extent, but the ads must be tasteful, artistic, and high quality. Not trashy, porny, or lame. And absolutely never degrading or abusive.” Whilst this seems like a simple, reasonable policy — it must be noted that it is a policy almost entirely subjective in its parameters. '[T]asteful', 'artistic', 'high quality', 'trashy', 'porny', and 'lame' are all adjectives that allow for value judgements about the content being assessed. Whilst a distinction

between, say, ‘artistic’ and ‘porny’ might seem a reasonable dichotomy, these ideas are famously nebulous. In this regard, Tumblr’s policies seem to borrow from former United States Supreme Court Justice Potter Stewart’s now infamous expression in *Jacobellis v. Ohio* (E. Warren et al., 1964). When passing judgement as to determine what constituted ‘hard-core pornography’, he opined “I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description [hard-core pornography], and perhaps I could never succeed in intelligibly doing so. But I know it when I see it...”

Notionally art and creativity-focused social networks like Flickr, 500px and deviantArt, also have similarly permissive policies, generally allowing nudity in sexual and non-sexual contexts, so long as the users submitting the content exercise appropriate self-regulation by flagging their content as ‘adult’, so that it will not be shown to underage users, or that it will not be shown to non-members of the platforms. These guidelines tend to be more permissive for photo-based content, as opposed to video content — where websites like Vimeo ask that users “keep Vimeo artsy, not porny”, by only allowing “depictions of nudity and sexuality that serve a clear creative, artistic, aesthetic, or narrative purpose”, and “non-sexual nudity, including naturalistic and documentary depictions of human bodies” (Vimeo, n.d.). Youtube maintains a similar policy but provides clearer examples of their regulatory boundaries for nudity — insisting that videos featuring nudity may be deemed appropriate or inappropriate on the basis of their context — providing the example of a breast cancer documentary as appropriate, whilst a video of excerpted nudity devoid of the documentary context would be inappropriate. It might be the case that Video platform providers are slightly less permissive in these regards for fear of being associated with video pornography

(and the perception that video pornography is more ‘graphic’ than photographed, auditory, or written erotica and pornography).

Non-permissive platforms

Google Plus’s policy is functionally similar to Youtube’s policy, allowing “naturalistic and documentary depictions of nudity (such as an image of a breastfeeding infant), as well as depictions of nudity that serve a clear educational, scientific, or artistic purpose”, but also making requesting that users avoid including nudity as a significant part of their profile photo. The Google+ policy also provides a clear admonition against users using the platform for the distribution of “sexually explicit or pornographic material” and instructing them “Do not drive traffic to commercial pornography sites”.

Facebook’s nudity policy is associated with their ‘Encouraging respectful behaviour’ community guidelines. The policy uses diplomatic and deferential language that avoids making value judgements about nudity in content but attempts to explain the platform’s reasons for policing — that “[we] restrict the display of nudity because some audiences within our global community may be sensitive to this type of content - particularly because of their cultural background or age.”

Interestingly, Facebook also pre-emptively apologise for the inaccuracies and inadequacies that they foresaw in their enforcement of this policy — the idea that their

uniformly designed policies can sometimes “be more blunt than we would like and restrict content shared for legitimate purposes.”

Whilst some guidelines are given as to notionally appropriate and inappropriate photographs, these mostly fall into the same dichotomous camps as Google+’s regulations. Naturalistic and educational content is more likely to comply with rules, whereas images that are devoid of this context are more likely to be removed by the platform operators. Like Google, Facebook also specifically forbids images of sexual intercourse, and vivid textual descriptions of sexual acts, with exemptions for educational, humorous and satirical purposes.

Instagram features many of the same policies again, and this is perhaps to be expected given Instagram’s relationship as a subsidiary platform and company to Facebook. An interesting difference is in Instagram’s reference to external regulation. As Instagram is largely used via smartphone apps on iPhone and Android smartphones, the platform’s content guidelines make specific references to the rules of those systems. To that end, Instagram implored users “While we respect the artistic integrity of photos and videos, we have to keep our product and the content within it in line with our App Store’s rating for nudity and mature content. In other words, please do not post nudity or mature content of any kind.” More broadly, the platform has functionally similar rules to Facebook, but communicates these rules in a significantly more paternalistic, perhaps moralising way, instructing; “If you wouldn’t show the photo or video you are thinking about uploading to a child, or your boss, or your parents, you probably shouldn’t share it on Instagram.” Instagram also includes an admonition towards posting pictures of naturalistic nudity of children under thirteen.

“We may remove images that show nude or partially nude children for safety reasons. Even when these images are uploaded with the best of intentions, this content could be re-shared and used by other people in unanticipated and inappropriate ways”. This disclosure is unusual, mostly because of its rareness — no other platforms analysed for this research made any reference to the nudity of children. Whilst this could be based on a very liberal reading of the United State of America’s *Children’s Online Privacy Act* (which restricts the information that online service providers can gather on users under the age of 13), it could also be a reference to concerns held by platform operators about child abuse material.

It is important to point out that the differences between these platforms and their policies are not simply emblematic of their different perspectives as organisations and businesses, but also of the environments of those social platforms, and the norms expected of their users.

Platform History and the Conflict

Early Controversy: before the publicised policy (2007-2012)

The first publicly recorded instance of Facebook censoring a picture of a mother breastfeeding occurred in the June of 2007. The platform's operators removed user Kelli Roman's profile picture — a picture of her nursing her newborn child. Roman complained about this treatment and created a Facebook Group to organise and lobby for change on the social network. Later that month, Roman created the 'Hey Facebook, breastfeeding is not obscene!' group, which would serve as a key organising point for protests. Outside of the Facebook platform some users took to other platforms, like Blogger, to create and record the history of their grievances with the policies that Facebook held and enforced. In 2008, a number of users in these groups noted that Facebook had begun removing photos of mothers breastfeeding. Irked by this, they protested outside of Facebook's Menlo Park campus bringing significant media attention to the policy. The movement to change Facebook's policies slowly gained traction in the news media and was reported on in blogs and news outlets as pressure builds for Facebook to change their policies. Across the following years, the reporting focuses particularly on key users and operators of Facebook protest groups like Kelli Roman, Karen Speed and Emma Kwasnica, many of whom also wrote about being affected by the censorship afforded by Facebook's policies.

In 2012 the popular website Gawker obtains and publishes two copies of the 'Abuse Standards' Guidelines (versions 6.1 and 6.2) used by Facebook moderators, courtesy of a leak from Amine Derkaoui, a disgruntled young Moroccan man

associated with the oDesk online outsourcing platform. Chen (2012a) considered the guidelines to be particularly censorious, instructing the deletion of posts and content on the platform that included “blatant (obvious) depiction of camel toes and moose knuckles” (slang terms referring to outlines of the labia majora and penis visible in tight clothing), “mothers breastfeeding without clothes on”, “sex toys or other objects, but only in the context of sexual activity”, “depicting sexual fetishes in any form”, “ANY photoshopped images, whether negative, positive or neutral”, as well as “images of drunk and unconscious people”. When it came to nudity, the guidelines instructed that “naked ‘private parts’ including female nipple bulges [visible through clothing] and naked butt cracks” with the distinction that “male nipples are OK”.

Although some of the guidelines allow for an astoundingly broad range of policing³, the key distinction that affects this case study is the gendered distinction between appropriate (male) nipples and inappropriate (female) nipples. Whilst this somewhat mirrors contemporary western social norms regarding male and female toplessness and the social stigma directed at breastfeeding mothers (U.S. Department of Health and Human Services, 2011), this policy simply replicates these norms, rather than challenging them in a way that affords male and female users equal rights of bodily expression.

Organising Protests

³ How, for instance, does Facebook define a ‘fetish’? Similarly, if the digital alteration of images was *verboten*, then why were advertisers (or any professionals using digital manipulation tools like Adobe Photoshop not struggling with the platform’s censorship more frequently?

Protests against Facebook's policies in this area occurred both on and off the social platform. Noguchi (2008) noted that a 2008 nurse-in, held by the organisers of the 'Hey Facebook, breastfeeding is not obscene!' group was a quintessential "Palo Alto protest: A handful of peaceful pickets discreetly tucked away in a University Avenue plaza; placards reading 'Hey Facebook, Breast-feeding is not Obscene', protesters chatting up the media; and indifferent passers-by." Although many news outlets deemed the protest newsworthy, it seemed not to be spectacular enough to warrant much further coverage (Lunceford, 2012 61). Whilst the nurse-in only attracted a handful of protesters in person, an organiser of the protest, Stephanie Muir (Kozyra & Muir, 2009) explained that the event and its media profile led to even more opportunities for media coverage and advocacy across outlets with large audiences like daytime television show *Dr. Phil*, and internationally-read newspapers like the New York Times.

For a number of users who were prominent in protesting Facebook's policies, blogging platforms provided an opportunity to critically engage with Facebook's policies in an environment in which Facebook had limited influence, and limited ability to interfere. From 2008 onwards, a host of motherhood and breastfeeding blogs covered Facebook's policies and the effects of their policing in detail. Blogs like OnCloudMom (Kozyra & Muir, 2009) interviewed people who had participated and organised protests, whilst Bliss Breastfeeding (Bliss Breastfeeding, 2007) and Unlatched (Lesteshen, 2013) reported on users being banned, and republished the photos that had resulted in users being temporarily or permanently banned from the Facebook platform.

Although the earliest coverage of Facebook's censorship of breastfeeding photographs dates as early as 2007, it seems to take years for mainstream news organisations to articulate from covering this conflict as an occasional 'outrage' against Facebook (headlines often indicated a perceived novelty-cum-newsworthiness of nurse-in protests), to covering the conflict as a campaign in earnest to change Facebook policies. In 2009 and 2010 articles on the New York Times Technology Blog (Wortham, 2009) and the TIME Website (Block, 2010) begin the newer wave of coverage in earnest, focusing less on the novelty of 'protests against Facebook', and more on the substantial political issue of Facebook as a censor, and the platform's seemingly obstinate policy position in this conflict.

In 2012, shortly before the leak of Facebook's moderation guidelines, a writer for the American politics and culture website *Slate* condemned the platform for its capriciously enforced policies and argued that Facebook was unable to define a 'dirty picture – or the difference between biology and porn'. Following this article, and the subsequent leak of Facebook's abuse guidelines, The Guardian started a brief pro-breastfeeding campaign aimed at challenging Facebook's policy. A raft of articles from Arthur (2012), Davies (2012), Reed and Gardiner, (2012) and Belinda Phipps (2012), the CEO of the United Kingdom's National Health Service Trust took aim at Facebook's policies, and the stigma that women who breastfeed in public often face.

Platform Responses

Shortly after the first Nurse-in at Facebook's Palo Alto Campus, Facebook spokespeople started to engage with the press, to discuss Facebook Policy. Barry Schnitt, a spokesperson for Facebook, told The Guardian in 2008 that that (Sweney, 2008);

“Photos containing a fully exposed breast, as defined by showing the nipple or areola, do violate [our] terms (on obscene, pornographic or sexually explicit material) and may be removed. The photos we act upon are almost exclusively brought to our attention by other users who complain.”

Later, Schnitt contested the idea that they had a policy at all, dismissing the concerns of a reporter (Ingram, 2011):

“No policy against this at all. [...] We have removed some pictures of naked women who happen to be holding a baby.”

Schnitt's comments are typical of Facebook's attempts to engage with dissenting users in the earlier years of this conflict. The public statements made by the organisation are few and far between, and often seem to consist of a spokesperson making a short statement to the press after an event has gained enough traction to become a news story. In April of 2012, Facebook purchases the photo-oriented social platform Instagram for approximately US\$1 billion (\$700m in stock offering). In the following month, May, Facebook begun its initial public offering (IPO) of stock on the NASDAQ Stock Market. Throughout the following months, Facebook seem to attempt awareness and education campaigns to more clearly communicate their

policies, and the ways in which they wanted users to understand the systems at work inside Facebook. This included blog posts from Facebook’s Safety Team, describing (and providing a flowchart of) what is involved in the process of reporting content on Facebook, and the teams that are responsible for reported content (Facebook Safety, 2012).

Similarly, in the December of 2012, Facebook updated the wording for their publicly available ‘Community Standards’ (Facebook, nd), communicating in more specific terms the platform’s stance towards photographs of mothers breastfeeding. This marks the first time that ‘breastfeeding policy’ is considered in its own right as a rule on Facebook, rather than a scenario covered by other policies that address explicit nudity or pornography. As such, I consider that it marks a significant point for the conflict analysed in this case study — the formalisation and publication of a policy that directly engages with the core conflict with users. It also marks what appears to be a strategic change in communication which will be discussed later in this chapter.

Formalised Facebook Policy (2013 - Present)

The formalisation of a ‘Breastfeeding photo policy’ is marked by three significant changes in Facebook’s publicly available policies and rules. In a December 2012 update to their Terms of Service, Facebook instructs that;

“You will not post content that: is hate speech, threatening, or pornographic; incites violence; or contains nudity or graphic or gratuitous violence.”

Meanwhile, on the Community Standards page, where most of the platform's rules are communicated in language that is more approachable and less legalistic, an updated passage relating to posting of content including nudity advises users that;

" Facebook has a strict policy against the sharing of pornographic content and any explicitly sexual content where a minor is involved. We also impose limitations on the display of nudity. We aspire to respect people's right to share content of personal importance, whether those are photos of a sculpture like Michelangelo's David or family photos of a child breastfeeding."

Finally, in a supplementary page that seems to explain Facebook's community Standards in a 'frequently asked questions' format, Facebook explains that;

Does Facebook allow photos of mothers breastfeeding? Yes. We agree that breastfeeding is natural and beautiful and we're glad to know that it's important for mothers to share their experiences with others on Facebook. The vast majority of these photos are compliant with our policies. Photos that show a fully exposed breast where the child is not actively engaged in nursing do violate the Facebook Terms. These policies are based on the same standards which apply to television and print media. It's important to note that the photos we review are almost exclusively brought to our attention by other Facebook members who complain about them being shared on Facebook."

A supplementary question elaborates about another aspect of non-pornographic nudity that had become contentious for Facebook users, photos from women documenting their experiences with breast cancer — often including mastectomy scars or photographs of breast reconstruction;

” Does Facebook allow post-mastectomy photos? Yes. We agree that undergoing a mastectomy is a life-changing experience and that sharing photos can help raise awareness about breast cancer and support the men and women facing a diagnosis, undergoing treatment, or living with the scars of cancer. The vast majority of these kinds of photos are compliant with our policies. However, photos with fully exposed breasts, particularly if they're unaffected by surgery, do violate Facebook's Terms. These policies are based on the same standards which apply to television and print media, and that govern sites with a significant number of young people.”

In June 2014 the policy changes again — but very slightly; an updated community standards page excises the phrase *“Photos that show a fully exposed breast where the child is not actively engaged in nursing do violate the Facebook Terms. These policies are based on the same standards which apply to television and print media”* and change the words *“It’s important to note”* to *“Please note”*.

Again, in the March of 2015, Facebook update their community standards again, primarily adding explanatory text to more clearly explain the nuances of their policies;

“We remove photographs of people displaying genitals or focusing in on fully exposed buttocks. We also restrict some images of female breasts if they include the nipple, but we always allow photos of women actively engaged in breastfeeding or showing breasts with post-mastectomy scarring.”

This re-introduces some of the detail from the page that was present in 2013 and excised in 2014, but with clearer detail and more conciliatory language.

Although each of these changes are quite minor, both in text and in effect, they each attracted a small storm of attention from the news media. Generally, this coverage is of a positive tone, and remarks that Facebook’s concessions in policies are a ‘fix’ that brings their policies closer into line with the expectations of mothers on the platform.

Mainstream Media as conflict battleground

Over the course of this conflict, there have been hundreds of articles written for newspapers, websites as well as broadcast reporting on radio and television.

For users and platform operators alike, the mainstream news media form a key battleground in this user-platform conflict, as professional journalists and newsmakers lend a credibility or legitimacy by simply covering a story — and as a result, communicating to their audiences that that conflict is important, significant and worthy of their attention. This is invaluable to the campaigns of users endeavouring to affect

change in the policy or organisation. It has many significant uses for platform operators — functioning not only as a basic form of marketing for their service (by telling non-users that it exists, and that is worthy of newsprint/airtime/bandwidth), but also as an opportunity for platform operators to communicate broadly, in a broadcast fashion, so that they can communicate to participants in their conflict, to non-participants in the conflict, and non-users alike.

Instagram — Subsidiary Platforms and Policy Overlap

Facebook's acquisition of Instagram in 2012 allowed the company to benefit from the growth of the popular image sharing service, as well as use it as an ancillary platform for sharing photography on Facebook. In comparison to Facebook, Instagram's policy changes and controversies have been relatively infrequent. Whilst a policy change in 2012 that related to the ownership and usage rights of images caused quite significant consternation and controversy, Instagram's community guidelines and content rules seemed to avoid any large-scale protest or controversy until 2014. Instagram's most significant controversy also related to the platform's policing of female nudity. After censoring the topless photographs posted by a number of high-profile celebrity Instagram users, the platform came under criticism for its censorship of female nudity.

Throughout 2014, a handful of high-profile female celebrities including *Vogue* Editor Grace Coddington, socialite Scout Willis, and singer Robyn 'Rihanna' Fenty each experienced censorship at on Instagram for breaching the platform's policy for

nudity. In each case, the images were removed because they featured nipples, although the images that were removed were each thematically different. Fenty's images were a set of professional fashion photographs from a photoshoot with a French magazine, Whilst Willis's removed image was a candid polaroid-style snapshot of her and a friend standing together shirtless, and in contrast, Coddington's image was a cartoonish sketch of a woman with red hair, sunglasses and heels, reclining naked, with legs-crossed in a folding sun-lounge. This suggests that the platform's policy is quite absolutist on policing the existence of nipples in an image, rather than flexible for the potential for an image to be considered contextually appropriate.

As noted earlier, Instagram's early community guidelines place restrictions on nudity and more nebulously defined 'mature content'. Whilst these guidelines are relatively straight-forward, the communication of this policy makes peculiar calls to social norms:

"If you wouldn't show the photo or video you are thinking about uploading to a child, or your boss, or your parents, you probably shouldn't share it on Instagram."

and a concerned conservatism with relation to childhood nudity:

"While we know that families use Instagram to capture and share photos of their children, we may remove images that show nude or partially nude children for safety reasons. Even when these images are uploaded with the best of intentions, this content could be re-shared and used by other people in unanticipated and inappropriate ways."

In 2015, shortly after Facebook updated the wording and presentation of their policies, Instagram also updated their community guidelines — adopting a more distant and discerning tone, that bears a great deal of similarity to Facebook’s policies of the time, with particular references to “active” breastfeeding:

“We know that there are times when people might want to share nude images that are artistic or creative in nature, but for a variety of reasons, we don’t allow nudity on Instagram. This includes photos, videos, and some digitally-created content that show sexual intercourse, genitals, and close-ups of fully-nude buttocks. It also includes some photos of female nipples, but photos of post-mastectomy scarring and women actively breastfeeding are allowed. Nudity in photos of paintings and sculptures is OK, too.”

Instagram, unfortunately, doesn’t elucidate on the “variety of reasons” for which they do not allow nudity on the platform. It may be the case that this allows Instagram no more easily market itself to more conservative users, or to justify exemptions from censorship in countries with more censorious governments and less internet freedoms. In the absence of Instagram’s justifications, we can only speculate as to the reasons for this policy’s development.

Governance Context

Users (Methods of Subversion and Protest)

Off-Platform: Blogs

Off-platform protests present users with opportunities to contest Facebook's governance in spaces that are not subject to Facebook's governance. This is particularly important for users attempting to campaign against censorship, as it provides a safety for their expression analogous to the kind of safety sought by activists and dissidents who campaign for change in their homeland from the safety of another country. In this case study, the conflict surrounding Facebook's breastfeeding policies, a number of blogs became prominent for their archival of the conflict.

Whilst most blogs included varying degrees of information about breastfeeding education and activism, Rachelle Lesteshen's *unlatched* blog focused quite significantly on Facebook censorship, and catalogued both Rachelle's personal experiences with Facebook's moderators, as well as sharing the accounts of others who had experienced censorship on Facebook. Lesteshen accomplished this with screenshots as a form of pictorial evidence. Her *unlatched* blog also served as a catalogue and advocate for using the platform in subversive ways, with attempts to game the systems of policing to expose the systems as ridiculous. A key example of this includes Lesteshen's 'Nipples in Nature' campaign, in which she posted photographs of naturally occurring objects like acorns, lemons and eggshells, where parts of the photographed objects looked somewhat reminiscent of human female

nipples (Lesteshen, 2013). A follow up to this campaign noted that Lesteshen received a temporary, week-long, ban from the platform for a photograph of a chicken egg, where the pigmentation of the eggshell vaguely resembled a nipple and areola.

Two of the publicly-shared images that led to her punishment are discussed in her blog (Lesteshen, 2013). The first, is an act of protest against Facebook’s policies — it complains “Tell Facebook to Stop Bullying Breastfeeders”, with a link to a petition and a picture that another user had reportedly received a 12-hour ban for. The second, was a picture of an egg, in which the darker pigmentation of the top of the eggshell somewhat resembled a nipple and areola it was captioned “#nipples in nature”, “#nipples are normal”. In both cases these reports allowed Lesteshen the opportunity to appeal Facebook’s decision — which she did, complaining that “One was a breastfeeding photo that in no way violated any rules, was previously reported on another page, and Facebook recanted and said it was removed by mistake, and the second photo was a picture of a chicken egg” (Lesteshen, 2013). Facebook dismissed the first report, but bizarrely decided to censor the second image (the egg) and banned Lesteshen for 7 days for “posting nudity or pornographic content”. Lesteshen’s reactions to her experiences of censorship seem to be informed partly by a basic reading of Facebook’s breastfeeding policy (as she notes on her blog, Facebook ‘allows images of breastfeeding’), and partly by her own view that Facebook is not permissive of breastfeeding photos in spite of their own policies. Closely read, the policy (Facebook, nd) that governs this is quite helpfully revealing of Facebook’s strategies, in spite of the fact that its wording seems to be intentionally misleading. It reads:

Does Facebook allow photos of mothers breastfeeding?

Yes. We agree that breastfeeding is natural and beautiful and we're glad to know that it's important for mothers to share their experiences with others on Facebook. The vast majority of these photos are compliant with our policies.

Photos that show a fully exposed breast where the child is not actively engaged in nursing do violate the Facebook Terms. These policies are based on the same standards which apply to television and print media.

It's important to note that the photos we review are almost exclusively brought to our attention by other Facebook members who complain about them being shared on Facebook.

The “Yes” is a notional (but conditional) approval of breastfeeding photos, followed by a not-particularly-revealing platitude. The third and fourth sentences expose the strategy and the rationalisations of the platform operators. “Photos that show a fully exposed breast where the child is not actively engaged in nursing do violate the Facebook Terms” — in this sentence the implication is that female nipples and areola are inherently sexual or pornographic imagery, so if the “actively nursing” infant does not obscure this part of the female anatomy from view then the image would be considered explicit by Facebook. The following sentences rationalise the censorship — alleging that the platform merely reproduces standards existing in traditional media and pointing out that Facebook does not go looking for nipples to censor — it merely censors the nipples reported by other, less permissive users. In this, Facebook’s operators try to be seen as a supporter, rather than a censor, and as a moderator of existing social values, not a constructor of new values. Taken as a whole,

the policy represents an effort from the platform operators to bet a progressively minded government, but it then conditionally endorses only certain types of photos of breastfeeding and blames its censorship on the standards of traditional media (we might suspect Facebook's operators are not familiar with the 'Page 3' practices of tabloid newspapers like The Sun) and other platform users — the political equivalent of a government suggesting their 'hands are tied'.

Similarly, the *Bliss Breastfeeding* blog run by a pseudonymous blogger (Bliss Breastfeeding, 2007) recorded and publicised the experiences of mothers, and was a point of reference for breastfeeding activists during the earlier days of the conflict. Whilst the blog focused more significantly on breastfeeding education and advocacy, it also provided readers with a boilerplate letter of complaint, and tips for contesting Facebook's judgement when photos were censored. In addition to users campaigning for change, a handful of the better-known toplessness activist organisations such as Canada's 'Topfree Equal Rights Association' (TERA) (Rapoport, n.d.), and the Raëlian⁴ 'Go Topless' organisation of California (Go Topless Inc., n.d.) used their websites and social media presences to spread news of this conflict amongst their followers, and in the case of TERA, to archive some of the photographs that had been censored by Facebook.

Off-Platform: Mainstream Media

⁴ Raëliism or Raëlianism is a 'UFO' (Unidentified Flying Object) religion founded in 1974 and based on the claims of former automobile journalist and race car driver Claude Vorilhon.

As discussed earlier in this chapter, the mainstream news media —newspapers, television and major websites alike— have been a battleground of sorts in this user-platform conflict, as professional journalists and newsmakers can lend a credibility or legitimacy to user’s campaigns against Facebook, but also be a valuable asset for platform operators like Facebook in delivering a message more broadly — to non-participants in the conflict, and non-participants on the platform alike.

That said, mainstream media platforms can also be active stakeholders in these conflicts. In 2012, As noted earlier, UK-based news organisation The Guardian started a week-long pro-breastfeeding campaign aimed at challenging Facebook’s policy. Emboldened in part by the Facebook moderation guidelines leak supplied by Derkaoui to Chen (2012a), many journalists published articles that strongly criticised Facebook’s censorious attitude to female nudity in maternal contexts. Davies (2012), in particular, encouraged action from his readers — “Let’s put Facebook’s ‘no nipples’ rule to test with your breastfeeding photos”, asking readers to email in their breastfeeding photographs with the promise that The Guardian would post the images on their own Facebook page to challenge the rules. The Guardian received more photos (hundreds, according to Reed and Gardiner) than would be reasonable to post of Facebook, and elected to post only a handful — instead uploading many others to an album on the more permissive Tumblr platform (Reed & Gardiner, 2012). For the photos that The Guardian posted to Facebook, they found that the notoriety and status of being a major news company did not afford any special privilege in posting breastfeeding photos — they were each reported and taken down in turn by Facebook moderators. After Reed and Gardiner (Ibid.) published their article, Facebook’s then Director of User Operations contacted The Guardian with a familiar comment;

“Like many Guardian readers, we agree that breastfeeding is natural and we are glad that many mothers choose to share their experiences with others via Facebook.

The vast majority of breast-feeding pictures posted to Facebook don't break our rules, which everyone can read by looking at our statement of rights and responsibilities, so they remain on our site. However, some photos, like ones that contain a fully exposed breast that do break our rules, so we will remove them if they are reported to us.

Our rules are based on the same standards that apply to television or print media and they are designed to fit the needs of a diverse 845 million people strong community which includes people as young as 13.”

The CEO of the United Kingdom's National Health Service Trust, Belinda Phipps, was openly critical of this response (Phipps, 2012), citing in an op-ed the importance of breastfeeding, and the existing pressures that women face in breastfeeding. Phipps took particular concern with Facebook's deferral to their 'diverse 845 million people strong community' and their implied values. She argued “the implication is that while Facebook may 'agree that breastfeeding is natural' and feel comfortable with pictures of mothers breastfeeding, there are many people who would not be. In fact, most adults are just fine with seeing mothers breastfeed. In a Department of Health (England) survey, 84% of adults did not have a problem with women breastfeeding in public places. This was repeated in a survey in Wales, where 85% of adults were not concerned about women breastfeeding in front of them.” These

figures are dramatically different from the public breastfeeding support rates in the United States referenced earlier in this chapter and point to the potential of a broader cultural disconnect between the United States-headquartered Facebook and their internationally diverse users.

On-Platform: Groups, Subversive Speech

Although less prevalent at the beginning of the case study, practices of on-platform protest seemed to grow in strength over time, particularly closer to the end of the research period. Whilst some users have leveraged the power of off-platform protest to work against, others simply continued to use the platform in bold or indifferent protest to the platform's policies. These users clearly demonstrate the civil disobedience extolled by Russell (1961), by breaking what they see as the unjust or unreasonable 'laws' of the platform. These practices are also not dissimilar to the 'political jiu-jitsu' practices documented by George (George, 2013 387-392) in protest of authoritarian government practices in Singapore. Intentionally or otherwise, these acts of overt protest and obedience tend to work in the favour of users, by highlighting the platform's zealous moderation, and contrasting that with the user behaviour.

Facebook's Policies and Policy Development

Informal Policy and Communication

In 2008, when this conflict started to gather significant attention from mainstream news outlets, it also started to receive responses from spokespeople inside Facebook. After the nurse-in protest outside of Facebook's Menlo Park campus brought of media attention to the policy a spokesperson for Facebook told the media at the time that (Schnitt in Sweney, 2008);

“Photos containing a fully exposed breast, as defined by showing the nipple or areola, do violate [our] terms (on obscene, pornographic or sexually explicit material) and may be removed. The photos we act upon are almost exclusively brought to our attention by other users who complain.”

I argue that this statement represents Facebook's first publicly communicated ‘*informal policy*’. The language of this expression in particular, is referenced by users in their protest (Kwasnica, 2009), with particular disagreement about the notional categorisation of breastfeeding alongside “obscene, pornographic or sexually explicit material”. Although this policy is not formally presented in Facebook's terms of service or community standards with direct reference to breastfeeding, it is publicly communicated by a spokesperson in an official capacity for the platform and is essentially treated as a *de facto* policy statement.

Three years later, the same spokesperson contested the idea that they had a policy at all, telling a reporter who argued that Facebook had a censorious influence on discourse that they had (Schnitt in Ingram, 2011):

“No policy against this at all. [...] We have removed some pictures of naked women who happen to be holding a baby.”

This is, notionally, a recanting of the earlier statement. Whilst it is likely that the aim of this statement was to dismiss or allay concerns about Facebook’s role as an intermediary for public discourse (and, by extension, their function as a censor in that role), this statement can also be read as a glib or flippant and dismissive response to the concerns expressed by users. The language in this ‘*non-policy*’ statement seems to authoritatively dismiss concerns about a policy and attempt to cast the censored photographs as ‘dirty’, disingenuous, or just plain inappropriate.

Facebook’s early forms of policy in this area, the ‘*informal policy*’ and the ‘*non-policy*’ were both made in response to criticisms or protest events that received media coverage. We can argue, therefore, that the public presentation of Facebook’s policies regarding photographs of breastfeeding were mostly reactive to media attention. It’s worth noting that we can only significantly discuss the public facing policies here, as the policies and systems used internally by Facebook Moderators are essentially unavailable to the public or to independent researchers. That said, we can consider that the internally used rules and policies have likely formally existed for longer, as evidenced by the references to breastfeeding and nudity in the leaked Abuse Guidelines for moderators provided by Chen. (2012a).

Formalised Policy — Language and Governmentality

In 2013, a formalised policy was announced in Facebook’s community guidelines. It largely enshrined into policy some of the responses Facebook spokespeople had deployed in response to conflict. This new, formalised policy told users (Facebook, nd):

“We agree that breastfeeding is natural and beautiful and we're glad to know that it's important for mothers to share their experiences with others on Facebook. The vast majority of these photos are compliant with our policies. Photos that show a fully exposed breast where the child is not actively engaged in nursing do violate the Facebook Terms. These policies are based on the same standards which apply to television and print media. It's important to note that the photos we review are almost exclusively brought to our attention by other Facebook members who complain about them being shared on Facebook.”

The following year (2014), Facebook amended this formalised policy, excising the part that distinguished between the photos that showed ‘active nursing’ and attributed their policies to broader media standards. Again, in 2015, Facebook amended this policy, by reintroducing the breastfeeding policy to their policy on nudity in images. Whilst the effect of the policy is not markedly different, the text of the policy changed significantly. The new text of the policy read (Facebook, nd):

“We remove photographs of people displaying genitals or focusing in on fully exposed buttocks. We also restrict some images of female breasts if they include the nipple, but we always allow photos of women actively engaged in breastfeeding or showing breasts with post-mastectomy scarring. We also allow photographs of paintings, sculptures, and other art that depicts nude figures. Restrictions on the display of both nudity and sexual activity also apply to digitally created content unless the content is posted for educational, humorous, or satirical purposes. Explicit images of sexual intercourse are prohibited. Descriptions of sexual acts that go into vivid detail may also be removed.”

For brevity’s sake, we will discuss the 2013, 2014 and 2015 changes as the ‘*formalised*’ policy, the ‘*amended*’ policy, and the ‘*re-categorised*’ policy respectively.

You can see some distinct changes in attitude between the formalised and informal policies. Where the initial statements to the media (the ‘*informal*’ and ‘*non-*’ policies) are almost confrontational in nature, with Facebook spokesperson Schnitt arguing an authoritative ‘Facebook’ perspective on nudity and dismissing notions of conflict, the 2013 ‘*formalised*’ iteration is markedly deferential. Whilst Facebook offers platitudes as to the rights of users and the emotional value of mothers sharing treasured photographs, they furtively blame their practices as moderator and censor on existing social norms vaguely attributed to print and screen media⁵ and the

⁵ Perhaps also erroneously attributed, given that nudity is not particularly uncommon in print media (where it appears semi-regularly in Fashion, Health and News magazines, among other sources) or

users who use the ‘report’ tool to flag the content for a moderator’s attention. At the same time, it couches Facebook’s intermediary role in comfortable platitudes, intimating the platform’s helplessness to create, shape or defy social norms.

The difference between the ‘*informal*’ policy’s description of nipples and areolae, and the more ‘*formalised*’ policy’s mandate that exposed breasts must be ‘actively engaged in nursing’ is largely arid semantics. A child ‘actively engaged’ in nursing will serve the same ends for the platform as the censorship of a photograph — acting as tiny human censor-bar for the image’s *verboden* nipple.

Facebook has become more sophisticated in their use of language to communicate their policy and their goals. From the ‘*informal*’ and ‘*non-*’ policies, Facebook made direct reference to their policies on nudity, explicit content and pornography when discussing breastfeeding photography. As explained by the critics of these policies, many mothers considered the imputation that feeding their child was an explicit pornographic act in the eyes of the platform a deeply offensive idea. By strategically using less direct language, Facebook is able to appease users and perhaps even to escape some degree of scrutiny.

The ‘*formalised*’ and ‘*amended*’ policies from 2013 and 2014 offered platitudes with affirming sentiments to users — like “We agree that breastfeeding is natural and beautiful and we're glad to know that it's important for mothers to share their

screen media (where it appears somewhat regularly in both drama and documentary film and television (regional censor regimes notwithstanding.)

experiences with others on Facebook”. By demonstrating an opinion that breastfeeding is good, Facebook afford themselves an opportunity to describe their role as an intermediary with insignificant social power. In this manner, Facebook can shift blame for their actions — telling its users “Sorry! It’s not our fault!” – as if their policies and their enforcement is mandated by existing social norms, or by the whims of their more prudish users.

Whilst Facebook is obviously the primary arbiter and executor of power in conflicts with users on its platform, they have strategically used both softer language and a position of deference in a way that serves to minimise the chance that their role as an intermediary will be seen as prurient or excessively meddling in the day-to-day affairs and minutiae of its users lives. By blaming both the existence of the policy and its enforcement on an ill-defined Other, Facebook’s operators can hide their own strategic processes of governing behind what Morozov (2013) and Alder (1999) describe as a ‘technocratic pose’ —a rhetorical stance that conceptualises a politically neutral relationship between a technology, its users and its engineers— that serves to ‘de-politicise’ the platform from the perspective of its users.

Whilst we can reasonably expect that platforms like Facebook create and enforce regulations and make judgements about the content that users contribute —given their notional ownership of the server and software infrastructure, if not the content— it is important to remember that this practice of moderation and governance is a work of political and social decision-making. Moreover, the pretence of appearing as a ‘neutral’ conduit for user desires does not absolve Facebook of responsibility for its policies and their consequences, and the role of Facebook’s operators in these systems.

Finally, the most recent policy, 2015's *re-categorised* policy demonstrates the most significant changes. In this policy, Facebook re-contextualises their judgements regarding photos of mothers breastfeeding and photos of mastectomy scars into their Community Standards regarding nudity. The language is clear and avoids condemning users who come afoul of the policy.

The first paragraph — essentially, a preamble to the policy — consists of four key parts: An acceptance (“People sometimes share content containing nudity for reasons like awareness campaigns or artistic projects”), a rationale for their actions (“We restrict the display of nudity because some audiences within our global community may be sensitive to this type of content - particularly because of their cultural background or age”), an explanation for their failings (“In order to treat people fairly and respond to reports quickly, it is essential that we have policies in place that our global teams can apply uniformly and easily when reviewing content. As a result, our policies can sometimes be more blunt than we would like and restrict content shared for legitimate purposes”), and a pledge to improve (“We are always working to get better at evaluating this content and enforcing our standards”). Whilst this paragraph still makes deference to a need or desire for Facebook to police content according to extant social norms, it also much more clearly emphasises Facebook’s role as an intermediary arbiter.

The second paragraph gets to the heart of policy and simplifies it to explain; “We remove photographs of people displaying genitals or focusing in on fully exposed

buttocks. We also restrict some images of female breasts if they include the nipple”. This clearly defines what Facebook’s policies consider as infringing content (specific displays of certain types of nudity), but the policy is immediately expanded with provisions where the platform will permit nudity in user-uploaded content: “We always allow photos of women actively engaged in breastfeeding or showing breasts with post-mastectomy scarring. We also allow photographs of paintings, sculptures, and other art that depicts nude figures. Restrictions on the display of both nudity and sexual activity also apply to digitally created content unless the content is posted for educational, humorous, or satirical purposes”. Hereby, the policy clearly lists the contextual affordances for nudity in the policy; breastfeeding, post-mastectomy scarring, painting, sculptures and art — as well as nudity for the purpose of education, humour and satire. The policy rounds out by explaining: “Explicit images of sexual intercourse are prohibited. Descriptions of sexual acts that go into vivid detail may also be removed”

As users have continued to come into conflict with the *‘formalised’* and *‘amended’* policies as they did with the *‘informal’* and *‘non-’* policies, it would be a to suggest that the changes between these policies are strategic and largely represented a change in policy rationale and communication, both to users and moderators, more than a change in deployment. As the change to the *‘re-categorised’* policy has been made near the conclusion of this research project, it is too early to tell if users will continue to come into conflict with the platform in this area, or if the policy is representative of a wider effort at reforming both the rationale and enforcement of Facebook’s policies affecting nudity and breastfeeding in photographs.

Perspectives of Conflict and Governance

Experiencing the Platform — Censorship and other strategies

For Facebook and most online platforms, censorship is a key (although sometimes contentious) staple of the tools used to moderate and govern the uses of their digital space. In the context of social platforms censorship can be used both overtly and covertly. Overt censorship, which Facebook seems to favour, involves the deletion or removal of content, as well as the notification of stakeholders related to that content — typically the user who created or uploaded the content, as well as the user(s) who reported that content for infringing upon the terms of the platform. Covert censorship, on the other hand, avoids disclosure of the censorship action and involves hidden manipulations akin to ‘gaslighting’ (Stern, 2007) — a style of psychological manipulation tactic wherein the abuser attempts to make a target question their memory, perception or sanity by selectively twisting, omitting or denying information⁶. Reddit’s ‘shadowban’ is a technique of covert censorship that bears some similarity to gaslighting, as it “doesn't notify the user [...] lets them continue to submit [... and] makes it look like they're submitting normally when they're logged in and viewing their content” (Fasnacht, 2015) — so that the affected users remain unaware of their ongoing censorship.

⁶ ‘Gaslighting’ is named for the 1938 play by Patrick Hamilton, in which an overbearing, malicious husband makes his wife anxious by withholding information and convinces her to question her perception of reality by questioning her observation of the flickering and dimming of gas lights.

Of the two, overt censorship is notionally a more transparent method to users (and external observers, like researchers) because of the disclosure involved in its use. This makes it a useful tool for governmental interventions intended to change a user's behaviour, as it details the contentious content or behaviour of the user in relation to the rules or policies that govern them. It could be notionally more efficacious in conflicts where the censored user either agrees with the policy being enforced or sees futility in opposing the policy's enforcement. That said, its efficacy is perhaps limited where users have no interest in obeying the policy.

As indicated, a benefit of this style of censorship — for researchers at least, is its disclosure. For the users who found policies onerous or preposterous, a common recourse was to document their frustrations externally to the platform — often on an alternative social network or a blog. One of the earlier examples of this is of Karen Speed's *Bliss Breastfeeding* blog. In her posts, Karen records her interactions with Facebook moderators from as early as 2007.

Karen's (2007a) posts detail a protracted disagreement with Facebook moderators about a personal image of her 'tandem breastfeeding' two of her children. The photo (Speed, 2007b) seems an odd candidate for removal, considering Facebook's many policies, as both her children and her clothing obscure the majority of her breast flesh, and —critically to the application of earlier policies— neither her nipples nor areolae are visible. Her contact from Facebook moderators seems to indicate a low-cost or low-effort approach to moderation, given the seemingly generic 'boilerplate'-style email she received:

“You uploaded a photo that violates our Terms of Use, and this photo has been removed. Among other things, photos containing nudity, drug use, or other obscene content are not allowed, nor are photos that attack an individual or group. Continued misuse of Facebook's features could result in your account being disabled.”

The broad explanation for removal and the general list of proscribed content seems to indicate a lack of specific attention paid by moderators to individual complaints. Generic statements like these might also account for the perception amongst platform users that moderation is ‘automatic’ or ‘algorithmic’, even where it is not. Karen’s (2007a) later posts describe a follow-up email from Facebook moderators after her photo dispute that threatens the disabling of her account for ‘misusing certain features’ of the Facebook platform:

“Our systems indicate that you've been misusing certain features on the site. This email serves as a warning. Misuse of Facebook's features or violating Facebook's terms of use may result in your account being disabled. Thanks in advance for your understanding and cooperation. If you have any questions, please contact warning@facebook.com from your login email address.”

Karen (2007c) indicates that her account was suspended shortly after receiving this e-mail, and posted her response to Facebook’s moderators publicly in a blog entry;

“I received this email from facebook today. I am very confused as to which services I have been misusing. On the contrary, I feel that as a contributing member

of facebook, I have valuable services to offer, such as a support group for children with cystic fibrosis, an online support group for new mothers having trouble breastfeeding their babies, as well as particiaption [sic] in various charities and causes.

I am very hurt that a moderator would accuse me of misusing facebook services, and I would appreciate a reply with a description of how I am doing so.”

A final post indicates a notionally permanent punishment for Karen (2007d), although little-to-no-clarification as to how she had earned the ire of Facebook in the first place;

“After reviewing your situation, we have determined you violated our Terms of Use. Please note, nudity, drug use, or other obscene content is not allowed on the website. Additionally, we do not allow users to send threatening, obscene, and harassing messages. Unsolicited messages will also not be tolerated. We will not be able to reactivate your account for any reason. This decision is final.”

These interactions are particularly revealing of the attitude of Facebook’s moderation team in some of its earlier years. They reveal an unsophisticated wielding of power, in which Facebook moderators seemed more inclined to punish users for disobedience, than attempt to engage users who disagree, or teach users who are misbehaving. This likely does not help the platform to placate upset users.

Emma, a Breastfeeding advocate who repeatedly had her breastfeeding photos removed by Facebook moderators told me in an interview that;

“I lost my entire account on Facebook, like, four separate times over the years. Shut down, no apologies — just, I was gone from Facebook... called a pornographer and all sorts of things, so... that was a bit shocking as how quickly I was flagged.”

Emma’s experience mirrors that of other mothers who took issue with Facebook’s policies and policing of breastfeeding photographs. It’s worth noting that whilst Facebook’s early policies did not particularly attempt to distinguish between sexualised nudity and non-sexual nudity, the generic, unspecific email admonishments to users are even less careful in describing the behaviour they address. This is particularly visible in Emma’s statement that she was “*called a pornographer*” in her dealings with the platform, as the platform’s contact with many users around this time seems to be characterised by a lack of effort in communication.

Fighting Back — Tactics of User Dissent

Documentation

As noted earlier, a recurring tactic for users attempting to affect change on Facebook includes the documentation of their individual conflicts. From individual users like Karen Speed (2007c) and Rachelle Lesteshen (2013) to institutional stakeholders like the Topfree Equal Rights Association (Rapoport, n.d.) and The Guardian (Reed & Gardiner, 2012), documentation variously includes either copied-and-pasted textual information, and from more technologically savvy users, screenshots.

Documentation is a significant user tactic that can occur both on-platform and off. In documenting their conflict with platform operators, users are able to chronicle and re-tell the story of their conflict and share their perceptions of injustice at the hands of platform operators to other users both on and off the platform. On-platform retelling entails greater risk for users (such as the potential for further censorship or punishment, if their retelling contains examples of the content they were originally punished for), whereas off-platform retelling mitigates the risk of ongoing censorship and punishment, by placing it outside of the ‘jurisdiction’ of platform moderators, so to speak.

Simone, a mother who has experienced censorship of her photos on the Facebook platform told me that:

“Taking screen-shots felt like the most convincing way of showing people how I was being treated, and what had happened. I posted screen-shots hoping that other people would notice, even if they were just my friends.”

Although screenshots can be manipulated or fabricated with sufficient Photoshop or web design expertise, for many users they seem to be a record of legitimacy, a literal recording to say ‘this is what I saw, this is what happened to me’. Though it is not without its flaws, the practice is reasonably widespread, and consistently appears not only in this the tactics of users in this case study, but also in the studies later discussed in this thesis.

Mainstream Media Coverage

Given the protracted length of this issue on Facebook, it would be folly to suggest the mainstream media was not influential to the changes and development of the conflict. In earlier stages, Facebook relied on their conception of media norms in the rationalisation and justification of their policies: “These policies are based on the same standards which apply to television and print media” (Facebook, nd); and in turn, they also seemed to be most receptive to user complaints when those complaints attracted the attention of the mainstream media. As Emma explained;

“I [kept] losing my account and people [on Facebook’s moderation team] wouldn’t talk to me. So I went to the media, and that’s when I got some action. I actually had this conference call with the Facebook team of the people who decide on these bans and the deletions of photos. [...] It wasn’t until it became very public and went to the media that we were able to achieve much concrete action. I don’t regret doing it — but I had to do that to get anywhere.”

It is difficult to ascertain the internal machinations that might prevent or allow policy change in Facebook. It would be reasonably fair to argue that Facebook changes its policies quite conservatively and does not consider policy change until negative attention from users or media makes their policy status quo untenable. Alternatively, issues in enforcement of policies might point to significant structural issues, like a cultural disconnect between users and operators, or different classes of operators, like the low-wage moderators tasked with enforcing the policies designed by Facebook’s policy team. Further research in this field should attempt to more clearly understand the machinations of moderation, policy development and policy enforcement on social media platforms.

Conflict Avoidance

An alternative tactic employed by users embroiled in both user-user and user-platform conflict is the avoidance of conflict altogether. Emma described her attempts to use Facebook in a way that did not attract the ire of Facebook, or of passing ‘troll’ users looking to cause trouble;

“I’ve kind of backed away from my public stance of my page or public groups... I’ve gone into a mothering group — I feel like I can help women in a more meaningful way, one-on-one in a closed group as opposed to being on a public page. Because, I get a lot of trolls — obviously — people that want to cause problems... It doesn’t feel worth it to me to battle this publicly on the internet because people hide behind a screen and say stuff that isn’t really helping, or is just distracting me of what I really need to do.”

This tactic, avoidance, reveals a lot about the enforcement priorities of Facebook as a platform. Facebook (Facebook, nd) claim that their moderation system is reliant on reports from users — and that it does not algorithmically identify offensive content. This is a simplification as it fails to acknowledge Facebook’s use of algorithmic or semi-algorithmic measures for policing copyright-infringing content and content that the platform has deemed inappropriate for the platform (e.g. ‘revenge porn’, where links to many infamous websites such as ‘IsAnybodyUp’ cannot be posted to the Facebook platform at all). That said, a significant part of Facebook’s policing does

indeed rely on the reporting tools built into the platform, and their use in good faith reports. Because of this, users can easily avoid policing by adopting less-public usage habits for Facebook, such as restricting privacy on their user profile, or joining closed or secret groups with likeminded users. Given the ease with which users can evade a significant part of Facebook's policing, it stands to reason that Facebook's policies about appropriate content are closely related to how the platform wishes to be perceived, more so than strict ideological decision about what should or should not be on the platform.

The practice of conflict avoidance is also symptomatic of a conflict where a practicality of the platform's operation is that it doubles as a significant apparatus of surveillance for users. In this apparatus, users are trusted as a small part of platform governance mechanisms given their ability to report content that they deem to violate Facebook's rules. Given this, the option of avoiding conflict by avoiding users becomes a tactic of pragmatic dissent, allowing users to continue their practices without attracting the attention of platform operators.

Evaluation

On Sufficing, and pleasing Stakeholders with Minimum Changes

As noted earlier in this chapter, Facebook has clearly made some changes in response to user criticism. Clearly, the operators of Facebook have noticed the criticism relating to the tone of their policing, and have quite dramatically adjusted their language over time, moving from an accusatory tone, to a complimentary and empathetic tone. Unfortunately, for the users who have criticised Facebook's policies more broadly, the underlying policy and the methods and mechanisms of policing said policy seem largely the same. Fundamentally, Facebook continues to only conditionally allow photographs of the act of breastfeeding, contingent on the obscuration of nipples and areolae. Facebook's lack of transparency in this domain leaves room for observation and speculation as to the methods and motives of their governance policies. The practice of enforcing this policy relatively uniformly across the world speaks to perhaps a globalising impulse or political perspective within the platform, given that the platform seems to ignore local and cultural significances or contexts in its judgement.

Whilst public breastfeeding is common, and relatively non-controversial in some of the countries and areas that Facebook operates, it remains a topic of controversy and consternation in other countries and political contexts. A factor in deciding to deploy a relatively uniform policy may well be the complexities that would be associated with international and cross-culture moderation: global communities that see some content as desirable or delightful might be at odds with other communities

that view the same content as amoral or illegal. Whilst Zuckerberg (Zuckerberg, 2017) acknowledges the complexities of these issues, his proposed solution is for Facebook to provide more individually-focused tools, so that users are empowered to more clearly define the edges of their own filter bubbles, rather than engage critically with the platform and the content shared on it.

Through Facebook's middling response to user conflict, and its user-empowerment solutions, we can see a social platform engaged in a practice of 'sufficing' — taking a slow, minor, and ultimately conservative approach to policy and governance in digital spaces. Whilst the private nature of Facebook means that we are not privy to the reasons for this — there are a handful of possible causes for this sort of conservative approach. Firstly, market pressures: Facebook, as a publicly traded company, must consider its business operations not only in relation to their users and commercial stakeholders, but also in relation to its stock market investors. Given the often quite libertarian philosophical notions associated with Silicon Valley notions of 'the Internet', Facebook's *light touch* approach to digital governance may well be for fear of too radically changing their relationship with users — and avoiding the risk of alienating them. Similarly, Facebook's approach to governance might be particularly modest given the lack of broader regulatory structure governing social media platforms, and their relationships with users. Too strong or overt intervention with the perceived 'rights' of users, may see legal and regulatory challenges from users or policymakers.

Facebook's approach has seen users adopt a range of measures in frustration. For users, opposition to this policy has tended to manifest in both off-platform and on-

platform protests. For some users, particularly earlier in Facebook's public life, off-platform protest in venues like personal blogs, the Facebook campus, and the mainstream media, have been particularly advantageous in drawing attention to areas of frustration with the platform and its policies. Whilst these off-platform protests have limited access to audiences in the platform, they have an analogue in the dissidents of autocratic states, who leverage the comparative safety of exile or refuge in another country to criticise and draw attention to unjust laws in the country they have left.

For other users, the social cost of platform-exile can be too great — as leaving the platform can entail weakening digital connections to friends, colleagues and family. To this end, we have also seen on-platform protest through rule-breaking and outright circumvention of the policies. Some users who have objected to Facebook's policies and policing, have sought to flaunt or ignore the policies — rejecting them as unfair or unjust, with the intention of revealing the policies as such. For other users, the most practical solution is to adopt a more 'covert' use of social media — sharing pictures in private groups, private messages, or with stricter privacy controls.

All of these tactics are indicators of different ways of engaging in conflict with platforms, ranging from idealistic protest and conflict issues, to pragmatic conflict avoidance practices.

¹ The terms ‘topless’ and ‘topfree’ are often used interchangeably, however — ‘topfree’ is used by some advocates as a way of distancing the freedom of dress from the term ‘topless’, which is considered more of a sexualised term given its relationship to ‘topless dancers’ and other adult entertainers whose chests are bared for the purposes of titillation.

² How, for instance, does Facebook define a ‘fetish’? Similarly, if the digital alteration of images was *verboten*, then why were advertisers not struggling with the platform’s censorship more frequently?

³ Raëlism or Raëlianism is a ‘UFO’ (Unidentified Flying Object) religion founded in 1974 and based on the claims of former automobile journalist and race car driver Claude Vorilhon.

⁴ Also erroneously attributed, given that nudity is not particularly uncommon in print media (where it semi-regularly appears in Fashion, Health and News magazines) or screen media (where it appears semi-regularly in both drama and documentary film and television (regional censor regimes notwithstanding.)

⁵ It is named for the 1938 play by Patrick Hamilton, in which an overbearing husband makes his wife anxious by withholding information and convinces her to question her perception of reality by questioning her observation of the flickering and dimming of gas lights.

Chapter 5: Case Study Two: Google+ and “Real Names”

During the design of this project, in 2013, Google+ seemed something of an ‘underdog’ platform – a social platform designed in competition to Facebook, and although it was not widely used, it offered some innovative features to users, and had some dedicated users. Over the interceding years, Google+ has become comparatively stagnant, and at the time of writing is no longer seen a particularly ‘exciting’ platform.

As of October 2018, Google has announced plans to end service for Google+, after a design flaw in Google+’s API exposed the data of hundreds of thousands of platform users. Observers noted that Google may be concerned about the regulatory scrutiny as a result of the privacy issues caused by the platform API (Wong and Solon, 2018).

Whilst Google+ is soon to be a defunct platform, and never saw the societal importance of platforms like Facebook or Twitter, it remains an important platform for study and research, particularly given this issue, the mandatory requirement of a ‘real name’ identity, faced by Google+’s operators, and the reoccurrence of similar ‘real name’ issues on other platforms, both before and after Google+.

Platform–Political Tensions

This chapter analyses the conflict surrounding identity, and in particular, the notion of ‘Real Name’ identities on the Google+ platform. To begin, I sketch the key notions of named identities, Anthroponyms and Pseudonyms, and the use these identities in Internet spaces and social media platforms. Later in this chapter, I examine

the relatively short conflict, the coverage of the conflict in industry and mainstream press, and the influence of platform strategies and user tactics in this conflict.

What are Anthroponyms and Pseudonyms?

An anthroponym, to quote the Oxford English Dictionary (OED), is ‘a personal name’. The term ‘anthroponym’ is generally used to refer to a person’s surname or given name. An anthroponym can be a birth name, a legal name, a mononym, and even pseudonym or stage name that a person is commonly known by. That said, an anthroponym does not particularly need to be a ‘legal’ name (by which a person is known to their government) or a ‘birth’ name (often conferred by parents or carers shortly after the birth of a child), nor does it need to conform to common standards of naming.

I use this term in avoidance of the more simplistic descriptions of ‘real names’ or ‘authentic names’ used by many of the sources cited in this case study. This is in part because the terms ‘real name’ and ‘authentic name’ make implicit value judgements about the use of pseudonyms — and situate pseudonyms in opposition as ‘not real’ or ‘inauthentic’. These judgements may well stem from the historic meanings of pseudonym — ‘a fictitious name’, according to the OED— which originated from the Greek *pseudōnymos*, and the root terms *pseudēs* ‘false’ and *onoma* ‘name’. To continue to situate the identities constructed with and around pseudonym as ‘false’, or less worthy of respect than anthroponymous identities is an unfair characterisation that does a disservice to the varied ways in which people can present and construct their identities in both physical and digital spaces.

The ways in which humans identify themselves both online and offline are many, varied and complex. There are many boundary cases — where a person’s identity does not cleanly map to the norms or expectations of their communities. In this chapter, I largely engage with the normative western expectations of anthroponymity (generally combinations of first and surnames, conferred by parents and recognised by the state), internet-related practices of pseudonymity, and internet-specific affordances of anonymity.

Identity on the Internet

Anthroponymous, Pseudonymous, and Anonymous spaces

For decades, internet cultures have endorsed and encouraged the use of ‘user names’ and ‘handles’ — abstracted identities in the form of pseudonyms that may or may not be directly related to a person’s offline identity (Moll, 2014). Pre-internet and early-internet networked spaces, like Usenet networks, bulletin board systems (BBS), and gaming systems like Multi-User Dungeons (MUDS) often allowed users a flexibility of identity — the permission and encouragement to construct identities with pseudonyms. These identities and pseudonyms ostensibly offered their users a freedom to participate in the network without fear of external social repercussion outside of these digital spaces.

In contemporary internet usage, spaces that provide comprehensive affordances for anonymity are rare somewhat rare. Some spaces of this nature are ‘imageboards’.

Imageboards are a style of Bulletin Board system that requires the initial post of a thread to include an image, and that tend to include an ephemerality in their design — often permanently (and automatically) deleting discussion threads that are not actively being participated in. The design of these imageboards is inspired by 2channel (also known as ‘2ch’), a popular Japanese textboard and Futaba Channel (also known as ‘2chan’), an Otaku⁷ subculture imageboard that permits explicit content.

The most popular western Imageboards tend to be inspired by 2channel and Futaba Channel, with significant focuses on Japanese popular culture — especially, video games, cosplay, anime, manga, and hentai⁸. Amongst these imageboards, 4chan is the best-known — particularly for its relevance in internet culture as a spawning pool for popular memes (such as lolcats and rickrolling), and its prominence in playing host to the loosely-organised “Anonymous” activist group(s) and their campaigns against the Church of Scientology and other targets. Derivative imageboards 8chan (also known as ‘infinitechan’) and Wizardchan have risen in alternative to 4chan, because of perceived issues with the platform’s management (Machkovech, 2015a), and culture (R. Jones, 2014) respectively. These platforms often keep limited to no traffic logs for visitors and posters on the site and require no registration or username

⁷ ‘Otaku’ (おたく / オタク) is a Japanese term for people with obsessive interests, especially in relation to anime and manga. It has historically been used as a pejorative term due to negative stereotypes, and association with Tsutomu Miyazaki, a serial killer, cannibal and necrophile described as ‘The Otaku Murder’ in Japanese mass media. In recent years, the term has lost its stigma, and many Japanese people self-identify as otaku. It still bears derogatory and self-deprecating connotations and would be harsher than the similar terms like ‘geek’ or ‘nerd’ in English.

⁸ ‘Hentai’ is the English genre categorisation of erotic or pornographic manga, anime and computer games. It is short for *hentai seiyoku* (変態性欲), which translates to ‘a perverse sexual desire’. However, the term’s narrow Japanese use (as a description of a perverse act) and catch-all English use (as a broad genre of published material) are generally incompatible. Thus, it would be incorrect to assume that the material that English-speaking audiences categorise as ‘hentai’ would necessarily be categorised as ‘perverse’ by Japanese audiences.

to post or access content. An affordance of limited pseudonymity, whereby users can easily construct temporary identities, is often available in the guise of ‘Tripcodes’ — the cryptographically hashed result of a user’s password string that is computationally difficult for other users to spoof. These tripcodes are available to all users, as a text input prompt. The text prompt is converted as by a complex tripcode algorithm, which does not store any identifiable information on the imageboard’s server. For example, 4chan founder Christopher ‘moot’ Poole’s tripcode is “*!Ep8pui8Vw2*”. Poole’s use of a tripcode for persistent identity in conjunction with his ‘moot’ pseudonym is atypical of tripcode usage on imageboards. Generally, tripcodes are not parts of constructed identities, but instead are a functional affordance for declaring a singular author across multiple anonymous posts. An example of this might include a user seeking advice or offering services to other users. This allows users to have their needs of anonymity served by the platform, but it also makes an affordance for semi-persistent identity in the cases in which this is needed.

On digital spaces where identities are required, there are very few internet platforms that mandate pseudonymity. For the most part, social platforms that do not require anthroponyms or ‘real names’ will allow users to use either an anthroponym or a pseudonym at their own discretion. Twitter, for example allows anthroponymous and pseudonymous use, at the user’s choice. Similarly, Twitter allows for a flexibility of -onymous uses, by allowing users to change both their user name and display name at will. For some platforms, particularly in subcultural spaces (like those associated with gaming) there are well-established cultural norms that expect the use of pseudonyms. This flexibility of identity affordances can highlight the purposes that

pseudonymity and anonymity serve for users, in allowing varied consistencies in identities to suit their environment.

One of the most well-known social platforms that mandates the use of Anthroponyms is Facebook, although other platforms, like Google+ (discussed at length in this chapter) and Sina Weibo also makes use of ‘real name policies’ to require users to present their ‘authentic identities’ online (Moore, 2012). Part of the reason for this design choice is the way in which social platforms like Facebook attempt to replicate existing social conventions and norms and insert their platform into the daily life and social connections of their users. As boyd (boyd, 2012) notes, in Facebook’s early days as a Harvard-specific social networking website, users participated with their anthroponyms or ‘real names’ because “they saw the site as an extension of campus life”. Whilst ‘real name’ policies are often touted as a boon for user safety, and a civilising force for users behaving badly (Zuckerberg in E. Galperin, 2011a), these justifications often rely on folk wisdom and opinion rather than any substantive research from platform operators (Rosenbloom, 2014).

‘Authentic’ Identities, and ‘Real Life’ online

Google+, Facebook, Twitter, or any social platform are important arenas for people choosing to represent their identity online. As Haimson and Hoffman (2016) argue, platforms like Google+ and Facebook in particular, actively position themselves as “a kind of online identity registrar — a “core social infrastructure” for the Web where users express their authentic ‘real selves’.” I’ve discussed how this dichotomy

of real and (implied) false identity is generally reductive. In particular, Marwick's (2013 356-7) discussion of identity is a useful counter-point to the notion of an authentic identity espoused by social platforms. Marwick builds on, and modernises, Erving Goffman's framework for contextualised identities — noting that “Identity is flexible and changeable, and people are highly skilled in varying their self-presentation appropriately. Identity is thus socially constructed in tandem with the people around us.”.

An interesting element to explore from Haimson and Hoffman's (2016) analysis is this notion of identity infrastructure and identity registrars. Considering social networking services as providing a regulatory role for identity online serves to underscore the prevalence of conflicts in this area, as the regulatory role platform operators take in turn enforces normative restrictions and policies upon users and their ability to express their identity.

Furthermore, users who have are experiencing significant life transitions (for example, gender transitions, survivors of domestic abuse — or people undertaking the radical changes of identity often required in Witness Protection programmes) are most likely to bear the brunt of a normative policy's reductive expectations of an ‘authentic identity’ (Haimson and Hoffman 2016). Of course, the platform doesn't *have* to be this way, enforcing normative ideas of identity on users. It has been designed and implemented this way, whether through assumption, or deliberate intent — but the relative lack of regulation or standardisation in this space means that a platform like Google+ could have taken a more expansive, exploratory approach to user identity — rather than integrating their assumptions and enforcing them heavy-handedly.

-Nyms on Social Networks

Twitter, Tumblr, Reddit and -nym friendly platforms

The cultures and affordances supporting the use of pseudonyms are available in many of the web's popular western social platforms, but perhaps most clearly in the microblogging platform Twitter, the blogging platform Tumblr and the forum platform Reddit. In each of these platforms, the cultures and affordances are different, lending to different styles of identity management.

Twitter's affordances for -onymous identity are both simple and flexible, a text field that can be used for an anthroponym or pseudonym. Whilst many users engage with Twitter under their anthroponym, there's a culture of pseudonymous use too. Pseudonymous can include parody impersonations, including users who create 'fake' satirical accounts representing politicians and public figures (J. Wilson, 2011). These uses are generally allowed by Twitter's rules, so long as the impersonation is not malicious and deceptive. Similarly, there is a 'bot' culture, wherein Twitter accounts are not representative of people, and are instead avatars for code of varying complexity. Artist Darius Kazemi's Twitter bots range from Kenosha Kid (@YouNeverDidThe), a bot which riffs upon a linguistic experiment from the Thomas Pynchon novel *Gravity's Rainbow*, to @TwoHeadlines, a bot which mixes two news article headlines for comic effect. Similarly, the infamous @horse_ebooks account, which purported to be a bot, amassed over 200,000 followers by tweeting unintentionally amusing non sequiturs, and was eventually revealed as part of a performance art project (Orlean, 2013).

Both Twitter's affordances and cultures in relation to pseudonymous use are broadly permissive of a range of identities that a person might use. And whilst Twitter accounts do not feature the ability to alternate between multiple identities, many twitter clients — including the official Twitter apps — include this affordance.

In Tumblr, the understanding that users may want or need to present a multitude of identities is underscored by its inclusion in the platform's affordances. Tumblr users can register and maintain multiple blogs and blogging identities under a single user account. In addition, the platform does not stress a *need* or requirement for social interconnection. For some users, a pseudonymous blogging identity can provide a style of functional anonymity through the obscurity provided by participation in a massively popular social platform like Tumblr.

Moreover, the 'Ask' function, which allows a Tumblr blogger to ask other bloggers questions (to be answered in the form of a public blog post) can allow users to ask 'anonymously', at the discretion of the blogger receiving the question. This distinction is made clear to the questioner, so they can clearly understand if their current Tumblr identity will be revealed to the blogger receiving their 'ask'.

This, along with Tumblr's loose restrictions adult content, likely contributes to the proliferation of Tumblr's 'NSFW' blogging community which reportedly makes up approximately 11.5% of the platform's most popular blogs (Perez, 2013). Tumblr's 'NSFW' affordances are relatively important to its users and attracted significant

media attention in 2013 when Yahoo acquired Tumblr. CEO Marissa Mayer publicly discussed the adult uses of its platform, addressing concerns that Yahoo might “screw [Tumblr] up”, advocating that the diversity of content on Tumblr allows the platform to reach a lot of users (Bosker, 2013).

For Reddit and Reddit users, the expectations surrounding identity are complex. However, a significant and vocal part of the Reddit community constructs Reddit identities as something of a shield identity, that allows users to more freely participate in the platform without fear of punishment or retribution. Disposable ‘throwaway account’ identities are also particularly common, especially in relation to uses of Reddit seeking sex or relationship advice, confessional stories, the sharing of sexualised content, and both belligerent and inquisitive ventures into political or ideological discussions.

In the past, Reddit has philosophically positioned itself as a platform for free speech, largely unfettered by the owners of the platform. This promise is ostensibly true but mitigated by the design of the platform — which rather than operating in the open, free-for-all engagement style of Twitter instead allows users to participate in a host of federated micro-platforms in the guise of ‘subreddits’. Each subreddit has super-users or ‘moderators’, who are users endowed with the ability to moderate, edit, censor, and ban other users from participation in their subreddit. Each Reddit user can create their own subreddits, in which they have these powers, or earn these powers through participation and acknowledgement from the operators of other subreddits. In this respect, the practicality of Reddit’s ‘free speech’ is complicated, to say the least, as the freedom is conditional upon whims of the operators in each subreddit board (Robertson, 2015).

There is however, a strong cultural appreciation of the value of free speech, and an accompanying understanding of the importance of pseudonymity and anonymity in facilitating unpopular or undesired speech. This can be seen in the Reddit community's relatively strict rules about 'doxing' redditors or revealing the identity of Reddit users publicly. This has been most notably demonstrated in the backlash against Gawker blogs and publications after then Gawker journalist Adrian Chen published the name of Reddit user 'violentacrez', who was a creator and/or moderator for a number of highly objectionable subreddits, including r/chokeabitch, r/niggerjailbait, r/rapebait, r/Hitler, r/jewmerica, r/misogyny, r/incest (Chen, 2012b). During the backlash, the moderators of many of Reddit's most popular forums disallowed linking to Gawker and Gawker-related websites (such as Kotaku, Jezebel, Gizmodo, Lifehacker, and io9) in solidarity with the named-and-shamed redditor, and in opposition to the way that Chen had doxed 'violentacrez'.

Facebook and mandated authenticity of identity

Facebook's 'real names' policy largely metastasised out of two factors, the social norms of its early users who self-represented with their anthroponyms for the ease of replicating class and college social connections, and the ideology of its creators like Zuckerberg who has rejected privacy as an "outdated social norm" (Zimmer, 2010), arguing that Facebook's successes reflected a move towards "radical transparency". Facebook's favour for what they have termed as "real" or "authentic" anthroponymous identities is perhaps symbolic of Zuckerberg's own beliefs about identity, given his

argument that “having two identities for yourself is an example of a lack of integrity” (in Helft, 2011).

As such, both Facebook’s operators and Facebook’s affordances eschew the complexities of human identities, and the multitudes of differing identities that people can present in their everyday lives, amongst friends, family, co-workers and strangers.

Worse yet, Facebook’s eschewing of multitudes of identities and the complexities of identity management through ‘real name’ and ‘authentic name’ policies ignore the people for whom the management of identities is a survival tactic. These people, who range from LGBTIQ people in fear of rejection or violence from their families or communities, and dissidents who fear state violence, are disadvantaged by Facebook’s policies, which in turn contribute to existing regimes of censorship and oppression. By mandating the ‘authentic’, ‘real name’ identities, Facebook requires that its users either expose themselves to risk of harms, surveillance or retribution from the state authorities or social belligerents, or that they cease participation in fear for their own safety (boyd, 2011).

In recent years, Facebook’s “real name” policy has come under criticism for these very reasons, which has resulted in some changes to the policy — requiring “authentic names” rather than “real names”. This distinction still requires users to self-represent with a singular, primary identity used in day-to-day life, but is designed to ameliorate issues surrounding birth or legal names being different to the names by a person used in day-to-day life. The policy’s enforcement however, is not meaningfully

changed, as users who are deemed to have breached this policy often have their accounts suspended until they can provide legally recognised identification to Facebook as proof of their ‘authentic name’.

On the Imperatives for Singular ‘Real’ IDs

As noted earlier, the role that Google (and other competing platforms) seem to want to offer in this space is not too far removed from that of an Identity Registrar, so it is likely that there is an underlying imperative that would be serviced by a platform operating in this kind of role. Given this, it is unsurprising that a recurring theme amongst user criticisms and concerns levied at ‘Real Name’ systems relate to the imperatives perceived by users for the deployment of these systems. Firstly, an economic imperative (especially given Google’s reliance on advertising as an income stream) — that advertisers or other third parties will find data associated with personal identifiers to be particularly more valuable than data associated with pseudonymous identifiers, or other anonymised data — and secondly, a political imperative — that ‘real’ or ‘legal’ name systems will allow for stakeholders like law enforcement agencies, and governmental operators to more easily access and enact surveillance on citizens either through cooperation with platform operators.

Whilst Google’s responses to criticisms did not particularly engage with either of these perceived imperatives, much of the coverage of the Google+ ‘real name’ conflict in industry press argued that advertisers saw more value in ‘real-name’ data (Ingram, 2011), and that advertiser desires for rich highly-individualised audience and

customer profiling could be a key element of Google+'s decision-making, given the Google search platform's significant advertising revenues.

The argument that platform data associated with legal identities is more valuable and compelling to external governmental operators, is much more compelling, although, there is scant evidence to support user speculation here. Whilst many platforms facilitate government requests or interventions at the request of law enforcement or a court order, no evidence has yet come to light revealing the structure of these platforms and their governance to be particularly shaped by the needs or desires of traditional governmental actors.

From a critical standpoint, the argument that this data is more valuable to advertisers is not particularly compelling — detailed personalisation could be meaningfully accomplished without a 'real' or 'legal' identifier. A pseudonym, or even a notionally anonymised identifier could still be used to organise data usually captured to profile audiences and prospective customers. An argument, well outside of the scope of this research, could even be made that data associated with anonymous or pseudonymous users could be more 'accurate' if users believe they are not being observed, or surveyed in a manner that links their online behaviour to their legal identity.

We can see the advent of this sort of data collection at work in Facebook's "Shadow Profile" system (Quodling, 2018), where Facebook collects and associates social graph data for users who have not joined the platform based on the already-

shared data of their friends who have joined Facebook. Critically, this is pre-consent, as a user has not agreed to use the platform, much less to have data about them used by the platform. That said, the use of this system and systems like it, point to lack of need for ‘real names’ at all in this space. So, we have to ask — why is this a big deal? For users, the answer is readily apparent — their names are a key facet of their identities, but for platform operators — is there a meaningful difference between ‘John Smith’ and ‘Jane Doe’, or a long, unpronounceable hexadecimal hash code used as an identifier?

Platform History and the Conflict

The Launch of Google+

Google+ was launched as a social platform in December 2011. It is Google's fourth social networking platform, after Google *Friend Connect*, the short-lived Google *Buzz*, and the acquisition of *Orkut* by Google Brazil. Its shared space in Google's wider array of services with a number of other attempts at identity management initiatives, including Google *Profiles*.

The platform was initially quite highly anticipated, especially by users who found Facebook's policies regarding identity and privacy frustrating (E. Galperin, 2011b), and had expected a more sophisticated system from Google. During the platform's invite-only beta phase, before the platform's release, some users were frustrated by the platform's inflexibility for gender identity, as the platform only allowed users to identify as 'Male', 'Female' 'or other', and it required all users to have this gender choice publicly displayed on their Google+ profile. After an outburst of complaint from users, Google quickly changed the Google+ platform to allow users to keep their gender private, and to use singular 'they' pronouns for users with their gender set to private (McCullagh, 2011).

In July of 2011, months before the public release of Google+, users began complaining of a "crackdown" on pseudonymous usage of the platform, and an enforcement of a policy that mandated the use of 'real', 'daily life' names. Users reported that the enforcement of the policy restricted the use to of their Google account

and its related products and services, including Gmail, Google Reader (a RSS newsreader client), the Blogger weblog platform, the ‘Docs’ collaborative office platform, as well as both Youtube and Google+ (GrrlScientist, 2011).

During this time, a number of high-profile Google+ users, including journalists, celebrities, and tech entrepreneurs found that they had been temporarily banned from the service, and in some cases, asked to change their names for not conforming to the Google+ Names Policy. In some cases, the users blogged or wrote about their experience with Google+, often criticising both the policy of requiring “the name by which you are commonly known”, and the enforcement of the policy. The policy continued to draw strong criticism in industry press (Blue, 2011), popular blogging websites (Doctorow, 2011), as well as industry bodies like the Electronic Frontiers Foundation (E. Galperin, 2011b).

As Limor Fried (Fried, 2011), a founder of the open-source electronics hardware company Adafruit Industries, demonstrates, Google+’s policy enforcement lacked finesse — or indeed, any user-facing information to explain their punishment. She wrote “Google suspended my Limor Fried ‘Ladyada’ Google+ profile, no show-and-tell tonight... It’s my real name, a picture of me – my posts and my nickname. I made sure it wasn’t [the company name] ‘Adafruit Industries’ and specifically me when google asked people to only use real names in their profiles. No explanation or reason why, just nuked.” Fried’s profile was restored in a similar manner, without explanation, leaving commenters to speculate as to the reasons for its removal. Actor William Shatner’s profile was similarly affected (Wauters, 2011), leaving the actor confused shortly after joining Google+: “My Google+ account was flagged for

violating standards. Saying hello to everyone apparently is against the rules maybe I should say goodbye?”.

Later in July, the Vice President of Product for Google+, Bradley Horowitz (Horowitz, 2012) to address some of the issues causing discontent amongst Google+ users. Horowitz offered that the platform’s system would be reformed in a few key ways: by providing warnings instead of outright bans for violations of the platform’s policies; by providing clearer instructions during the sign-up process; and by adding formal support for optional ‘nicknames’. Horowitz made no comment or judgement on the criticisms from users who desired the ability to use pseudonyms on the platform.

Google+’s “real name policy” would become a founding part of the platform’s governance. That being said, calling it a “real name policy” is something of a misnomer, because unlike Facebook’s policies at the time, the Google+ policy formally allowed the use of pseudonyms — with a condition. The conditional policy allowed users to use pseudonyms on the service, so long as the pseudonym is established elsewhere, and the user is commonly known by it. This assessment is left to Google+ moderation teams, and the enforcements seemed arbitrary to users.

Coverage in Press

The conflict drew much coverage amongst tech press and tech enthusiast websites, and eventually more mainstream news outlets. The conflict eventually grew to be known as ‘nymwars’, a neologism that was frequently used in hashtags relating

to discussions and criticisms of Google+'s "real name policy" on Twitter. Across both tech and mainstream reporting, the responses to Google's policy were overwhelmingly similar — a rejection of the policy and its goals. In WIRED, Carmody (2011) wrote of the policy that it represented a business model for the regulation of people's identities, noting that the technological affordances of platforms like Google+ "solicits information from you that, were it viewable by everyone in your networks, you would most likely keep to yourself" with the promise of advanced privacy and identity management. Madrigal (Madrigal, 2011) argued that the adoption and promotion of 'real name' policies represents a "radical departure from the way that identity and speech interact in the real world." Leaving aside Madrigal's distinction between social platforms and 'the real world'⁹, Madrigal's attempt to puncture the narrative of "real names" as similar to existing analogue experiences of society is a useful reminder — after all, users having a discussion on a platform like Google+ will gain access to each other's social graph, and public data — a similar discussion in a public space like a café would likely begin with participants knowing much less about each other. Whilst a platformed interlocutor can see your profile, posts, and friends — an interlocutor in the public sphere of a bar or café would be unlikely to know much more than a name, if you've offered one.

The conflict grew strongly around this time, as Google's "new social network" had been relatively highly anticipated amongst Google fans, and internet users who had begun to grow uneasy with the growth and prominence of Facebook. In particular,

⁹ Social Media platforms, and Internet Services exist in the world, and are thus a part of 'the real world'. The popular dichotomy between 'cyberspace' and 'real life' is a false dichotomy that often serves sensationalise and undermine solutions to social, cultural and technological issues related to the internet.

the conflict saw a lot of coverage in tech-related blogs like *Boing Boing* and *Ars Technica*. Writers at both of these publications were quite strongly resistant to the ‘real name’ ideology of the platform. *Boing Boing* in particular often imagined users in the position of superheroes in comic books — relating the policies to the comic-book trauma of Batman or Spider-man being unmasked as Bruce Wayne or Peter Parker respectively. It was similarly around this time that the ‘Nymwars’ moniker was coined as a succinct descriptor of the conflict.

Google+’s Eventual Concessions and fall into Obscurity

Years later, Google+ eventually conceded the policy, allowing users to go by the names or pseudonyms they desired. (Quodling, 2014) This concession seemed to be seen as “too little, too late”, as the platform had failed to attract a significant and regular user base. Unfortunately, the changes to the policy were not accompanied at all by any change to technological affordances. Because of this, many of the issues raised during the conflict stayed un-resolved, as the Google+ system essentially only allowed pseudonyms that conformed to Google’s relatively close-minded proscriptions for names — that they exist in pairs of ‘First’ and ‘Last’ names, and only contain certain patterns or characters. This left people with atypical legal names, (and of course, atypical pseudonyms) with the same ongoing frustrations with the platform’s policy and the affordances that the policy governed.

Although Google+ showed some technological promise during its beta, it ultimately seemed to fail to live up to the expectations both of its users and of its

operators. Usage data has often been difficult to find for the platform, as the notion of counting or identifying active Google+ accounts is somewhat complicated by Google+'s integration into other Google services, like Gmail and Hangout's profiles systems, and Youtube's commenting systems.

Google+ seemed to operate without project leaders, after the departure of prominent Google Executives associated with the project. Whilst outside of the scope of this case, there has been some discussion relating to Google+ and a recent Google hire — 4chan founder Christopher 'moot' Poole. Time will tell whether or not Google learns or interrogates Poole's radically different perspectives on names policies and affordances for anonymity.

Governance Context

Google+ as Platform

Policies and Affordances of Identity

Google+'s "Real Name" policy, was enshrined in the support documents for the platform in 2011, when it read:

"Google+ makes connecting with people on the web more like connecting with people in the real world. Because of this, it's important to use your common name so that the people you want to connect with can find you.

Your common name is the name your friends, family or co-workers usually call you. For example, if your legal name is Charles Jones Jr. but you normally use Chuck Jones or Junior Jones, any of these would be acceptable."

For some critics, Google+'s above policy was an example of a company that had more regard for their commercial business interests than for the safety and wellbeing of their users. boyd (2011) argued that these policies did little to empower the users of these platforms, and instead were "an authoritarian assertion of power over vulnerable people". As boyd notes; "many people are far less safe when they are identifiable — and those who are least safe are often those who are most vulnerable." Critics of the policy also noted that Google+'s deployment of a "real name" policy betrayed an unfamiliarity with the social context of these policies in online spaces, and an ignorance of the history of these issues in online spaces.

After Google+ eventually conceded the aforementioned policy, it changed to the shorter, less restrictive text that follows:

“Google+ profiles are meant for individual people. That's why we recommend using your first and last name on your profile. It will help friends and family find you online, and help you connect with people you know.”

Whilst the latter policy still contains the Google+ “recommendations” and therefore Google+'s implied value judgements about legal or ‘real’ names, it is a significantly more flexible policy, as it offers a suggestion, rather than restrictions.

The affordances of Google's identity-related systems have been rather inflexible, and rely primarily on common, normative assumptions about western names. These assumptions include the notion that people have no less than two names, often a ‘first’ given name and a ‘last’ family name. Whilst this format is particularly common, it is not a comprehensive standard for people's names across the world. Indeed, many people have legal, ‘real’ names that do not conform to the norms expected by Google+'s software engineers. Even after Google+'s policy changes to allow for the use of pseudonyms, these affordances largely stay the same — with the presumption that pseudonyms will follow the same conventions that Google engineers had designed their systems for.

Subversions and Reactions to Google +'s Affordances

Google’s policing of user identities generally saw two key responses deployed by users — attempts at circumvention, and abandonment.

Early attempts at circumvention tended to involve attempts to simply shape pseudonyms or legal but non-conforming names to suit the ‘first name’, ‘last name’ schema required by the platform. For mononymous users, this sometimes involved repeating the mononym in both text fields or putting unrelated punctuation marks in one of the fields. Users with pseudonyms that conformed to the schema were in for an easier task, but report from the time indicate that Google+ moderators tended to be relatively zealous in policing names on the platform.

Because there’s no regulatory mandate for auditing or transparency in moderation processes, it is difficult to make judgements as to why and how users were targeted and policed in regard to these policies. We do not know, for example, if seemingly pseudonymous users were identified and reported by vigilant Google+ users, or if they were identified by Google+ engineers or support staff. If the latter was the case — it is still not possible to know if these users were identified systematically, or simply by chance — and to what extent the engineers or support staff did any due diligence on uncommon names before punishing or policing users.

For some users, the policy implemented by Google+ and the strict enforcement of this policy was alienating. For these users, the platform became undesirable, and

ultimately became a platform that they either abandoned, or abandoned hope of joining.

Google+ as a component of the Google ecosystem of platforms

Google+ as component — Identity Management in +, Mail, Youtube

Part of the drive for Google+'s "real name" policy may well have been an internal drive for Google to leverage its products into a more thorough and detailed system for identity management — a system that, like the then new OpenID framework, allowed for a single user account to be used to access and customise a range of products. In the Google stable, this included the management of identity across Gmail, Google Talk, Google Hangouts, Google Drive, the Google 'Docs' productivity suite, Youtube and the Android mobile platform.

Beyond the initial focus on using Google+ as a key component in an Identity Management system, Google+ eventually become a testing and deployment ground for features in other Google services. Often, new features for Google Photos or Youtube are deployed for testing on the Google+ platform, rather than on the more popular Photos and Youtube platforms.

Perspectives of Conflict and Governance

User Perspectives — Identities and Policy Flaws

When I interviewed Kaliya, a Google+ user who had been affected by Google's "real names" policy during the course of this research she described Google+'s policing as a highly determinative 'shoot first, ask questions later (or not at all)'-style practice. For Kaliya, the tone and severity of Google+'s style of enforcement was a point of particular frustration, as the strength of Google+'s response (disabling access to her Google+ account and advising that she had breached their terms) was disproportionate to her action (attempting to use Google+ with a name that she considered to be both hers, and real).

Kaliya documented her experiences with Google+ on her professional blog, where she has developed a profile as an industry expert on matters of Online Identity. She wrote several times about experiencing this from Google+, and her frustrations with the policy and enforcement of the policy. In one particular post, she annotated a formal notification from Google+ that instructed her that "The name you use violates our community standards" (Kaliya, 2011).

Google+ staff largely responded to individual cases in this conflict in the rote style of customer-support attendants without the ability to make or influence policy. Google+'s team was not meaningfully consultative with users and seemed to avoid the notion of consultation altogether when prompted by users to start a dialogue about identity on the internet. This, for people like Kaliya, was a critical failure of Google+

as a platform — as the Google+ had essentially coded normative western expectations of ‘name’s and identity into platform’s affordances. Given this, there was a large chorus of prospective users who’d no major qualms with Google+’s policies, and every intention of complying — but had become frustrated with the platform’s names policies, and the platform’s lack of understanding for both non-western and less-common practices of identity.

Kaliya noted in her discussion with me that the policy seemed controversial amongst the people working at Google, describing a handful of instances of “backchannel” correspondence, where Google employees, especially those not associated with the Google+ team, reached out to her personally or in confidence, to extend their support for her campaign, and their frustration with the policy implemented for Google+. This is a particularly important, if somewhat unsurprising revelation — as it reveals what we might have already expected about the organisations that operate platforms. Platforms, platform operators, and the commercial organisations surrounding platforms are not monolithic. Whilst these organisations may often present highly prescriptive policies as *de-facto* laws, these policies may also be subject to internal scrutiny, or frustrations in their development. They may even be subject to even internal dissent should a stakeholder like a moderator choose to disregard certain guidelines or embrace a more vigilante-style course of action, by acting beyond their ambit, and censoring social content that is not prohibited by platform rules.

For Kaliya, this seemed like a poisoning of the well for Google+ users. In our discussion, she noted that the lead-up to Google+’s public availability was

characterised by frustration and anxiety about Facebook’s growing role and presence as a Social platform. Google+ had, at the time, something of a reputation as an ‘enlightened’, socially conscious and ethically conscientious technology company — so at the time, prospective Google+ users like Kaliya seemed to relish the possibilities of a ‘better’ Facebook-style platform. Indeed, much of the tech press preceding Google+’s public release characterised the platform as a potential ‘Facebook-Killer’, a web app that could unseat the dominance of Facebook’s platform.

During this conflict, Google+ head Vivek ‘Vic’ Gundotra was a particular focus point for criticism, given his prominent role at in guiding and operating the platform.

Whilst the earliest communications with users about Google+’s naming policies were relatively simply put: “To help fight spam and prevent fake profiles, use the name your friends, family or co-workers usually call you” (Google, nd), later communications stressed notions the idea that by mandating “real names”, Google+ was working to foster a “real” and trustworthy community — although Google operators seemed either ignorant or indifferent to user criticisms about this position, as it created a corollary implication that users who were not using their “real names” were not “real” or “trustworthy”.

When Google+ eventually announced the end of this policy, they acknowledged that the policy was isolating to people some people, although mischaracterised their grievances by simplifying the conflict to one around pseudonymy:

“When we launched Google+ over three years ago, we had a lot of restrictions on what name you could use on your profile. This helped create a community made up of real people, but it also excluded a number of people who wanted to be part of it without using their real names.”

In a broadly apologetic statement, Google+ also offered that “Today, we are taking the last step: there are no more restrictions on what name you can use.” (Google+ 2014) This is perhaps true at an institutional policy level, but the technological vestiges of this policy remain in place — namely, the platform’s requirement for both ‘first’ and ‘last’ names. This continued to be a frustration for prospective users — that the platform, despite professions to allow names other than “real names”, still restricted users to the platform operator’s normative expectations of identity.

Relatively promptly after the announcement, Google+ head Vic Gundotra left the Google+ team and Google more broadly. Google co-founder Sergey Brin characterised both Gundotra and his own involvement in Google+ as a mistake, admitting in an interview with industry press that he (Brin) was “kind of a weirdo” and that, “it was probably a mistake for me to be working on anything tangentially related to social to begin with” (Blue, 2014). Brin’s *mea culpa* here echoes the frustrations of many Silicon Valley critics — that the mantra of “move fast and break things” is profoundly irresponsible for enormous companies whose products and services intersect with the lives of millions of people.

Evaluation

Ultimately, this case study highlights a notable issue with social platforms — that they are often designed around the perspectives of the engineers that build them. Ullman (1995) terms this as the ‘cult of the engineer’ — “assumptions and presumptions”, designs and systems that “sum up the intelligence and intentions of all the engineers who worked on the system”. The nature of many social media platforms often means that there is are particular brands of *-centrism* in the design of these: Western-centrism, USA-centrism, and even Silicon Valley-centrism.

Whilst these engineers may be profoundly intelligent at the level of software engineering and technological work, they may not necessarily have broader sociological understandings of the ways in which people interact with each other and the world around them in cultures and communities around the world. Whilst these issues can become particularly clear at the level of software engineers, where technological affordances are designed for software platforms, these issues are no less significant throughout other the companies operating these platforms. From Board and CEO, to policy development, to customer support and moderation, company cultures, and company-specific understandings of culture can clash and conflict with the world more broadly.

Google’s particularly restrictive philosophy of identity clashed with their prospective users by either not understanding or not seeking to support the ways in which people construct and deploy their identities in everyday life. For many, identity

practices can be socially contextualised — with people acting and presenting themselves differently in professional, social, and familial environments. For some people, this involves multiple names and identities altogether.

Part of this failed relationship between Google+ and users could be attributable to commercial pressures or commercial instincts. As noted earlier, a significant proportion of Alphabet's (née Google) revenue is associated with Advertising sales, it may be the case that the focus on 'legal' or 'common' names was partly driven by a desire to enrich the data they can provide to prospective advertisers for the sake of targeted advertising.

Whilst the text of this case study has been relatively short, the cause of this is simply the relative unpopularity of this social platform. Like the Facebook case before it, Google shows a similar late strategy in sufficing — providing a solution to users that was not optimal, but instead largely acceptable to the remaining users.

The complete lack of change to the platform affordances to suit the change in policy is a clear example of this 'sufficing' strategy, as it shows a solution that is under-delivered, through either lack of time, lack of resources or lack of willingness to solve the problem.

Chapter 6: Case Study Three: Twitter, Abuse and Harassment

Content Note: In this chapter I quote and reproduce text that could be *charitably* described as hateful and abusive. These include violent threats, incitements to suicide, white supremacist rhetoric, sexist and racist slurs as well as host of other plainly offensive language. I reproduce this text in this work not to be gratuitous or offensive, but to give context to the discussion of the policies, cultures, and technological structures that support the publication of these texts. ‘Hate speech’ can be more easily dismissed as an abstract term, but in order to understand the dynamics of its effects on people and on discourse we must be ready to engage these texts and challenge the systems that support their propagation.

Disclosure: Some of the writing included in this chapter was originally published in a chapter that I contributed to the book *Negotiating Digital Citizenship: Control, Contest and Culture*, edited by Anthony McCosker, Son Vivienne and Amelia Johns, and published by Rowman & Littlefield in 2016.

Platform–Political Tensions

Whilst Twitter has become known in recent years for their problem dealing with belligerent and harassing users, it is worth noting that this issue has plagued Twitter for almost a decade. The cases analysed in this chapter have not occurred in a sociological or subcultural vacuum, in fact, they fit within a trend of abuse that has grown over time to now dominate the narrative of Twitter as a company. Whilst

Twitter was known in its earlier days for a particularly *laissez faire*, free speech absolutist attitude, their stewardship of a relatively open platform which has since seen dozens of high-profile cases of abuse and harassment, and countless lower profile cases, have seen this corporate outlook begin adapting to the practicalities of their governance roles.

In particular, this case analyses the bullying, abuse and harassment of users, and their expectations of safety, moderation and intervention from platform operators on Twitter. With this study, I provide a sketch of Twitter's history. Twitter's issues with harassing users and content first attracted press attention in 2008 and have continued and evolved to this day.

The issues of bullying, abuse and harassment in this case are not unique to Twitter. That said, factors in the design and governance of the platform likely contribute to the continued difficulties they face in addressing this issue. Twitter, as both company and platform is bears similarity to other websites and platforms with similar with similarly liberal attitudes to user governance.

Early Harassment on Twitter

The earliest case of harassment on Twitter to attain significant media attention related to technology consultant Ariel Waldman, who, after having months of difficulty with a stalker that pestered her with gendered insults and innuendo through a host of social networking platforms, reported her stalker to Twitter's moderators.

She found the response from Twitter lacking, and wrote on her blog about her frustration:

“The reality of Twitter is that they refuse to warn and/or ban people who use their service to ‘abuse, harass, threaten, impersonate or intimidate other Twitter users’ (to quote their fourth line item on their TOS [Terms of Service] page). What does this mean? In short, anyone can use Twitter to consistently harass you and ruin search results for your identity and Twitter won’t execute any means of community management” (Waldman, 2008a).

Waldman contrasts this response to that of the photo-sharing website Flickr, which responded to similar reports of harassing and threatening behaviour by deleting the reported content and banning the reported user. The initial response from Twitter indicated that “We can’t remove [this user’s] profile or ban [this user’s] IP address; [they’re] not doing anything illegal” (Ibid.). Given Twitter’s relative infancy, Waldman was able to talk to founder Jack Dorsey about their policy. Dorsey indicated to her that Twitter was hesitant to ban users from their platform, “because they’re scared if it turned into a lawsuit that they are too small of a company to handle it” (Waldman, 2008b). Waldman continued to report harassing content to Twitter and eventually received an email from Dorsey reading; “We’ve reviewed the matter and decided it’s not in our best interest to get involved. We’ve tasked our lawyers with a full review and update of our TOS. Thank you for your patience and understanding and good luck with resolving the problem” (Waldman, 2008a).

Unsurprisingly, Waldman found this response unsatisfying and continued to blog, and talk publicly about this problem, and advocate for Twitter to uphold their Terms of Service. In a thread on Twitter's (then active, now-defunct) customer service forums, she drew attention to her case, and drew attention to Section 230 of the Communications Decency act. Waldman (and her legal counsel) considered that "This pretty clearly shields them from liability for removing posts and users who tweet offensive or harassing messages." (Ibid.). Ev Williams, a founding member of Twitter's executive admonished Twitter users around this time, tweeting "Before joining a mob, you might want to check if everything they're saying/assuming is true" (E. Williams, 2008a), and "Our response to the Ariel situation" (E. Williams, 2008b) with a link to a response to Waldman's forum thread from Biz Stone, another Twitter executive.

On GetSatisfaction, Stone (referenced in Schiffman, 2008) wrote:

"Twitter recognizes that it is not skilled at judging content disputes between individuals. Determining the line between update and insult is not something that Twitter nor a crowd would do well.

We have deleted accounts for more straightforward violations of the existing Terms. That being said, we are engaged in an editorial review of those Terms to make it more clear what actions we will and won't take. Essentially, Twitter is a communication utility, not a mediator of content.

It's great that everyone's talking about this because it helps us make important decisions. After all, Twitter is a new medium that we're all figuring out as we go."

Stone's response indicated two key things. First, it indicated that Twitter's policymakers and moderators were uncomfortable in their role as arbiters of content: "Twitter recognizes that it is not skilled at judging content disputes between individuals. Determining the line between update and insult is not something that Twitter nor a crowd would do well". Second, it indicated that Twitter's reaction to the assertion that they were not upholding their Terms of Service as rules of their platform was to change their Terms of Service, rather than to more dutifully uphold their established terms.

This decision was received poorly by other, then-influential users of Twitter, who saw Twitter's decision to revise their Terms of Service as an abrogation of their responsibilities. Jeffrey Zeldman, a well-known web developer and prominent Twitter users noted that "Instead of dealing with the harassment, the CEO decided to alter Twitter's TOS. [...] One expects corporations to behave in cowardly and callously self-interested ways, but one expects more from one's heroes and friends" (Zeldman, 2008).

This response is an indication that Twitter, as a platform and as a company, have considered this problem — and determined that they do not want the responsibility of addressing it. It seemed, particularly in the platform's second year of operation, that they were uncomfortable with the prospect of being entrusted as arbiters of their user's speech.

This is in part a manifestation of the platform’s political ideologies — and its inclination to present itself as a “free speech” advocate — but it also suits the legal agenda of the company. By presenting themselves as a much simpler conduit-like technology, instead of a complex platform, Twitter attempts to limit the perception of their culpability in the production and distribution of harmful content. In the years since, this sort of distancing effort has become less effective for Twitter, as campaigns of harassment have developed to become larger and more systemically intricate. Similarly, the media attention attracted by these campaigns has made it difficult for Twitter to avoid a frank discussion of the roles it plays in adjudicating the platform, and the roles its users expect it to play.

Trending, and Affordance-Aware Harassment

Years later, we now see campaigns like ‘Austengate’ (the campaign against Caroline Criado-Perez in the wake of her advocacy of a Jane Austen Bank of England banknote) and ‘Gamergate’ (an amorphous and loosely organised campaign that began by targeting women in video game subculture, and later expanded its scope to target men, LGBTIQ+ people and people of colour who voiced critical opinions about video games in relation to social justice issues). In each of these campaigns, and the many others like them, the harassers and abusive users on social media platforms are becoming more expert, and their tactics of harassment are becoming more complex and affordance aware of the technological affordances of the social platforms.

For Twitter in particular, harassing users have made use of the functionality associated with hashtags and Twitter’s search API to intrude into discussions they were not already following.

This phenomenon of intruding repeatedly and uninvited into conversations under the pretence of politely asking questions was ascribed the name ‘sea-lioning’ after web-comic artist David Malki published a comic strip (Know Your Meme, 2014a) in which a woman confides in her friend that she can not stand sea lions. “Don’t say that out loud! Now you’ve done it!” he cautions, as a sea lion appears in the remaining five panels of the comic, pestering the woman at all-hours; “You made a statement in public for all to hear. Are you unable to defend the statements you make? Or simply unwilling to have a reasoned discussion?”.

To the targets of harassment in the Gamergate campaign, as well as observers of the campaign, the behaviour of the fictional sea-lion was reminiscent of the behaviour of Gamergate proponents (Doctorow, 2014). The behaviour is a form of disruptive conversation, that often proposes questions in bad faith. It bears similarity to a rhetorical technique often abbreviated to ‘JAQ-ing’ — Just Asking Questions; in doing this, it attempts to make wild accusations acceptable by framing them as questions rather than statements (RationalWiki, n.d.).

Another tactic associated with online harassment is ‘brigading’ (often short for vote-brigading), whereby the members of one online community organise *en masse* to influence another online community. Although it is against the platform’s rules

(Massanari, 2015), this happens quite frequently on Reddit, where users can post links to other subreddit forums as tacit calls-to-action to down-vote and lower the prominence of other users posts. Similarly, this behaviour can be used to organise off-platform brigades, where organising harassers provide links to users on other platforms that they find objectionable, as well as complaints or talking points. By leaving the call-to-action as implicit and avoiding any explicit encouragement, harassing users can skirt the rules of certain host platforms, to organise malfeasance on other platforms. Other examples of these include so-called ‘raid’ and ‘invasion’ chat rooms and forums. A well-known example is 8chan’s */baphomet/* board, which hosts and supports a number of crowd-enabled efforts towards releasing personal, private information (or doxing) and making false reports to law enforcement to elicit armed response to a target’s location (often called SWATing) (J. Allen, 2015b).

Online harassment has also been supported financially through crowd-funding and patronage platforms. In 2015, the patronage platform Patreon drew criticism for allowing the 8chan website, and Andrew ‘weev’ Auernheimer to use their infrastructure for raising money, because Auernheimer’s activities and the 8chan platform have been linked to harassment on many other platforms (J. Allen, 2015a; Fleishman, 2014).

As noted with the discussion of ‘platform politics’ in the literature review, the term ‘platform’ is a particularly flexible descriptor for social networks and online software, as it represents a number of implicit political notions. Key amongst these is a presumption of technology as ‘neutral’.

The ‘neutral platform’ ideology that argues that web-based platforms should be treated like ‘common carrier’ telephony providers, or simple ‘conduits’ for their users, comes to a head when issues of legality, fairness and freedom for user’s expression, and the rights and responsibilities of platform operators clash. When faced with unrest from users over Auernheimer’s usage of Patreon, Patreon’s founder, Jack Conte reflected that “This is so hard - as a society, we must let the fringe have a voice - it's so important, even if we disagree [...] Even murderers get a right to a fair trial, right?” (Fleishman, 2014). Conte, a well-known musician whose fame as part of the *Pomplamoose* band grew on Youtube, seemed to resent the monolithic enforcement of copyright restrictions on Youtube, and expressed a reticence to make an executive decision to judge a user’s off-platform behaviour. Later in the year, an updated set of Community guidelines eventually saw the removal of both Auernheimer and 8chan from Patreon.

Platforms and Abuse

Reddit, Youtube, and smaller platforms

In tandem with Twitter’s issues regarding harassment, Reddit’s users and operators have faced controversy across a number of issues relating to harassment, hate speech and sexual photos of people either taken or shared without their consent¹⁰.

¹⁰ These images are often characterised as ‘revenge porn’ but associating this ‘revenge’ behaviour with the legal and consensual practice of creating adult entertainment seems counter-intuitive and unhelpful.

I highlighted earlier in this thesis the 2012 case of ‘violentacrez’, a Reddit user whose avid use of the platform earned him moderator roles in a number of forums. It is worth returning to his case here, to examine abuse in the context of the Reddit platform. In 2012, Brutsch was ‘doxxed’ — in this case, his name ‘Michael Brutsch’ was revealed — by then Gawker journalist Adrian Chen. Chen found the subreddit forums that Brutsch moderated or established reprehensible (Chen, 2012b). These forums included r/Chokeabitch, r/Niggerjailbait, r/Rapebait, r/Hitler, r/Jewmerica, r/Misogyny, r/Incest. Brutsch left Reddit shortly after the exposé, although the incident started a broader discussion about freedom of speech on the Reddit platform (Baker, 2013). Speech controversies continued on Reddit, with the release of nude celebrity self-photographs reportedly hacked from iCloud accounts, and with the growing prominence of a number of explicitly or belligerent subreddits, including r/CoonTown, r/GasTheKikes and r/FatPeopleHate (Sottek, 2015). Both Reddit’s reticence to police these communities, and the community’s visceral, vituperative response to the spectre of regulation speaks to a sort of free-speech absolutism that underpins Reddit’s ethos. In spite of this, system itself is not entirely a grand democratic enterprise, as the moderators of forums are given the abilities to prevent users from participating through blocks, through automated censorship, and through the use of non-obvious policing methods, like shadow-bans (which give the user the impression that they are still participating, even though their contributions are not seen by any other users (Fasnacht, 2015).

Facebook and Abuse

Facebook on the other hand, has dramatically different affordances which in turn affect the ways in which harassing behaviour can occur on the platform. For the most part, harassment on Facebook tends to be restricted to semi-private delivery systems (Through the easily filterable Messages system), and to the comments of publicly visible content.

The extent to which harassers can impersonate their targets is also partially ameliorated by Facebook's so-called 'Real Names' policy. This is a polar opposite of Twitter's policy, where although impersonation accounts have been frowned-upon in both the Rules and Terms of Service, in practice they tend to be treated as 'parody' accounts and left alone.

Although the disinhibition effect associated anonymity is often a convenient explanation for online misbehaviour, it's worth noting that many of the Facebook users who do participate in harassing and unlawful behaviours on the platform do so under their 'real name'. This may make it easier for legal recourse, or for a target of harassment to learn more about their harasser as Ford (2015) and Pearce (True, 2014) demonstrate.

Platform History and the Conflict

Policy Beginnings

Twitter's earliest policy towards abuse, bullying, intimidation and harassment appears as a single line in Version 1 of the Twitter Terms of Service. This version,

which served from around the time of Twitter's formation in 2006, to the terms were revised in 2009;

You must not abuse, harass, threaten, impersonate or intimidate other Twitter users.

Other parts of this first Terms of Service document provide supporting, and corollary rules, such as; “You may not use the Twitter.com service for any illegal or unauthorized purpose. International users agree to comply with all local laws regarding online conduct and acceptable content”, and “You must not, in the use of Twitter, violate any laws in your jurisdiction (including but not limited to copyright laws).” The punishment indicated for breaking these terms, was simply “the termination of your Twitter.com account.” (Twitter, 2006)

Whilst it's difficult to identify the earliest cases of bullying or harassment, the earliest incident large enough to warrant a public response from Twitter's operators seems to be the aforementioned case of Ariel Waldman, who complained to Twitter in 2008 about a stalker who continued to send her harassing messages across a number of social media platforms (Waldman, 2008a). As noted earlier, Waldman wrote about her experiences on her blog, and on Twitter's Get Satisfaction¹¹ forums, arguing that although the account of the user stalking her was clearly breaking Twitter's Terms of Service, Twitter seemed too timid to act punitively against her harasser (Ibid.). Twitter's responded to Waldman's criticisms by expressing disappointment at the

¹¹ Get Satisfaction is a third-party company that provides troubleshooting and customer support infrastructure to other companies. Twitter used Get Satisfaction and competitor Zendesk during its earlier years

situation and articulating their role as “communications utility, not a mediator of content” (Stone in Schiffman, 2008). As Waldman (2008a) notes, this attitude towards harassment on the Twitter platform is peculiar. Whilst Twitter openly discussed their mediation and moderation of other content, especially illegal or copyright infringing content, they also took a hands-off approach to harassment on the platform that effectively rendered their discussion of harassment in their Terms of Service as an empty platitude.

Key Conflicts

As indicated earlier, harm, abuse and threats the safety of others are not a particularly new phenomenon, and have a history dating back to well before the advent of personal computing.

To that end, this case study focuses on two key conflicts that are particularly revealing in providing opportunities to unearth the dynamics of expression and safety in digital platforms, and how these concepts are navigated by both user and platform operator alike. The ‘Austengate’ and ‘Gamergate’ conflicts present cases where groups of belligerent users campaigned to bully, harass and intimidate other users on Twitter. In these cases, Twitter’s strategic power — wielded in the architecture and policies of the space can reveal their power on the platform. Similarly, responses of users to harassment and to Twitter’s governance of the platform reveals the expectations and powers of citizens in these digital spaces.

#AustenGate

In 2013 the departing Bank of England Governor Sir Mervyn King (Bank of England, 2013) announced that from 2016 onwards, the Bank of England would introduce polymer notes into circulation. Coinciding with this change was the announcement that Sir Winston Churchill's portrait would replace Elizabeth Fry's portrait on the reverse side of the £5 note (K. Allen, 2013). This change drew the ire of some observers, who noted that with Fry's replacement, there would no longer be any women represented on the reverse side of any Bank of England Currency (Criado-Perez, 2013b), which has been used to celebrate the achievements and contributions of historic Britons since the introduction of Series D notes in the 1970 (Bank of England, 2015)¹². After complaining to the Bank of England about the resulting lack of female representation and finding their response "inadequate" (Hinge, 2013) Caroline Criado-Perez launched a legal challenge under the *Equality Act, 2010 §149(1)* against the bank for failing to have "due regard" to the need to "eliminate discrimination" and "advance equality of opportunity" in its operation as a Public Authority (Ibid.). In conjunction with this, Criado-Perez also started a petition and public campaign to suggest Jane Austen for consideration as a portrait on the reverse side of a banknote (Criado-Perez, 2013a). The campaign was successful, with over 36,000 signatures on a Change.org, and supporting letters from members of parliament. The Bank of England announced its intent to feature Austen on the reverse of the £10 banknote on 24th of July 2013. Criado-Perez hailed this as a victory for

¹² Queen Elizabeth II remains on the obverse side of all Bank of England banknotes; however, this is a function of her role as Head of State, rather than a recognition of her accomplishments or contributions to British Society.

women — but the celebrations were quickly overshadowed by a backlash, counter-campaign against Criado-Perez herself.

A litany of grotesque and intimidating threats ensued on Twitter, primarily delivered through public tweets. Many employed the use of sexist slurs and violent, threatening language;

“this Perez one just needs a good smashing up the arse and she’ll be fine” (in Criado-Perez, 2013c)

“Everyone jump on the rape train > @CCriadoPerez is conductor” (Ibid. c)

“Wouldn’t mind tying this bitch to my stove. Hey sweetheart, give me a shout when you’re ready to be put in your place” (Ibid. c)

Criado-Perez wrote that at the height of the campaign, she would receive one or more intimidating tweets every minute (Ibid. c).

The campaign against Caroline Criado-Perez drew significant media attention as journalists, columnists and politicians who reported on or spoke out against the campaign of threats and harassment were often also targeted. Writers India Knight and Laurie Penny wrote of a specific bomb threats made against them that was treated as credible by Metropolitan Police (Penny, 2013b; Knight in B. Quinn, 2013). Similarly, Labour MP Stella Creasy, who was involved in Criado-Perez’s bank note campaign, received dozens of threats from Twitter users.

“I’m going to pistol whip you over and over until you lose consciousness, then burn ur flesh.” (in Roiphe, 2013)

“I will rape you tomorrow at 9pm, Shall we meet near your house?” (Ibid.)

“You better watch your back, I’m going to rape your arse at 8pm and put the video all over.” (Jozuka, 2014)

Over 147 Twitter accounts (Creasy, 2014) were associated with this campaign of harassment and intimidation against Creasy, Criado-Perez and other prominent women in the United Kingdom. It’s not possible for external observers to determine how many users were behind these accounts as it is relatively trivial for users to repeatedly create new accounts on Twitter, but the police investigations into the threats resulted in the arrests and convictions of three Twitter users under §127 of the *Communications Act* (2003) for sending messages that were “grossly offensive or of an indecent, obscene, or menacing character”: Peter Nuun (Creasy, 2014), Isabella Sorley and John Nimmo (Riddle, 2014). For Sorley (P. Smith, 2015) and Nimmo, the bilious counter-campaign was more about the sense of celebrity and community they earned from participation in an online mob: “I saw it was trending, so I looked into what it was about and, stupid me, I decided to join in. And I was getting, like, retweets, I was getting favourites and all that” (Nimmo in P. Smith, 2015). Nimmo (2014) wrote in his blog “The irony of it all is that I wasn’t even passionate about the subject or the people I was bullying. I was simply bored, saw what was trending and leaped on to the bandwagon.”

Only a year later, this style of misogynistic and explicitly anti-feminist would appear again in the shape of ‘*Gamergate*’, an amorphous crusade against women,

people of colour, and LGBTIQ+ people who were seen as interlopers in video gaming culture (Mantilla, 2015 137).

Gamergate

The *Gamergate* conflict began early in August 2014 after Eron Gjoni, the ex-boyfriend of independent game developer Zoë Quinn, published ‘the zoe post’, a “rambling online essay” (Wingfield, 2014) designed to cause “maximum pain and harm” (Gjoni in Jason, 2015) in which he accused her of having a relationship with a journalist associated with Gawker Media’s popular gaming website *Kotaku*. Although Gjoni’s forum posts promoting this blog were deleted from the Penny Arcade and Something Awful forums (Gjoni, 2014), Gjoni’s post was still shared on several websites, including 4chan’s /r9k/ ‘Robot 9001’ board (GjoniAnonymous /r9k/ Users, 2014) and the gaming and culture website *The Escapist*’s forums (Rebant, 2014).

This prompted a loosely organised campaign of harassment (Thomsen, 2014), that continued unabated for the remainder of the month, focusing largely on the harassment of Quinn and the allegations made by Gjoni — with some gaming commentators alleging that Quinn’s relationship with *Kotaku* journalist Nathan Grayson was a *quid pro quo* for favourable press coverage and reviews. This assertion was quickly revealed to be demonstrably false, as the *Kotaku* website has never reviewed Quinn’s ‘Depression Quest’ game, and as Grayson had only written a single article relating to Quinn’s involvement in game development before their relationship began, and not at all since (Totilo, 2014). In spite of a clear lack of evidence of any

malfesance on Quinn or Grayson’s part, these allegations continued, and the scope of harassment extended not just to Quinn, but also supporters like Phil Fish (Grubb, 2014) and other women who identified as or were perceived as either a ‘feminist’ or a ‘social justice warrior’ (often presented as the acronym *SJW*)¹³, like cultural critic Anita Sarkeesian (Johnston, 2014). The campaigns clear and overwhelming focus on harassing Quinn and other women in the gaming industry earned it an early admonition as misogynistic and hateful, though its proponents often insisted that their campaign was “Actually about ethics in gaming journalism”, a refrain that would eventually become the punchline of an absurdist series of humorous images (Know Your Meme, 2014b).

By the end of August, the controversy earned a name; ‘#Gamergate’, ascribed in a tweet from television actor and conservative activist Adam Baldwin (Dockterman, 2014). A host of video gaming websites, including the developer-focused website *Gamasutra*, had begun to criticise the campaigns against Quinn, with a number of writers railing against the notion of ‘gamer’ as a consumeristic subcultural identity (Alexander, 2014).

The identity of ‘gamer’ has been particularly fraught in this controversy, as the proponents of Gamergate often discuss their identity as ‘gamers’ in a manner that

¹³ Social Justice Warrior is often used as a pejorative epithet to dismiss the arguments of feminist, anti-racist and social justice-focused campaigners as radical or ludicrous.

The term is also adopted by many of those it is intended to disparage, who argue that being described as a fighter for social justice is not an insult — and occasionally remix the insult with other fantasy-style character classes common to role-playing games — e.g. ‘Social Justice Mage, Paladin or Cleric’.

mimics and appropriates minority politics and language to construct ‘gamer’ as what Kunyosying and Soles (2012) might term as a “simulated ethnicity”, with reference to the melodramatic construction of ‘gamers’ as members of a marginalised identity so that they can claim a validity or authenticity for their concerns. Similarly, the ‘gamer’ identity is constructed as a supra-identity that subsumes other identities. As Cross (2015) argues, women and minorities are welcomed to the identity in a limited capacity, contingent on “checking all other identities in at the door”.

Gamasutra writer Leigh Alexander (2014) and others opposing the ‘gamer’ identity argued that producers should attempt to meet the needs of a broader, and growing audience of people who play games, rather than just the most dedicated, ‘hard-core’ audience members that wear ‘gamer’ as a part of their identity. These sentiments, incidentally, mirror the conclusion of theorist Ian Bogost’s (2011 154) book ‘How to Do Things with Videogames’;

“As videogames broaden in appeal, being a ‘gamer’ will actually become less common, if being a gamer means consuming games as one’s primary media diet or identifying with videogames as a primary part of one’s identity. [...] Soon gamers will be the anomaly. If we’re very fortunate, they’ll disappear altogether. Instead we’ll just find people, ordinary people of all sorts. And sometimes those people will play videogames. And it won’t be a big deal, at all.”

Contrasting the thrust of the Gamergate’s efforts with this sentiment, it is not difficult to see the Gamergate campaign as a conservative, reactionary revolt against

the perception that newcomers and progressive interlopers have come to drastically change their way of life (Cross, 2014).

Gamergate was quickly excommunicated from 4chan, one of its original hubs of activity, after the website's owner Christopher 'moot' Poole decided that its proponents were too ready to violate 4chan's 'rules' regarding the publication of private, personal information (Kushner, 2015). The movement consolidated its discussion and organisation to Reddit's '/r/KotakuInAction/' forum, and 8chan, a 4chan-style imageboard (Machkovech, 2015a) that is notable mostly for its association with Gamergate (Cush, 2014) and its removal from Google search results for reportedly hosting 'suspected child abuse content' (Machkovech, 2015c).

The Gamergate conflict has been characterised as a new, technologically-advanced style of 'culture war' (Chituc, 2015) (Ernst, 2014) (Shackford, 2014) (Kyle Wagner, 2014) — a protracted, ideological conflict between advocates for change, and proponents of the status quo. In light of this, it is not surprising to see that influential figures in conservative (Chituc, 2015), reactionary and 'neoreactionary' (J. Allen, 2014) circles —including Breitbart columnist Milo Yiannopoulos (Schubert, 2014), American Enterprise Institute scholar Christina Hoff Sommers, and notorious hacker, self-described "white nationalist" (Auernheimer, 2014) and troll (Penenberg, 2015) Andrew 'weev' Auernheimer (Biddle, 2014b)— are regularly afforded celebrity status amongst Gamergate proponents on 8chan's boards, and the /r/KotakuInAction forum (OMalley, 2014).

There's much more that can (and should) be written about Gamergate — from the cross-platform loose organisation of campaigns, to the semiotics of 'Vivian James' the movement's coded rape joke-cum-mascot¹⁴ (Melendez, 2014). For now, the core allegations of 'unethical journalism' alleged by 'Gamergate' proponents have largely been exposed as mess of misinformation (Lomas, 2014), wilful misunderstanding (Petit, 2015) and '*chart brut*'-style¹⁵ conspiracy theories — thin veils used to obscure troublingly sexist perspectives and behaviours (Wofford, 2014). Although 'Gamergate' wanes in public significance¹⁶ a dedicated base of harassers and culture-warriors remain tightly associated with the movement and its media.

Responses to Conflict

Twitter's response to these cases, and to broader issues of abuse and harassment has been relatively meek, and has drawn criticism from users and external observers alike (Tiku & Newton, 2015). Whilst Twitter spokespeople have acknowledged the

¹⁴ The colours of the 'Vivian James' mascot often pictured in Gamergate iconography are a reference to an animated GIF file and early 4chan meme called "Piccolo Dick" that had regularly circulated on 4chan from 2008 onwards. The image was shared so regularly that it was often termed a "Daily Dose", to which users would respond "Thanks Doc."

In "Piccolo Dick" the green-skinned, purple-clothed, Piccolo character from popular anime *Dragon Ball Z*, rapes another character, Vegeta. After 4chan moderators started to ban users who shared or uploaded the image, users rebelled and began posting abstract green and purple images in reference. The creators associated with the Vivian James character publicly contest this interpretation, however a significant amount of the early responses to the character's design included the reference "Thanks Doc" in reply.

¹⁵ Biddle (2014a) and Rugnetta (2015) term the aesthetic of crude "Microsoft Paint"-annotated images associated with online conspiracy theories and mob activity as 'Chart Brut', citing examples of the ill-conceived /r/FindBostonBombers subreddit, the 'Fappening' leak of private nude celebrity photographs, and Gamergate.

¹⁶ For example, a Google Trends search for 'Gamergate' reveals the term's steady decline in search traffic from late 2014 onwards. (For example: <https://trends.google.com/trends/explore?date=2013-01-01%202018-12-31&q=gamergate>)

existence of this issue, the policy or technological responses have not affected meaningful change on the platform (Gadde, 2015). The technological responses, including a ‘Report Abuse’ button, and the ability to export and import lists of blocked accounts (Romano, 2013b) have been criticised as susceptible to abuse, and inflexible to the needs of users. Whilst it’s difficult to discern if any internal changes at the company have influenced the ways in which Twitter deals with harassing and abusive users, observers like Brianna Wu—who has been repeatedly targeted in Gamergate related abuse and harassment—have noted differences in attitudes towards moderation at different stages in the Twitter’s operation.

Governance Context

Platforms — Strategies

Policies

As an organisation, Twitter has long been a fierce advocate of free speech for its users — with former Chief Executive Dick Costolo (in Barnett, 2011) asserting that Twitter strives to be the “free speech wing of the free speech party”. This is represented both in the rules of Twitter, which are broadly permissive of user expression —even to the point of allowing nudity and pornography on the platform—, and their actions in contesting United States Department of Justice subpoenas for user data relating to WikiLeaks affiliated users (Barnett, 2011). Similarly, their reticence to censor anti-semitic tweets in response to requests from German (Spiegel Online, 2012) and French (Beardsley, 2013) governments resulted in the development of a location-based censorship mechanism, which created affordances for certain tweets to be made unavailable for viewing in the countries in which the governments had sanctioned them.

Whilst permissiveness and freedom of expression are still important values held by the platform’s operators and communicated in the company’s blog and public statements, recently the subject of user safety become a more readily discussed priority for Twitter, in both internal and public communications. Dick Costolo’s leaked communique (Tiku & Newton, 2015) provides a clear example of this attitude:

“We suck at dealing with abuse and trolls on the platform and we’ve sucked at it for years. It’s no secret and the rest of the world talks about it every day. We lose core user after core user by not addressing simple trolling issues that they face every day. I’m frankly ashamed of how poorly we’ve dealt with this issue during my tenure as CEO. It’s absurd. [...] We’re going to start kicking these people off right and left and making sure that when they issue their ridiculous attacks, nobody hears them.”

Twitter General Counsel Vijaya Gadde (2015) argues that the recent focus on user safety is core to the company’s platforms value of free expression, noting that they are allowing that value to be undermined if they allow belligerent users to threaten, harass and intimidate other users into silence. With this, Twitter begins to embrace the nuance that Owen Fiss (1996) explored in *The Irony of Free Speech* — that censorship, can modestly enhance freedom, to the extent that governmental actors can use censorship to ensure that the “underfunded, underrepresented, or disadvantaged voice” (Bernstein, 2014 2) can be heard.

Design

Twitter’s policies speak clearly and definitively about the importance of user safety but are undermined by organisational issues and design issues. At an organisational level, Twitter may not prioritise user safety to the extent it does other commercial, technological or ideological concerns. An example of these priorities can be seen in Twitter’s short-lived attempt to change the ‘blocking’ function in 2013 (Sippey, 2013). At the time, ‘blocking’ another user on Twitter would remove their

ability to view your user profile or tweets whilst logged in and would prevent them from using your username to @reply and @mention you. Twitter briefly changed the function of ‘blocking’ to be more akin to a ‘mute’ function, which would allow blocked users to follow and interact with the users who blocked them — but would simply not send notifications the users who blocked them (Davidson, 2013).

Whilst both ‘block’ and ‘mute’ functions are useful to users in different ways, the outright replacement of the more powerful ‘block’ with a less effective ‘mute’ was of great concern to users, many of whom were strongly critical of the change. Twitter changed the function back within twelve hours of its launch (Sippey, 2013). It was later revealed that Del Harvey, Twitter’s Vice President of Trust and Safety had internally counselled against the change, warning that it was a “terrible idea” that “would make cyberbullying easier” (Hill, 2014).

As a technological structure, Twitter’s platform has affordances that can support uses that are abusive to users. This can be seen in the ease with which serial harassers can set up multiple accounts (Twitter, nd). By having multiple accounts, at the ready, users inclined to harassment can simply log into new accounts and continue their campaigns after being banned or punished for breaking Twitter’s rules, or use multiple accounts at the same time to create the illusion of a more populated campaign (Seife, 2014). This is a tactic of abuse that has been deployed in the aforementioned Austen and Gamergate campaigns (Carter, 2014).

Even Twitter's 'Promoted Tweets' advertising system is susceptible to abusive behaviour. The system adds tweets to user's Twitter timelines based on the demographic options, email addresses or Twitter usernames provided by the advertiser, as well as providing flexible affordances not available to ordinary tweets. This can allow for both broad or small-scale targeting for the production of tweet campaigns, for the purposes of harassment and abuse.

One such campaign, offers a clear example of how Twitter's 'Promoted Tweets' system can be used for harassment:

@Caitlin_Roperr: Trannies, getting surgery won't change your gender. It will only turn you into a deformed freak. End your miserable existence. Kill yourself (in Gibbs, 2015)

@Caitlin_Roperr: Trannies, everyone is laughing at you. You will never be able to change your chromosomes. End your miserable existence & kill yourself now. (in Payton, 2015)

@Caitlin_Roperr: Trannies, you will never be the opposite gender. (in Gibbs, 2015)

@Caitlin_Roperr: Trannies, your families will never love you. You are living a lie & you know it.

End your miserable existence.

Commit suicide now. (in Gibbs, 2015)

The preceding four tweets were targeted at users of Facebook who were identified as transgender¹⁷ all originated from ‘@Caitlin_Roperr’, an account set up to impersonate Caitlin Roper, an Australian feminist campaigner. On 4chan’s */pol/* board, an Anonymous user claimed credit for the campaign (“*/pol/* - Politically Incorrect,” 2015; Hathaway, 2015). It is alleged that this anonymous 4chan user was Joshua Goldberg (Plenke, 2015; Potaka & McMahon, 2015), a young man who reportedly used similar tactics in a campaign to impersonate Australian lawyer Joshua Bornstein’s name and photographs in an elaborate ruse to contribute a blog post to the Times of Israel’s website calling for the genocide of the Palestinian people (Times of Israel Staff, 2015) (Bornstein, 2015), as well as impersonating him in promoted tweets (Knibbs, 2015) that called for similar acts of violence and genocide.

At the time of writing, Goldberg has been arrested on suspicion of encouraging an attack in Kansas City, Missouri on the 14th anniversary of September 11 (Zavadski, 2015) and has been charged with “distribution of information relating to explosives, destructive devices, and weapons of mass destruction” (Moyer, 2015). He had been sent for a psychiatric assessment to determine his competence to stand trial (Hussein, 2015). In 2017, he pleaded guilty to federal charges of attempted malicious damage and destruction by an explosive of a building. On June 25, 2018, Goldberg was sentenced to ten years in federal prison and lifetime supervision.

¹⁷ The anonymous 4chan user who claimed credit for this campaign did not reveal whether he supplied a list of user accounts for ad-targeting that they suspected to belong to transgender users, or if Twitter’s advertising system directed it to users with interests in transgender issues — as the Twitter platform does not offer users the ability to indicate their gender in their user profile and settings pages.

The progenitor of these efforts to use Twitter’s Promoted Tweets for targeted harassment seems to be Andrew ‘weev’ Auernheimer (2015a; 2015b), who used Promoted Tweets to target tweets about ‘white discrimination’ and ‘white pride’ to women and minorities (Brogan, 2015).

@rabite: Whites need to stand up for one another and defend ourselves from violence and discrimination. Our race is dying.

@rabite: White pride, world wide. Do you know the 14 words?¹⁸

As both Auernheimer (2015b) and Baio (Baio, 2015) note, the affordances available to advertiser’s tweets are significantly more flexible than those available to non-paying users, and include the ability to make a ‘nullcasted’ tweet — that is, a tweet that does not appear in the creator’s timeline, nor in the results for Twitter search¹⁹. Nullcasted tweets are difficult for the average Twitter user to find, unless they are @mentioned in the tweet itself, or they follow the chain of replies from a tweet the timeline of a user who was @mentioned in the nullcasted tweet.

To be clear, Promoted Tweets is neither a tool created for abuse, nor a tool solely used for abuse, and much of its usage is likely for fairly traditional advertising

¹⁸ The ‘fourteen words’ is a well-known reference to the phrase “We must secure the existence of our people and a future for White Children.”, popularised by white supremacist David Lane. (Anti-Defamation League, nd). This phrase is reportedly derived from a passage in Adolf Hitler’s autobiographical book *Mein Kampf*.

¹⁹ Although, at the time of writing, some third-party applications with access to the full gamut of Twitter data –including promoted tweets– can find these ‘nullcasted’ tweets.

purposes. The design and implementation of this system however has provided avenues for abuse without appropriate safeguards.

Users - Tactics

Users who have borne the brunt of the hostilities in campaigns like that against Criado-Perez and Gamergate deploy a range of tactics in navigating Twitter and negotiating the affordances of the platform. We can classify four key tactics as follows:

- **Advocacy** - The use of rhetoric and argument to advocate for changes in policy and/or enforcement on a platform.
- **Circumvention** - The use of third-party software and/or the platform's application programming interface (API) to alter the harassed user's experience of the platform.
- **Intervention** - The use of social intervention or social engineering to appeal directly to a harassing user's friends or family.
- **Litigation** - The use of legal threat or legal argument against either a social platform or a user of the social platform.

These types of tactics are seen in many other types of platform-user conflict but are particularly common to conflicts relating to harassment and user safety.

Advocacy

In the aftermath of the campaign of tweets and threats following Caroline Criado-Perez's banknote campaign, users turned their attention to the platform, with over 140,000 people contributing signatures to a petition demanding better reporting tools (Graham, 2013). Whilst Twitter was not criticised for hosting or facilitating the campaign, users saw both a failure of moderation and a failure of affordances (Elgot, 2013), and lobbied accordingly for Twitter to make changes to address these flaws. Similarly, the lengthy Gamergate campaign saw users and observers not only campaign for better protections against harassment and harassing campaigns (Eleventh, 2014) on Twitter and other social platforms, but also for the response and intervention of traditional law enforcement agencies (Clark, 2015; Machkovech, 2015b). A key part of these campaigns is attracting the attention of 'main-stream' media sources and the sense of public legitimacy that widely published coverage can confer. This is often accomplished by writing publicly about their experiences and being interviewed by in mainstream news publications. The attention of policymakers and major news outlets often necessitates responses from the operators of social platforms. The *meae culpa* from Twitter's United Kingdom General Manager Tony Wang (Politi, 2013) and then Chief Executive Officers Dick Costolo (Tiku & Newton, 2015) are key examples of these responses — both coming after significant public consternation about Twitter's failures to protect users.

Circumvention

Twitter's *relatively* extensible platform infrastructure allows users the opportunity to navigate the platform in atypical ways. A notable example of this is the rise of third-party blocking applications that use Twitter's API to out-source or

collaborate on blocking accounts that are known to or deemed likely to harass users. Some examples of these third-party systems include *The Block Bot*, *Blocktogether*, and *Good Game Auto Blocker*. Though each of these systems have different implementations, and different designs, their goal is fundamentally quite similar.

The Block Bot is an application that includes an updating set of block-lists. It is largely developed and organised by members of the ‘Atheist+’ community, a movement that attempts to accommodate diversity of gender, race, sexuality and class in Atheist discourse. It is operated by a team of moderators who classify Twitter accounts under three tiers based on levels of harm and aggression. The highest ‘Level 1’ tier criteria include a bevy of behaviours that are already against Twitter’s rules (Twitter, nd), including “threats, harassment [...] impersonating someone, [...] encouraging self-harm, [...] promoting hate speech” (The Block Bot, 2013). The lowest tier ‘Level 3’, expands the list to include accounts that the moderators deem likely to be “tedious or obnoxious”, especially from accounts that “appear to frequently engage in micro-aggressions [... or] show a sense of entitlement to have a conversation”. Users who activate *The Block Bot* with their Twitter accounts can choose their preferred level of severity to pre-emptively block according to their desires.

In comparison, *Block Together* provides a more flexible system for users “to reduce the burden of blocking when many accounts are attacking you, or when a few accounts are attacking many people in your community” (Hoffman-Andrews, n.d.). *Block Together* provides the option for users to privately manage their own block lists or share them amongst others. It also provides a number of blocking affordances

designed to counteract behaviours commonly associated with harassing accounts. In particular, *Block Together* provides options for users to block “young” Twitter accounts that are less than a week old, as well as Twitter accounts with less than fifteen followers. Both of these options provide some remedy to the harassing tactic of ‘sockpuppeting’²⁰ — creating new, disposable Twitter accounts to avoid the punishments associated with breaking Twitter’s rules.

The *Good Game Auto Blocker* (often shortened to GGAutoBlocker) uses the underlying *Block Together* system, along with an associative algorithm that adds Twitter users to a shared block list on the basis of who they’re following. If a user follows more than one of a handful of twitter accounts associated with the Gamergate movement, including Breitbart columnist Milo Yiannopoulos and blogger Ethan Ralph (Harper, 2014a), they’re added to the list. This dragnet method is efficient at weeding out potentially harassing users but flawed as ‘following’ a Twitter user does not particularly indicate support for them or their views. To counteract this, *Good Game Auto Blocker* has an appeals process by which Twitter users can publicly appeal their ban and request to be added to a whitelist of accounts that the algorithm will not block (Harper, 2014b).

²⁰ Sockpuppet is a relatively old term that became common in USENET groups in the mid to late 1990’s. It refers to an online identity used with the intention of misleading others. Sockpuppet identities are often used to praise, defend or criticise another member of an online platform under the guise of offering an independent or correlating third-party viewpoint.

The term refers to a rudimentary style of puppet made from a sock and controlled by the wearer’s hand.

At a less-comprehensive level, third-party Twitter clients like Tweetbot and Twiterrific allow flexible ‘mute’ and ‘muffle’²¹ options to create personalised filters that hide or make less visible tweets from certain users, clients, including certain hashtags, or tweets that conform to user-specified regular expressions (Gonzales, 2014). Although Twitter now offers rudimentary block list sharing (Zhang, 2015) and both ‘mute’ and ‘block’ functions (Twitter, 2015), these third party predate Twitter’s provision of these services and functions, and continue to offer more flexible options to users who find Twitter’s affordances lacking.

Intervention

For some users, an effective tactic is to directly or indirectly engage their harassers through social means. A particularly well-known instance of this is of writer Alannah Pearce’s response to a harasser who turned out to be a young boy. Pearce used Facebook to investigate the harassing user (who had seemingly used his real name) and used Facebook’s messaging function to contact the boy’s mother; providing a screenshot of the threat and asking her if she’d like to discuss his threat with him. The mother replied “omg little shit”, “IM SO SORRY”, “YES ILL TALK TO HIM” (True, 2014), and reportedly asked her son’s school to talk about online harassment and bullying (Pantozzi, 2014). This sort of social intervention can be a powerfully discouraging factor. Zoë Quinn (Z. Quinn, 2015), one of Gamergate’s key targets,

²¹ The distinction between ‘mute’ and ‘muffle’ is relates to the design of the functions. Whilst ‘mute’ features do not display tweets that conform to mute filters, ‘muffled’ tweets still appear in a user’s timeline but are given a visual and design treatment that shows that a tweet has been hidden in the timeline and indicates the filter responsible for the removal. The user can then tap a ‘muffled’ tweet to see the hidden tweet (The Iconfactory, 2013).

reflecting on her own online behaviour as a teenager and her discussions with self-identified ‘former trolls’ argues that “more often than not [trolls] expressed that someone they were closed to, respected, or looked up to said that wasn’t cool — that the social network supporting this kind of feeding frenzy was no longer reinforced.”

Litigation

Although targets of harassment have had difficulty in getting platform operators to respond to harassing behaviour and content, social media platforms regularly provide an example in which their moderations are particularly prompt, effective — intellectual property. Given that platforms like Facebook, Twitter and Google police their social platforms for potentially and presumptively infringing copyrighted content, some targets of harassment have sought to make use of these efficient systems for combatting harassment.

In 2012 an inflammatory anti-Islamic film titled ‘Innocence of Muslims’ was uploaded to Youtube. The film drew condemnation almost immediately. Cindy Garcia, an actress who was tricked into appearing in the film and dubbed over in post-production sought to have Youtube remove the film on copyright grounds, as she was receiving harassment and threats of harm for her role in the film (Jeong, 2015). Youtube declined to remove the film at her request, as she was not the producer or copyright holder. Garcia attempted again to have the film removed by suing Google with a copyright claim and requesting the content’s removal with the *Digital*

Millennium Copyright Act (DMCA). Although her case was briefly successful, it has since been overturned on appeal.

Jeong (2015) notes that “usually, when seeking to delete something off the Web, wronged individuals go after the platform that hosts the content. “Whilst internet social platforms are generally not held liable for the conduct of their users given Section 230 of the Communications Decency Act (CDA), “the biggest gaping hole in CDA 230” is provided by the DMCA — where most of the legally-mandated deletion and moderation occurs.

The DMCA is particularly effective given the relative absence of policing effort from traditional law enforcement agencies. Whilst methods of online harassment are illegal under federal and/or state laws in many countries²², the enforcement of these laws is patchy at best. This problem is twofold. First, it can be difficult to target a harasser directly, especially if that harasser is anonymous, or using multiple accounts.

²² Under *Title 18 of the United States Code*; § 875 outlaws any interstate or foreign communications that threaten injury. § 2261A outlaws any interstate or foreign electronic communications with the intent to kill, injure, harass or intimidate another person – especially conduct that creates reasonable fear of death or serious bodily injury, or that attempts to cause substantial emotional distress to another person.

The United Kingdom’s *Communications Act 2003*: § 127 makes it an offence to send an electronic message that is grossly offensive, indecent, obscene or of menacing character.

Similarly, under Australian state and commonwealth law: § 474.15, .16 and .17 of Australia’s *Criminal Code Act 1995 (Commonwealth)* makes it an offence to use electronic communications to threaten to kill or harm, to send hoaxes about explosives or dangerous substances, or to menace, harass and cause offence.

They may be outside of the jurisdiction of local law enforcement, or they may regularly create and deleted their own harassing content in order to attempt to ‘gas light’²³ their targets. Second, there are likely issues with the currently afforded training for police departments for these kinds of crimes, as many anecdotal reports describe police officers who dismiss the internet as ‘not real’, or counsel harassment targets to ‘turn off the computer’ (Hess, 2014).

²³ As noted earlier, Gaslighting is a tactic of psychological abuse whereby the abuser attempts to make their target question their own memory, perception or sanity by selectively twisting, omitting and denying information. It is named for the 1938 play by Patrick Hamilton, in which an overbearing, malicious husband makes his wife anxious by withholding information and convinces her to question her perception of reality by questioning her observation of the flickering and dimming of gas lights.

Perspectives of Conflict and Governance

Shaping the Platform: Strategy and Design at Twitter

The changes in Twitter's policies, affordances, and public communications provide a history of the platform's evolving political stances. Twitter's initial terms of service, which ran to a length of just over 700 words and served as the only demarcated rules of the platform until Twitter's 'Rules' were formally created in 2009. These terms of service were relatively spartan and laid out a bevy of disclaimers regarding intellectual property, a reservation of rights to change and refuse service to users, and a short list of 'Basic Terms' that laid out the responsibilities of users. These terms required users: to be over 13, to be responsible for the activities associated with their screen name, to be responsible for password security, to "not abuse, harass, threaten, impersonate or intimidate other Twitter users", to "not use the Twitter.com service for any illegal or unauthorized purpose", to not hack or modify the Twitter website, to not send spam email to Twitter users, to not transmit worms or viruses, and to not use Twitter to break laws in the user's local jurisdiction.

With the exception of a handful of clarifications made regarding copyright, intellectual property and the ownership of Tweets, the most significant case of political unrest seems clearly to be the aforementioned of Ariel Waldman's harassment on Twitter, and Twitter's response. As Biz Stone clearly states on the Twitter blog, and in their GetSatisfaction forums, the platform perspective was that Twitter's role is largely as a hands-off facilitator of a communications platform. To borrow Stone's

phrasing; “Twitter is a communication utility, not a mediator of content” (Schiffman, 2008).

As Twitter grew in public use, public significance and as a commercial entity, the policies and perspectives of the Platform and its operators noticeably changed. Both the ‘Austengate’ and ‘Gamergate’ campaigns drew attention to disparate user groups and user concerns on the platform — harassers and their targets alike both have dramatically different perspectives about appropriate behaviour on the platform, and as a result, positioned Twitter as an uncomfortable mediator. The head of Twitter’s UK Office Tony Wang (Politi, 2013; B. Quinn, 2013) foreshadowed changes to Twitter’s perspective towards their own roles on the platform with his apologies to the people targeted with harassment in the ‘Austengate’ campaign, in which he indicated that threats and harassment were “not acceptable in the real world, and [...] not acceptable on Twitter”, noting that “there is more we can and will be doing to protect our users.” This developed more significantly in response to the ‘Gamergate’ campaign. Dick Costolo’s leaked *mea culpa* memo, Twitter’s formative partnership with *Women, Action and The Media!*, and public statements from Twitter executives about harassment and safety on the platform each demonstrate that Twitter’s operators are receptive to the changing demands of their user base.

As discussed earlier, statements from Vijaya Gadde, Twitter’s General counsel, and Del Harvey, Twitter’s Head of Safety, begin to shed light on what appears to be relatively nuanced thinking from the platform’s operators about proactively and reactively protecting users, and about the rights and responsibilities of users and of Twitter as a platform.

As Koh (Koh, 2015) notes, Harvey joined Twitter in 2008 as its “one-woman safety department”, a role that has since grown to be an influential shaper of platform development and policy. Before her involvement with Twitter, Harvey worked with a non-profit organisation called ‘Perverted Justice’, which attempts to have a “chilling effect” on the perpetrators of child sexual abuse by monitoring and documenting sexual abuse and ‘grooming’ in regional chat rooms and on social networking platforms. The organisation is most prominently known through its association with the NBC television show “To Catch a Predator”, which televised ‘sting’-style operations to catch people in chat rooms attempting to groom underage targets for sexual abuse. Given her experiences with Perverted Justice, Harvey seems keenly aware of the challenges in policing harassment and abusive behaviours and cautions against the “Silicon Valley” tendency to attempt to engineer technical solutions to social problems (Ibid.): “I don’t think that this is something that can be solved solely by software or solely by people. It has to be a combination of the two.” To this end, Harvey discusses the importance of contextual awareness, especially in relation to user harm. Whilst Harvey notes that much of Twitter’s spam-combatting work can be automated, reports of abuse, or suicidal ideation require manual review.

Harvey’s public dealings give the impression of an expert well acquainted with her subject matter. It bears consideration that although Harvey joined Twitter in 2008, the platform has continued to struggle to deal with abusive content and harassment. This begs the question, why has Twitter continued to struggle with abuse and harassment? Unfortunately, without internal sources or observation we can only speculate as to the reason for Twitter’s continued struggles. It may be the case that

there are still institutional impediments to policy reform and change, such as the prioritisation of, and allocation resources to other business interests, such as sporting or major media partnerships, or the development of news user-facing features.

Significantly, these perspectives on combatting harassment seem now to be broadly present throughout the company. Former CEO Dick Costolo's memo on combatting harassment (Tiku & Newton, 2015) demonstrates a new, perspective. During Costolo's tenure at Twitter, the platform's rhetoric has moved away from the "free speech wing of the free speech party" and free-speech absolutism towards a more nuanced, European perspective on protecting freedom of speech through the policing of hateful speech (Rosen, 2013).

Twitter's General Counsel Vijaya Gadde (2015) demonstrates this more holistic, less absolutist approach to speech, by arguing that "Freedom of expression means little as our underlying philosophy if we continue to allow voices to be silenced because they are afraid to speak up. We need to do a better job combating abuse without chilling or silencing speech." As noted earlier, this is a concept that Fiss (1996) explored — the notion that censorship, can enhance freedom, by ensuring that the underrepresented and disadvantaged have voices that can be heard (Bernstein, 2014 2). Gadde notes Gamergate as a key example of belligerent users deploying harassing and abusive tactics to chill speech and silence other users (Gadde, 2015). Interim Trust and Safety Director Megan Cristina succinctly describes Twitter's current policy perspective; "We believe that protection from abuse and harassment is a vital part of empowering people to freely express themselves on Twitter" (Cristina, 2015).

New Normal: Tactics of Agitating Users

We can contrast Twitter's developing politics with the perspectives of users who deploy harassment tactics against other users, and their professed ideals. For belligerent users, the self-professed trolls, and the agitators of online conflicts, the politics associated with their actions tend to demonstrate a simplistic absolutism with regard to online rights, often accompanied by repulsive perspectives on race and gender. In the fashion of Poe's Law²⁴, it is difficult to discern whether or not these perspectives are deeply held or affected for shock value.

The political values contributed by harassers are often reactionary, and often particularly conservative about perceived social changes, some to the point of adopting it as a political culture in the manner of the "Neoreactionary movement" (J. Allen, 2015c). Given this, it is unsurprising that Twitter users associated with harassment campaigns and abusive behaviour on the platform have espoused racist viewpoints (often under the guise of 'race realism'), nostalgia for misogynistic behaviours deemed unacceptable by modern society, and a rejection of 'progressives' and 'progressivism' as political projects (Schubert, 2014).

²⁴ An adage akin to Godwin's Law. Poe's Law argues that "Without a winking smiley or other blatant display of humour, it is impossible to create a parody of fundamentalism that someone won't mistake for the real thing".

To the extent that belligerent users espouse arguments about their rights as internet users and participants in social platforms, we can begin to see an oddly-nihilistic, *laissez-faire*, approach to speech rights and perceived norms not dissimilar to the ‘chan culture’ associated with imageboards like 4chan and 8chan. These boards tend to have views that free speech protections should extend to violent threats, or extreme pornography²⁵. To this end, the desire for harassing users is often not simply that speech is free of restriction, but also that it is free of consequence. On one hand, this is revealing of a faith in the ability of a discursive group to self-regulate (Machkovech, 2015a), even if that self-regulation comes through abuse rather than argument and discourse. On the other hand, this is also a distinguishing feature of a culture that is irony-heavy, perhaps to the point of a presumption of bad-faith (J. Allen, 2015d). These values are evident both in the ‘Gamergate’ campaign, which was essentially started on a 4chan imageboard, and in the ‘Austengate’ campaign, where some of the convicted harassers expressed views that they did not take Twitter seriously, and that they were “trolling” for “a laugh, mostly” (P. Smith, 2015).

For the most politically motivated amongst Twitter’s belligerent users, there is a tactical awareness to the efficacy of their actions in conflict, and the broader influence that the simple existence of conflict can have in media and culture more broadly. Keegan Hankes (Hankes, 2016), a researcher with the Southern Poverty Law Center, describes this with reference to the ‘sock puppet’ tactics adopted by some users as a form of culture-jamming. ‘Sock-puppeting’, as Hankes notes, is an effective tactic in that it preserves the user’s relative anonymity, it artificially inflates the perceived

²⁵ ‘Extreme pornography’ here refers to explicit imagery or videography that would be considered illegal in many jurisdictions — such as child abuse material, or recordings of bestiality.

support an issue is receiving, and it provides an avenue for ‘satire’ or irony-based criticism (much like the Culture jamming tactics that are typically associated with anti-consumerist movements). At a smaller, more personal scale, Twitter users can utilise abusive and harassing tactics in their day-to-day interactions with other users in attempts to make them feel unwelcome on the platform, or to create a negative experience that is a disincentive to participation.

Seizing Control: Tactics of Dissatisfied Users

A key element of activism and advocacy in response to harmful campaigns like Austengate and Gamergate is that of awareness-raising. With this, users who have experienced abuse and harassment attempt to communicate the seriousness and extensiveness of the problem to other stakeholders and observers.

For Anita Sarkeesian, who had been a target of online harassment long before the Gamergate campaign started, awareness-raising is a concise but confronting explanation of her experiences, often documented with screenshots and other pictorial evidence.

“For the last two years they have viciously and aggressively come after me by any means necessary[...] I have been attacked with a constant torrent of grotesque, misogynistic harassment, everything from gendered hate speech to weaponised pornography — where they superimpose my face on pornographic images and then spam me with them— to bomb threats, rape threats and death threats directed at me,

and my family, and my colleagues. [...] There are ongoing investigations taking place for the worst of those crimes.” (Sarkeesian, 2014)

Far from simply documenting her direct experiences with harassment campaigns, Sarkeesian also uses her experience to raise awareness about different methods and tactics of harassers. In discussing more subtle forms of harassment, Sarkeesian has spoken repeatedly about online impersonation and the development and dispersal of conspiracy theories, by deconstructing her own experience to demonstrate the practices of harassment, and how falsely constructed information created to justify harassment becomes creates viral memes that reinforce a folk demon narrative constructed around feminist interlopers (Ibid.). These memes range from garden-variety sexist stereotyping, to more outrageous claims of skin-bleaching, neuro-linguistic programming and brainwashing, or allegations of con-artistry (Ibid.).

Zoe Quinn, another key target of ‘Gamergate’-related harassment responded similarly, channelling her experiences as a target of massively internet-enabled, loosely-organised harassment into her pro-bono organisation *Crash Override Network*, which attempts to provide resources for people who are experiencing online harassment, with explanations to help demonstrate issues to family members, employers, or law enforcement agents who are not technologically savvy, how to react or respond to s, how to act to defuse tensions when being SWATted, as well as advice about strong passwords, multi-factor authentication (to prevent unwanted access to online accounts), and social engineering. She explained that “So many people have been through this and not known what to do and felt helpless. [...] So much of our control over our own lives has been taken away from us. This is one way we can take

it back, to decide what happens to us, and try to help other people decide what happens to them” (Hudson, 2015a).

A common axiom given to people who face abuse and harassment online is the instruction “don’t feed the trolls”. It’s worth deconstructing this advice here, given its prevalence — this advice relates to what Noelle-Neumann (1974 44) describes as a ‘Spiral of Silence’. This “don’t feed the trolls” advice is spiral-forming, as it admonishes people seeking support or change, and criticises them for continuing to express themselves. In doing this, this advice and the people who proffer it work to reinforce barriers against targets of harassment and victims of abuse, working to prevent them from finding support or justice. In turn, it may well have a chilling effect on the speech of those targeted. Thankfully, in the cases of harassment seen in this case study, people targeted with harassment and abuse have recognised this axiom as spiral-forming and have sought to criticise and repudiate it.

Unfortunately, because tactics of harassment and abuse are significantly more severe than garden-variety ‘trolling’ and are also often used with the intent of chilling or silencing speech, any form of action taken other than surrender often prompts even more abusive messages. This places the targets of harassment in an uncomfortable position, where the “Don’t Feed the Trolls” axiom condemns them to suffer in silence, whilst the act of speaking out, almost inevitably draws criticism for ‘feeding the trolls’ or being attention seeking. Indeed, Quinn, discounts it entirely as “garbage advice” (Z. Quinn, 2015) For Quinn, Sarkeesian, and countless others, drawing attention to threats of harassment and abuse often draws the criticism that they have courted the abuse and “tricked poor, innocent kids on 4chan” (Sarkeesian, 2014) into harassing them, or that

they have confected their harassment, through an elaborate assortment of sockpuppet accounts. When social media users confront belligerent users, and seek remedy to their behaviour outside of the platform, they receive similar messages of condemnation and criticism, as Ford (C. Ford, 2015) discusses:

“Over the past 24 hours, I have been (unsurprisingly) berated for ‘ruining a man’s life’ and ‘causing him to lose his job’. But I didn’t force Michael Nolan to come to a post in which a woman shared a screencap of a man telling her she’d ‘jibber less with a cock in [her] mouth’ and make his support for that kind of misogyny known by simply writing ‘Slut’. I didn’t tell him to include his employer’s linked page on the Facebook account where he had also joked about migrants being lazy, unemployed welfare cheats and thrown around the word ‘nigger’. He is responsible for the words he chooses to write and the views he chooses to hold.”

Evaluation

Strategies

Lawlessness and Governance

Platforms like Twitter have policies that are ignored by belligerent users, and technological affordances that can be abused to cause harm to others. In response to the perceived failure of operators like Twitter to effectively and efficiently govern their platform, users can also use the affordances of a platform and subvert their intended uses to mitigate and combat abuse. In doing this, they bring attention to the ways in which platform operators fail to effectively regulate the uses of their space.

Both the failures of platforms to effectively govern their space and the tactical advances of users to address these failings are indicative of the challenges present in creating digital realms wherein users can truly participate as citizens. Whilst social networking platforms like Twitter grew spectacularly in their early years as notionally self-regulating entities — this is largely because the platform was left to its own devices as it grew in use and cultural relevance. Recent experiences suggest a need for a less *laissez-faire* approach to belligerent users' behaviour. Indeed, these *laissez-faire* approaches to users seem mostly helpful in constructing somewhat lawless digital spaces, in which users feel empowered to menace and harass one another without fear of recourse.

The structural empowerment of harassing content

Could the design of platforms like Twitter make them structurally complicit in harassment? Quinn (2016a; 2016f; 2016e; 2016d; 2016g; 2016c; 2016b) highlights that recommendation engines and other social networking apparatuses are a part of social networking platforms that perpetuates and normalises online harassment;

“Any time I create something or link to it, I’m promoting my abusers channels directly to my own audience. I raise their profile by existing. The suggestions are content-neutral and only care about how many clicks they get. They turn me into an advertisement for my own hate.

For Quinn, this makes the major platform operators an active tool for unchecked abuse, given that the affordances of these platforms can create distressing and harmful experiences even without the direct intention of a belligerent user. The ‘indifferent cruelty’ of platform algorithms mean that belligerent users do not need to directly contact their targets in order to contact them. As Quinn notes “People aren’t sending these videos *to* me[,]...] The platforms themselves are doing it for them. If this shit involved pirated Game of Thrones episodes though, none of this would be an issue.”

This may not be intentional, but it is an outcome borne of the thoughtlessness of algorithms. As Meyer (2014) and Quinn (2016f) both note, these systems are (or, at least, they appear to be) content-neutral to observers. The code and algorithms that govern these systems are anything but neutral. Code inhibits, enhances, constrains and forbids certain kinds of speech and content on the basis of variables and programming logic. If the algorithms that recommend Tweets, users, videos and links to users treat

belligerent, anti-feminist conspiracy theories with a one-size-fits-most approach, then platforms become complicit with perpetrators of harassment through indolence.

Given that algorithmic systems are fixtures of the digital landscape of platforms like Twitter, they represent an affordance that can be hostile to the participation of some users. As Quinn argues, these systems “do little to nothing about perpetually abusive content/users[but,] they help strengthen their connections to me. [...] So my choice is disappear, or help abusers profit off my own body of work.”

It’s not just the subterranean, algorithmic affordances that can empower harassing content, but also the user platforms user interface design. Consider the @-reply and @-mention functionality, which often sends push-notifications to user’s mobile devices and emails to their inbox. As Jeong (2015) notes, “an @-reply feels offensive because it ‘invades’ your online space; it’s a message delivered straight to your face because the Twitter user interface is designed to make @-replies visible”. Again, the term ‘thoughtlessness’ is appropriate. The design of some of Twitter’s most basic functions can be delightful in good-faith interactions with other users but distressing and outright harmful in sustained harassment campaigns. The platform’s affordances are ill-designed to compensate for these campaigns and leave the company ill-equipped to protect its users.

A key example of this is the push notification, which can be configured to alert users of when their tweets are ‘liked’ (formerly favoured), re-tweeted, replied to, or when a user is @-mentioned or sent a direct message. In a sustained harassment

campaign, users can receive anywhere from dozens to hundreds of @-mentions or messages each minute. (Given that push notifications also tend to wake phones from sleep, and signal sound and/or vibration alerts, this can also have a negative effect on the phone's battery life. For some harassment campaigns, the attack is multiplied by these affordances existing across multiple social networks. For Twitter's official client applications, the affordances available to users wanting to change their notifications are reasonably inflexible, offering users a binary, all-or-nothing choice.

Roles: How Users Leverage Labour and expertise in conflict

Earlier in this chapter, I identified trends in the way users operate in these styles of campaigns. For belligerent, harassing users and for the users targeted by the aforementioned users, it seems apparent that there are tactical decisions made as to the usefulness of individual skills, and labour in their conflict. For many of the users involved in conflicts, individual efforts are leveraged through use of a number of roles that are performed through their behaviours on the platform.

How loosely-coordinated harassment uses roles that eschew direct hierarchy

The Austengate and Gamergate harassment campaigns analysed in this case demonstrate both a disorganised campaign (Austengate), and an organised campaign (Gamergate). For disorganised campaigns, the practices and labours of harassment are not strategically organised. Instead, belligerent users take social cues from other, likeminded users, with similar perceptions about their digital space.

Gamergate, on the other hand, demonstrates a much more organised system, wherein the knowledge and roles of abuse and harassment are shared. In this respect, we can chart roles and responsibilities of users in this style of campaign.

Thought Leaders and Campaign Organisers

At the 2015 XOXO Conference, Zoë Quinn spoke about the tendency for observers and reporters to use the anonymity of some harassers as a convenient explanation for their belligerent behaviour. Quinn (2015) dismisses this explanation, offering her experiences as evidence:

“My worst and most egregious harassers use their legal names and make a living off of this. They have Patreons²⁶, they get paid for YouTube clicks, there’s a number of people [...] that have made bank off of the backs of harassing people like me... And they use their names. They build their brand.”

With this, Quinn is likely referring to a number of high-profile people associated with Gamergate, or other anti-feminist movements, like the Breitbart columnist Milo Yiannopoulos, the prominent anti-feminist YouTubers Carl ‘Sargon of Akkad’²⁷ Benjamin, Phil ‘thunderf00t’ Mason, or the handful of amateur documentarians

²⁶ Patreon is a crowdfunding/patronage service, where users can pledge recurring amounts of money to registered creators on a ‘per creative work’ or ‘per month’ basis.

²⁷ Benjamin’s channel name (pseudonym, online persona, and personal brand) ‘Sargon of Akkad’ is an overt reference to the to the Akkadian emperor Sargon, who is known for his conquest of Sumerian city states in the 24th and 23rd centuries BCE.

reporting on Gamergate from the ‘gamer’ perspective, like the pseudonymous ‘MundaneMatt’, who’s videos about Quinn played a significant role in the early days of Gamergate, or Jordan Owen or Davis Aurini who attempted to crowd-fund a documentary called ‘The Sarkeesian Effect’. Discussion of the roles of these stakeholders, perhaps, is where definitions of what constitutes abuse or harassment becomes a little complex. In most cases, these people have little or no direct communication with Quinn, Sarkeesian or other targets of harassment.

In many cases, their tweets, articles and videos are strongly worded critiques relating of feminism, political correctness, social justice campaigns or ‘progressives’, although strident criticisms of individuals are not uncommon. These articles and videos would likely not satisfy any conventional legal definitions of harassing or abusive content, and they do not satisfy the definitions of most platforms, either. However, these videos accomplish two things that seem to establish them as a part of the abuse and harassment ecosystem. First, these videos become a part of social network’s content ecosystems, and as a result, fodder for algorithmic recommendation engines. The result of this is that these sorts of videos become often presented as a “background radiation” (Z. Quinn, 2016h) to their online presences.

Second, the videos provide talking points — arming their audiences with both a perspective on recent events, and furnishing them with new arguments (Pless, 2015a). In providing these talking points, the thought leaders and figureheads of the Gamergate movement can essentially rely on their audiences as a proxy for their own involvement in harassment. Pless (2016a; 2016b) documents examples of this behaviour, and the

complex interweaving relationships between the Thought Leaders of these movements, and the audience members who follow them.

‘Foot soldiers’ and ‘Useful Idiots’

As indicated in the preceding paragraph, a factor in the influence of the Thought Leaders in harassment campaigns like Gamergate is the willing participation of their audiences. In a number of cases, audiences have responded to the videos of thought leaders by acting as ‘foot soldiers’, by “trolling, hounding, and even stalking” people criticised in Youtube videos on the behalf of the thought leader, if not at their behest (Pless, 2016b). In other cases, thought leaders rely on the efforts of their audiences, and other individuals, as volunteer researchers for their work. Pless (2015b) highlights a handful of cases in which Breitbart columnist Milo Yiannopoulos has written articles about social justice campaigners or targets of Gamergate, that seem to use information obtained by hackers on 8chan’s notorious *Baphomet* board — a board where ‘raids’, hacks, and doxes are planned and publicised. In the same way that targets of harassment have sought legitimacy for their concerns through coverage in mainstream media sources, the Thought leaders of movements like Gamergate provide an alternate media narrative, and an avenue in which the advocates of Gamergate can seek affirmation in their cause.

Whilst asking “to what extent did Gamergate *know* that it was an anti-feminist hate mob?”, Ian Danskin (2015) asserted that there were essentially two ‘kinds’ of participants in the Gamergate movement: One group, who were not necessarily

interested in journalism, ethics, or even video games, but who found the group a useful cover for their harassment of women and minority participants in gaming spaces, and a networking opportunity to meet other likeminded belligerents. The second group was actually interested in the ‘ethics in games journalism’ arguments. The second group were ‘useful idiots’²⁸, providing the appearance of a reform movement for the first group, although the influence of the first group likely resulted in the ‘ethics in games journalism’ concerns that tended to manifest in criticisms of the roles and opinions of women in games writing, and not more systemic concerns, like the firing of writers who are critical of advertiser’s products (Plunkett, 2012), or the economic influence of review scores (Kain, 2012) on games publishers.

How Targets of Harassment and Their Supporters work to mitigate abuse

For the targets of harassment, the division of labour is largely ad-hoc or post-hoc. Often, users are targeted as individuals, rather than as a part of a group, which makes the experience an isolating experience. That said, there are a number of identifiable roles that seem to regularly be filled in the conflicts observed for this case.

Campaigners and Engineers

The tactics discussed earlier, of Advocacy, Circumvention, Intervention and Litigation are often pursued to both seek redress or reform for the targeted user, but

²⁸ This phrase, ‘useful idiots’ refers to people who a) are perceived as propagandists for a cause whose goals they are not cognisant of, and b) are used cynically by the leadership of that cause. This phrase is often misattributed to Vladimir Lenin. The provenance of the phrase is unknown.

also to provide support for other users who find themselves the target of similar styles of campaign. This is particularly clear with both Sarkeesian and Quinn, who both created guides and support websites designed to help targets of online harassment: “*Speak Up & Stay Safe(r): A Guide to Protecting Yourself from Online Harassment*” (Feminist Frequency, 2015), and Crash Override Network (Z. Quinn, 2015). Similarly, Randi Harper’s development of tools to ameliorate Twitter’s poor blocking affordances provides another toolset for users that are experiencing these kinds of abuse.

Supporters and Reporters

When people who have been the target of harassment have friends, colleagues or a supportive audience, that group can provide moral support through the experience, but they can also perform assistive roles. Guides like Feminist Frequency’s ‘Speak Up & Stay Safe(r)’ and Crash Override both extoll the usefulness of delegating some social media responsibilities to a trusted friend or colleague who can volunteer as an intermediary for abuse in their stead. Twitter does not formally support delegating access to personal user accounts between users — which means that supporters who act in these delegated roles are often entrusted with passwords and other private information. Supporters can also make use of the ‘report’ functionality of Twitter and other platforms, either in isolation, or as part of a larger campaign to flag abusive content. This is sometimes stymied by the policy affordances of platforms. In particular, Twitter has been criticised for policies that only allow those directly affected by abusive and harassing content to report it — thus preventing the possibility

for users to act as ‘good citizens’ and to have their reports of infringing content or behaviours acted upon (Feminist Frequency, 2015; Matias et al., 2015).

Governance and Belligerence

For Twitter as a governing platform, the difficulty of policing abusive and harassing behaviours represents a deeper political and philosophical discussion about the boundaries of acceptable behaviour and speech that the platform’s operators must negotiate with users.

Defining or identifying an end to this case and conflict is a challenging task, the boundaries of the conflict are porous — as I will highlight in the following discussion chapter, many of the key figures associated with the Gamergate campaign turned their attention to the Trump 2016 election campaign, or to other culturally adjacent focuses. Twitter remains particularly well-known for a deeply belligerent subset of its users. For the people targeted in the Gamergate campaign, the trauma and disruption are not necessarily a fixed, finite and isolated chapter of their lives— rather, Gamergate was the heralding event for months, or years of harassment ranging from targeted campaigns, to spurious rumour-mongering, to drive-by insult, and beyond.

In 2018, reflecting during on a review Zoë Quinn’s memoir *‘Crash Override: How Gamergate (Nearly) Destroyed My Life, and How We Can Win the Fight against Online Hate’*, Leigh Alexander expresses a profound frustration at the glacial pace of reform on social platforms, and how the discourse has responded to her and Quinn’s experiences:

Zoë's experiences were awful, but the vulnerable aftermath, which everyone carefully tongues like the wound left by a missing tooth, is awful too. Here's the thing that's really awful: No one is going to let us forget it. It's now been four years since I was targeted for an essay I wrote, essentially in Zoë's defense, identifying the source of the harassment behavior as cultural displacement on the part of a coddled, insular male audience [...], but I'm still being asked to appear on Harassment Panels, to write uncompensated Harassment Articles and uncompensated Harassment Book Reviews, to perform emotional labor on the topic of Harassment, all in the name of feminism. The implication is I cannot refuse because I owe people *lessons*. In addition to what I experienced, I also owe my observations and expertise, and it seems I will owe it for some time. And yet what's the social value of all these discussions we are constantly being asked to revisit? People keep tenderly and breathlessly asking us to unearth these things, so they can speak in hushed voices about our bravery and heroism—what are they actually getting out of our Experience? I'm not any sort of hero because I was harassed. Nor is it the fact that I was harassed that makes me a feminist. These things seem obvious to me, and yet this is where the conversation stops.

Have platforms changed their policy? Has anyone built tools to improve the experience of marginalized people online? [...] Everyone seems to imply I really ought to keep donating my labor on this topic because it will supposedly help others. People keep saying they "learned so much" from Zoë's book, they keep saying how it's an important work, but what will they fucking *do* having read it? Anything at all save wringing their hands and asking for more titillating stories of abuse to pad out their arts festivals and make them seem progressive? (Alexander, 2018)

I've reproduced Alexander's passage at length to end this chapter and lead into the following discussion chapter — as it provides a powerful underline to this case study — a case study that attempts to meaningfully learn the Gamergate campaign, and the responses to conflict and trauma, and the interplay of user and platform responses. Further, Alexander highlights the need for research like this to be leveraged in pursuit of structural and policy reform on social platforms. Alexander's statement here should serve as a reminder to researchers whose work touches this field that while scholarship in at its most essential when leveraged to influence reform.

This chapter provides a succinct, thorough analyses of elements of Twitter's governance practices, and the tactical manoeuvres of users in both Gamergate and a significant preceding event in 'Austengate'. Whilst I researched and later wrote this chapter, I did so with a hope that research works like this could be used to influence policy decisions — and affect platform reform. Unfortunately, Twitter, as an organisation, has become a rather conservative entity — choosing to make small, 'sufficing' changes to occasionally address the most critically damaging issues to public perception of the company, In the following chapter, I will underline the extent to which the Twitter's policies have largely stayed the same in the years since the advent of Gamergate. That said, Twitter's organisational complacency should not be allowed to be the end of a search for improvement in this area. Rather than looking to platforms for solutions, users' have demonstrated an aptitude for remix and repurposing of platform spaces — this, perhaps, is our path to meaningful change in these spaces — a renewed focus not just on the paths designed by platform owners, but the paths built and improvised by users.

Chapter 7: Discussion

Strategies

Governance

Governance is managed on social platforms through a mixture of formalised policies and platform design. This manifests largely through the development and deployment of institutional strategies for governing and negotiating with the platform's user populace but is more subtly seen in the technological affordances that platforms provide — either through the user-interface of the platform, or the application programming interface (API) of the platform. Given this, we might consider that platforms have two key pillars of strategic power — the first, as institutional policies, and the second, as platform design.

In this thesis, institutional policies were clearly demonstrated as a key component of platform governance in both the Facebook and Google+ case studies, as these policies served as governing documents, and thus *de facto* laws for the platforms, as well as *prima facie* justification for platform policing and other governmental behaviours. Essentially, by having these policies enshrined in a digital list of rules or policies, platform operators could use the veneer of bureaucracy in an attempt to deflect suggestions of capricious or inconsistent governance. Whilst the text of a policy can be a particular issue in a conflict, often the most troubling issues for users relate to the enforcement of policies — or, as in the Twitter Case study, the lack of enforcement of certain policies. Conversely, not all of a platform's institutional policies will be user-facing. We have now seen leaks of guidelines used by Facebook's moderators a handful of times in reports from sources like Gawker (Chen, 2012a), and

The Guardian (Hopkins, 2017). For some, these leaks have been particularly galling, illustrating the boundaries of what platform operators have deemed appropriate for the space. A recent leak (Hopkins, 2017) illustrated this clearly with examples of what the platform regards as “credible violence” — posts where a target of violence is identified, and a statement of intention is made. This relatively simple set of criteria meant that posts like “Someone shoot Trump” were subject to moderation and removal, but posts like “To snap a bitch’s neck, make sure to apply all your pressure to the middle of her throat” were not.

An often-minimised concern of these policies, user-facing or otherwise, is the human element — the opinions, ideologies, and social norms influencing the creators of the policies, the moderators enforcing the policies, and the users governed by these policies. That element will always be present — even if algorithms and systems of automation are deployed to minimise (or eliminate) the role of human moderators, the platforms and their operations will always be shaped by those who create the platform —the engineers who create and code prescriptive technologies (Franklin, 1999 17–19) according to their understandings (Ullman, 1995)—, those who moderate the platform, and those who use the platform.

The other pillar of platforms governance is the platform’s design — that is, the affordances of the platform for users and their behaviour. This manifests in two ways — the overt design of the platform (the user interface), and the underlying design of the platform (the platform’s programming, and application programming interface (API)). Governance through design can be challenging to define through this style of academic research, as we can only truly analyse the user-facing effects and influences

of design. Given this, we can clearly see the repercussions of prescriptive design decisions that mandate both first and last names (as investigated in the Google+ case), or conformity to gender or sexuality norms. Less visible, however, is the influence of the platform's subterranean elements — the “algorithms” that programmatically shape and govern interaction on these platforms. Application Programming Interfaces (APIs) somewhat complicate this, as while they do not necessarily reveal the underlying workings of social platforms, they present the opportunities for even more diverse experiences of that platform. As noted in the Twitter case, some users turned to tools that leveraged the power of Twitter's API, or to third party Twitter clients that offered more useful features than the platform itself.

‘Sufficing’

In each of the cases analysed for this thesis, users have often had a goal or desire for the operators of platforms to change either the policies, the practices of moderation and governance, the technological affordances, or a combination of these three overarching themes. To this end, it is useful to look to the end of these cases and ask: “What changed?”. For better or worse, the answer is “not that much, really”. Platforms and their operators have become proficient and gained institutional knowledge and experience through repeated conflict with users. Given this experience, we can suggest that platform operators have become canner at negotiating conflict from the perspective of the platform. To wit, many of the responses seen in this research are emblematic of ‘sufficing’ strategies. ‘Sufficing’ (Gross & Guerrero, 2000) is a conflict management strategy, that essentially works to find a path of both least resistance and least effort — the minimum possible change to address a conflict. As such, sufficing is a comparatively reactionary strategy, a strategy born in response to

tactical conflict. Sufficing strategies, challengingly, allow for platforms to construct the appearance of change, through policy announcements, press coverage (like Facebook's many statements about the platform's policies relating to photographs of breastfeeding), and performed contrition (like Dick Costolo's leaked *mea culpa* for the Twitter's harassment issues).

Of course, the practicalities of change and reform can be complex, and users tend not to expect overnight change — but it bears consideration that each of the cases analysed in this research have effectively spanned multi-year timelines but seen relatively little in the way of systemic change. The text of Facebook's policy shows particular development over time, but the implementation remains unchanged — photographs including nipples and areolae remain forbidden, regardless of their proximity to nursing infants. Google+ notionally abandoned its policy during the course of this research but made no systemic changes to the affordances of the Google+ platform's affordances for naming. Twitter have perhaps made the most drastic changes, although many of these have only begun outside of the research period for this thesis. Twitter's recent efforts notwithstanding, users remain frustrated with the platform's inability to address their concerns for safety and abuse on the platform.

'Sufficing' is perhaps the least-challenging response for platform operators, because it provides the least disruption to existing business operations. Seeking the minimum-effort course of action allows for platforms to deploy the simplest form of redress, without upending any commercial plans. That said, as user frustration's with platforms grow, we may yet see a decrease in the efficacy of sufficing strategies for platforms — leading to eventual changes through social pressures (users seeking

change through organisation, and collective action), political pressures (users seeking change through regulatory redress and reform) or market pressures (users leaving, and in doing so, devaluing the platform). This is unlikely given the current dynamics of these platforms and will require more significant organisation and tactical work at the hands of users.

Tactics

Earlier, in the case studies of this thesis, I posited themes and sets of user behaviours we could classify as tactics. I have highlighted on- and off-platform protests, and different thematic clusters of tactical behaviours, ranging from the ‘advocacy’ (appeals to others for political support), ‘circumvention’ (the subversion of platform infrastructure), ‘intervention’ (the use of external social pressures in inter-user conflict), ‘platform-hopping’ (the use of multiple platforms to skirt the reach of a platform’s rules and influence), ‘regulation’ (appeals to external legal or regulatory pressure or reform), and ‘belligerence’ (hostile, interpersonal uses of tactics). Here, I have refactored these themes into two over-arching categories: Organisational tactics, which appeal to make broad, political change to a platform and its governance; and Circumventive tactics, which often rely on a user’s first-hand knowledge or understanding of a platform’s spaces, a platform’s limitations, and the platform experience as a tool for subverting the platform’s governance. It should be noted that neither of these overarching categories are reliant on motivations (extrinsic, intrinsic or belligerent), as the tactics described in these categories can be used for personal benefit, social benefit, or social harm or interpersonal harm. Similarly, the harms

(proximal or distal, significant or mild) that might encourage, necessitate or occur from these tactics are not key factor for the purposes of this categorisation.

Organisational Tactics

Through user behaviour in each of the cases we can see organisation—in the sense of collaborative or collective action—is key focus of tactics in agitating for policy change with social media platforms. Users tend to engage in practices of both regulatory, interpersonal advocacies with a mixture of both on- and off-platform spaces, as is contextually appropriate for the conflicts they are engaged in. As noted in the Facebook case study, users have deployed on-platform advocacy to communicate their concerns and frustrations to other users, in an attempt to highlight their experiences of the platform, and the injustices that they experience. Similarly, users also practiced off-platform protest, to communicate more fully and in more detail without the risk of punishment by the platform. This has been a particularly effective system for communication in the Facebook case study, where users could document the platform’s policing, and contextualise their frustrations and criticisms with documentary evidence that illustrates their criticisms of the platform. Moreover, once users are aware of the policies and the methods with which they are policed, they can use the platform in a way that essentially solicits platform interventions. Borrowing from George (2013 389–392), we can describe this is a form of ‘political jiu-jitsu’, wherein users have leveraged the difference in political power to throw platforms off-balance and revealing some of the mechanics of their attempts to wield power through policy and policing.

In many of these cases users also appeal to other authorities, seeking coverage in industry or mainstream press, or an audience with their elected representatives. As a form of off-platform protest, these efforts can be particularly effective in shifting popular opinion, but not necessarily effective in any soliciting meaningful reform. Whilst external regulation is sometimes sought to remedy the issues with these platforms, there are some recurring issues, relating to the regulation and external governance of a globalised platform on the internet (namely, issues relating to legal jurisdictions, differing legal and regulatory frameworks, and the varying powers (or willpowers) of external authorities.) For users, existing external regulation can be a valuable source for potential redress if either the operators of the platform, or other users of a platform are using the platform in a manner that would run afoul of legislation. This is particularly clear with cases of abuse and harassment, where existing legislation in many countries already provides for the types of abuse and harassment encountered online under criminal law. This can be complicated by a lack of training or interest from law enforcement officials (Hess, 2014), but the recent mainstream surfacing of online harassment in previous years may yet herald more focus on these issues for police training in the future. That said, a further complication arises when the key participants in a conflict are not in the legal jurisdiction of the reported-to enforcement agency. New South Wales police might be able to find and charge a law-breaker in the same state of Australia, but if the law-breaker is further away, interstate or internationally, the reach and practicality of policing diminishes.

As a corollary to this, some users tire of pressing for changes in governance of a platform directly, and have instead sought reform in the regulatory environment that platforms operate in. This is particularly evidenced by occasional user-led campaigns

to change or reform Section 230 of the *Communications Decency Act* in the United States, a piece of legislation that provides ‘safe harbour’ for internet service providers and digital platforms, absolving them of liability for the content produced by their users.

Circumventive tactics

We can also identify some key behaviours present in the tactics of users who attempt to circumvent parts of social platforms: the circumvention of strategies — circumventions that react and respond to issues in the construction or operation of social platforms; the circumvention of failures — circumventive measures that attempt to ameliorate the perceived failings of the governing platform; and the circumventions of protections. For users’ different styles of circumvention become tactically appropriate depending on the policing they intend to avoid or compensate for. In both cases, we can draw parallels to de Certeau’s (2002) conception of tactics used by the inhabitant of a city —as circumvention is perhaps the clearest example of users (or walkers, in de Certeau’s city) deploying their skills and knowledge of a platform to better navigate and survive inside it.

Circumvention of Strategies

Circumvention of strategies tends to occur when users have become frustrated with the formal structures and procedures of governance on a platform and subvert them. In this thesis, I noted that this occurred through the use of both on- and off-platform actions. For some users, the circumvention of strategies is accomplished through techniques of conflict-avoidance, like articulating from making public posts on a Facebook page, to semi-private posts in a ‘Closed’ or ‘Secret’ Facebook Group.

These sorts of tactics understand the weakness in platform operator's methods of policing and exploit them.

Circumvention of Failures

The Twitter case study in this thesis provides excellent examples of users circumventing the failure of social media platforms. In that case, users primarily grew frustrated with Twitter's poor response to harassment and abuse on the platform, and used or developed third party tools to circumvent the platform's failings and augment their experience of the platform with tools designed to mitigate abuse and harassment that were at the time of development significantly more advanced than the offerings in Twitter's official client or web interface.

Circumventions of Protections

Finally, the circumvention of protections is most keenly used by belligerent users, who test and become aware of the weaknesses and tensions in moderation systems and interaction systems on platforms so that they can deploy them in conflict with other users. For some belligerent users, this is as simple as exploiting the relatively lax requirements associated with opening an account on the platform, and thus having an arsenal of disposable social accounts — mitigating the cost of punishment. For other users, this involves experimentation with systems like Twitter's self-service advertising tools (Auernheimer, 2015a) to determine the affordances of these tools, and then wield them for the purpose of abuse.

Boundaries of Platforms

The boundaries of social media platforms are often quite porous, as web developers use platform APIs to build cross-service features and technologies, and users often travel from platform to platform using different features and meeting different users along the way. Given this, we should consider that the issues of governance in an expansive, occasionally intermingling ecosystem of internet platforms means that many user-user and user-platform conflicts may have trans-platform governance issues — issues where the boundaries of a platform’s governance or the jurisdiction of their policing come into play. When boundary-aware tactics are used in the pursuit of reform, this tends to involve the use of non-platform spaces like blogs, or conventional media platforms as opportune locations from which to criticise platforms, or bring light to platform issues, without risking censorship, or another moderation tool in platform arsenals.

The Twitter case study, particularly Gamergate, saw the deployment of platform-hopping as a belligerent tactic, as hostile users seemed to use different platforms for different purposes — planning harassment with the plausible deniability of highly anonymised and ephemeral chan-style imageboards, delivering talking-points and subcultural news through Reddit, and partaking in either conflict or abuse on platforms used by their targets, Twitter, Youtube and many others, sometimes even by hyperlinking to off-platform content to further complicate governance. This posed a challenge for Twitter, as the platform’s moderators seemed slow to grasp the significance of off-platform content as an influence and a factor in abuse. This is seen in microcosm in what Reddit users term as ‘brigading’ — where users of some of the platform’s semi-federated message boards being to use boards they dislike in hostile,

belligerent ways, in an attempt to make the users of that board miserable. More broadly, platform-hopping tactics throw into sharp relief the technological boundaries of a platform's reach and power. This may signal a need for the development of inter-platform regulatory frameworks, or reforms to platform policy that stress the importance of contextual awareness.

One aspect of this behaviour is rather more interpersonal — occasionally leveraging the social connections illustrated on social networking platforms so that one can better see the possibilities for the use of social pressures in inter-user conflict. This was noteworthy in the Gamergate case, where Pearce (in True, 2014) described the use of Facebook's social graph to report harassers to their mothers, but was not seen in other cases due to less overt inter-user dynamics. Further research could analyse the dynamics of user-user conflict on social platforms with the hopes of identifying how users leverage social pressures in online conflicts.

Tactics for ill

Belligerence is not a tactic per se, but it is motivation of tactics that bears discussion here. Whilst it may be tempting to view platform-user conflicts as narratives of user's struggling against platforms as modern-day David and Goliath conflicts, but it's worth noting that many of the tactics noted in this thesis can be used for ill. Indeed, off-platform protest puts users lobbying for platform reform outside of the reach of a platform, but it also puts users scheming to maliciously abuse other users outside of the platform's control too. Circumventive tactics might allow for users targeted by

harassment to deploy Twitter's API in a manner that helps them manage the incoming waves of hostile and abusive Tweets, but similar tactics have seen belligerent users circumvent Twitter's advertising platform to wield the platform in a way that undermines some of the platform's existing safety and anti-harassment affordances.

The response to the misuse or circumvention of a platform's architecture should not necessarily be to address that misuse, or to make that circumvention impossible in the future. A reactionary response to online harassment might have seen the obliteration of rights and affordances for pseudonymous or anonymous users on some platforms, but this works to the disadvantage both of users who would use the guise of anonymity or pseudonymity for misbehaviour and social ill, and the users who seek the protection that these affordances allow to find relief from the abuse or harassment that might already follow them. Platform operators must learn to account for good and ill in their responses to users circumventing the platform. Although platforms may deploy their policies as de-facto laws in their digital spaces, there is value in puzzling out the civil disobedience from generalised misbehaviour—in identifying both good-faith and bad-faith actors amongst platform dissidents—so that platform operators can more readily engage with criticism and so that platforms can become more keenly aware of the ways in which they can be instrumentalised for abuse.

The Interplay of Strategies and Tactics

Complications of the Strategies and Tactics Framework

During the planning and research of this doctoral project, I looked to Michel de Certeau's (2002) theoretical lens of 'Strategies' and 'Tactics' as a useful lens and framing mechanism for understanding the dynamics of power in platform spaces. de Certeau's conception originally communicated these power differences with two metaphors — firstly by comparing 'the city' with its inhabitants, and secondly, comparing employers to workers. Both of these metaphors have some appropriateness to the somewhat similar dualism between platforms and users — as platforms operate not only as a digitally inhabitable space, but also as an affective employer of sorts — as social platforms often include systems that provide opportunity and reward (although (with a few exceptions) generally not fiscal remuneration) to the majority of its users for their affective labour. As noted in the literature review of this thesis, this is perhaps an under-stressed element of the platform for social media users — that the fruits of their labours are one of the most tangible ways in which social media platforms can hold value for other users. In this respect, social media platforms are remarkably similar to multiplayer video game platforms. As Greene and Joseph (2015 227) note, this is a result of shifting economic environments, that demand and re-purpose audience labours into affective constructions and even the creation of digital commodities. That said, whilst de Certeau's 'Strategies and Tactics' framework has been useful in the analysis of these conflicts, this thesis has perhaps highlighted that application of de Certeau's framework is best suited to the analysis of hierarchical

power relationships — like those originally investigated by de Certeau. Whilst there are certainly still hierarchical relationships in this space, the nature of these social platforms is often also highly networked — that is, conflicts in platformed spaces tend to have both vertical (hierarchical) and horizontal (interpersonal) tensions, as changes to governance would be likely to affect many users — including those who had not been involved in a conflict, or (in the case of belligerent users) those who may have directly contributed to conflict through their actions. This highlights the complexity of platform and user relationships, and presents an opportunity for more complex research, focusing on these relationships. To that end, it may be useful for future researchers to engage with literature and discussions of collective action and unionisation, with an eye to both the collaborative dimensions of these arrangements, and the interpersonal conflicts that occur in union/non-union worker disputes. As Suzor (2010 88-89) noted, users in gaming platforms have previously worked to form disruptive protests — with some working to exploit technological weaknesses and force game servers to crash, which in turn makes the game server unavailable for all players during the outage they cause. Similar tactics could be conceivably deployed by users working to overwhelm social platforms — perhaps by finding poorly-automated or labour-intensive processes in the platform and overwhelming the moderators or customer service attendants with a bevy of bad-faith reports, complaints or issues.

Towards a Theory of User-Platform

Whilst de Certeau's (2002) concept is perhaps most keenly observed in the relationship between factory operators and workers (a relationship which has some parity in platforms and users), we can also see similar relationship used in order to

explain and contextualise the platform and user relationships. Perhaps chief amongst these is the “Product / Consumer” relationship which often situates the platform as a product and the user as the consumer — especially in the ‘love it or leave it’ encouragements from observers critical of user attempts to affect change on social platforms. This metaphor is often severely reductive when deployed in this context, as the networked effects of social media platforms mean that the realities of its “product-ness” are vastly different from that of any other conventional consumer product. Soft drinks and sports shoes, for example, do not ‘lock in’ the friends and behaviours of their customers to a proprietary system, or prevent customers from playing sports with customers of other shoe brands. Alternatively, some explain this relationship with users fulfilling the role of product, and advertisers fulfilling the role of customers (Rushkoff, 2011)²⁹. Whilst this is, again, reductive, this characterisation serves to highlight the complexity of platformed spaces and the differing goals of their many stakeholders. Of course, as platforms grow in reach and in political capital, comparisons to democratic states and citizens become more apt.

This conceptually allows a more thorough understanding of a user amongst users — the role of individual citizen in concert and contrast with the roles of citizens as a group. Similarly, this allows for a more honest engagement with the notion of power and control in these relationships — platforms and their operators often operate at a distinct imbalance of power in relation to their users. Moreover, they can design and operate the levers of control for the platform — creating roles, rules and even punishments for users. Whilst it may be tempting for some to diminish these rules and

²⁹ Notably, Facebook describes its support channels for advertisers as a “customer service” division.

punishments in comparisons to the laws and punishments of a nation state — it's worth considering that systems like prisons, or punishments like death, incarceration and exile all have had some degree of analogue in platformed digital spaces³⁰. However, as noted with the product/consumer metaphor — the comparison to democratic states is not entirely apt. Whilst a state-like understanding of platforms would serve to more fully understand the network effects associated with being a 'citizen' of these digital spaces — it must be noted that these many of the most popularly used platforms are often not meaningfully democratic. Grimmelman (2009) and Schneier (2013) have suggested looking to feudal societies for an appropriate metaphorical relationship — situating platforms as both kingdoms and feudal lords of sorts, and users as the serfs.

There are aspects and dimensions to each of these relationships and their use as frameworks to describe the dynamics of platform-user relationships, although our understanding of this relationship may be struggling against the confines of these hierarchal frameworks – especially given so many differentials in political power amongst stakeholders with varying levels of fiscal, political, or sociotechnical capital. Feudalism could serve as a useful, if imperfect, metaphorical lens for this type of research, especially given the tendency for these relationships to be seen and regulated a more conventional consumer-vendor relationships. The cases in this thesis have revealed the weaknesses of using these hierarchal frameworks for understanding. To

³⁰ A particularly interesting example of this is the Square-Enix massively multiplayer online role-playing game *Final Fantasy XIV*, which features an in-game location known by players as “GM Jail”, a virtual location that player avatars are whisked away to when their in-game behaviour has been noted as objectionable or against the terms of service by the game platform's moderators, the Game Masters (GMs). Players can be confined to the dingy virtual dungeon-space for hours of in-game time, although this 'sentence' is often commuted if they show contrition.

better understand user and platform relationships, we will need to look beyond hierarchy and towards models of networked power relationships. To that end, further research should interrogate the rights, behaviours, and crucially, the expectations of both users and platforms in these disparate relationships of power. It is apparent that there are existing norms and existing expectations for the operation and governance of social networking platforms, just as the platform operators have expected norms of user desires and behaviours (although operator's norms can often be US-centric or even Silicon Valley-centric).

This thesis has observed in particular the conflict of norms regarding identity, speech, safety, nudity and sexuality. In each of the cases analysed, users chafed against the norms and restrictions set in place by platform operators — from Facebook's and Google's normative prescriptions for identities, to Facebook's zealous policing of breastfeeding, and Twitter's slow progression from free-speech absolutism. Beyond these, concepts relating to property (intellectual and otherwise) have been similarly complicated in user-platform relationships by both practical and legal means. In essence, many social platforms make little-to-no effort to clearly establish or communicate labour, property, rights, responsibilities or entitlements in their digital spaces, leaving their interpretation to users, and their enforcement subject to the norms or potentially capricious decisions of moderators and developers. This is both a critical area for further research, and an important avenue for potentially reforming the platforms of today into more deliberative digital spaces.

Theorising Governance

The involvement of users in deliberative processes?

Inter alia, network effects, corporate institutional power, and effective lobbying have led to the creation of what Schneier (Schneier, 2013) terms “the feudal internet” — a status quo in which a handful of internet companies (Amazon, Apple, Facebook, Google, and Microsoft) have interrelated market duopolies, and between them, a market capitalisation of approximately three trillion US dollars. This poses a problem, Schneier argues, as users are required to essentially pledge themselves as serfs to these platforms (as feudal lords). Schneier’s use of feudalism as a metaphor for this platform-user relationship is insightful, given feudalism’s system of obligations in both directions. Grimmelmann (2009) makes a similar judgment of the Second Life video game. Whilst Second Life clearly creates a digital proxy for land in the virtual properties portrayed in the game, the internet companies Schneier focuses on have constructed property in relation to products and services. As noted earlier in this thesis, social networking platforms like Facebook have courted a loose understanding of property on their platform, by referring to features and services as belonging to users — “your timeline”, “your profile”, “your page”, “your notifications” and so on.

As highlighted earlier, this is a construction of ‘notional ownership’ that platforms deploy, so that users can understand certain interpersonal and software-constructed boundaries in the digital space — situating some zones as particularly user-related (“your timeline” or “my profile”). This pulls double-duty by implicitly reinforcing the platform’s argument that users ‘own’ the data that they provide to the platform — both at a level of intellectual property and legal culpability. Therefore, looking at this through Schneier’s lens of ‘feudal’ internet arrangements, we can see

that users have essentially pledged their services as both consumers and content creators, and in response — platforms perform the roles of feudal lords, pledging plots of digital ‘land’ in the form of pages and timelines (although, in the case of Second Life the ‘land’ metaphor is perhaps more literal) services of convenience, social support and security.

Whilst ‘feudalism’ is particularly apt as a comparative metaphor for the current power dynamics and relationships between users and platforms, it fails to fully encapsulate the relationship between platforms and users. Thus, it is critical that researchers also look beyond the status quo of this particular congealment of capitalism, market forces and network effects, to the possibilities that remain for future and alternative structures of governance and power in digital spaces.

A proposal to substantially change the power and governance structures of social platforms could involve either platforms or external regulatory pressures elevating users to as more significantly recognised stakeholders in matters of governance, and in turn, implementing systems and processes designed to support and accommodate this. This could involve not only empowering users, but getting them involved and invested in the platform, its goals, and its business practices. As noted earlier in this thesis, Facebook deployed a somewhat underwhelming attempt at user-deliberative governance, by using a poorly-advertised voting system that required a significant proportion of users to vote on a given issue for their decision to be binding (Schrage, 2012). This system of abstracting community and governance issues to a popular vote was flawed in a number of ways, but chief amongst them was the platform’s lack of a compelling argument for user involvement. Whilst the idea of affecting policy and

governance is worthy in the abstract — It will be challenging for a platform like Facebook communicate to an individual user (any one in millions or billions of users) that they have a meaningful role, or stake in platform policies?

Perhaps part of this can be solved by articulating the role of ‘platform user’ less as a customer to the platform’s service, and more as a citizen of a platform’s space. With this conception of user as citizen, we can imagine roles and responsibilities for users in platforms that have shaped their governance structures to better facilitate deliberation, democratic processes, and civic responsibilities. Much as the citizens of a nation might be required to perform jury duty in relation to criminal legal proceedings, platforms could work to involve users in moderation processes. Whilst this sort of concept might seem novel in the context of social media platforms like Facebook, and Twitter, this practice has some precedent in online gaming platforms — specifically, *League of Legends*’ ‘Tribunal’ system. In the Tribunal system deployed in Riot Games’ *League of Legends*, players of the game are given the ability to participate in a moderation framework that gives players a role in adjudicating reported conflicts, and in punishing other users who have been reported to the *League of Legends Tribunal* for having transgressed the game’s social rules —which are often less about gameplay mechanics, and more about the transgression of politeness or civility in social interaction, such as the presence of overtly hostile behaviour— harsh language, slurs and epithets for example (Scimeca, 2013).

Researching the implementation and reception of the ‘Tribunal’, Kou and Nardi (2014) found that it had become something of a ‘hybrid’ system for governance, whereby players constructed formalised understandings of the game platform’s rules,

but also leveraged and contextualised these rules with the social norms they had come to see as players of *League of Legends*. Riot Games' experiences with *League of Legends*' 'Tribunal' presents an alternate approach to moderation for social media platforms. Of course, the dynamics and technological affordances of a major social networking platform are different from those of a video game platform — the simple fact that users on many social platforms are empowered to upload their own files, where the structures of video games can be significantly more limiting, means that there is perhaps a higher immediate level of risk for users (and users performing moderator duties) to be exposed to extreme or illegal imagery from unsolicited 'dick pics' to brutal beheadings. That said, Koster (2017) notes that players can be inventive in finding ways to undermine or exploit systems for the purposes of abuse, harassment and/or garden-variety mischief, citing Julian Dibbell's example of sexual assault in text-based Multi-User Dungeons (Dibbell, 1993), users uploading and programming swarms of malicious penises in *Second Life*, and the practice of dropping items on the ground in *Ultima Online* to spell out banned words or to crudely draw penises.

These sorts of images and files are already commonplace on some platforms, although platform operators have often out-sourced moderation to countries where moderators can be hired and trained *en masse*, and at relatively low labour costs. Adrian Chen's (2014) attempts to unearth and investigate the content moderation industry have produced some of our only insights thus far into how these platforms manage content, and the into effects of this responsibility on their employees. Chen's investigations paint a picture of out-sourced distress and discontent — as a former Youtube moderator exclaimed to him (Ibid.): "Everybody hits the wall, generally between three and five months, [...] You just think, 'Holy shit, what am I spending

my day doing? This is awful.” Chen’s interviews and investigations of content moderation organisations (some of which are seemingly contracted or subcontracted to work for major platforms like Facebook, Twitter) have revealed that the role of content moderator is gruelling — and subject to high labour turnover, as moderators grow to resent the role and the content they’re tasked with policing.

A radical change to this role could see profound change to the social contract and social dynamics of major social platforms. Put simply — what happens when the responsibility for a good, safe community is held by the users in concert with the platform, instead of the platform alone? Riot’s example of hybrid model for participatory governance is instructive. As Kou and Nardi (2014) explained, not only did the ‘Tribunal’ model allow for players to become arbiters and — it also re-affirmed and communicated the social contract of the gaming platform, as users became invested in working to reduce abuse and hostility in their gaming communities.

Alternative Business Models

At what could be seen as the height of Twitter’s political impact, during charged political events like “Occupy” and the “Arab Spring”, Twitter briefly seemed to position itself as an institution that could be a bulwark against the state, protecting the privacy and the free expression of its users by refusing to bow to demands from authoritarian governments, and attempting to refuse and challenge a court-ordered subpoena to release user information about a protestor in the Occupy movement in the US (Ax, 2012). That being said, principled opposition and objection to governmental pressure may only go so far in a commercial, publicly-listed organisation. Indeed,

Twitter's strength in the face of authoritarian regimes has weakened somewhat, as it has begun to cooperate with requests for censorship from the government of Turkey under Recep Erdogan (Tufekci, 2017 243). Pressures to maintain or increase profitability may see principles like these hollowed out or eschewed entirely in favour of arrangements that make operating a business in certain markets more convenient, or less risky. To find an example of the pressures of capital altering the political machinations of an online platform, we need look no further than Tumblr. Whilst Tumblr had been a vocal opponent of proposals for legislation that sought to weaken 'Net Neutrality', its presence in these legislative issues has lately been muted to the point of absence — an absence that has been noted in industry press as a direct result of the influence of Verizon, the telecommunications company that recently acquired Tumblr's parent company, Yahoo! (Tiffany, 2017).

Given the stresses and complications associated with operating commercial social platforms (especially those with significant venture capitalist investments to justify), we can consider alternative business models for social platforms. Directly state-owned and state-operated platforms would be undesirable, given the possibilities (or perhaps, proclivities) of a government with authoritarian intentions to abuse a state-operated platform as a vehicle for surveillance. That said, independent state-owned, state-funded broadcasters might be illustrative of a model for state platform operation (at a figurative 'arm's length' from the government), building on the notion of a 'Civic Commons in Cyberspace' proposed by Blumler and Coleman (2001 16). Given the examples of the Australian Broadcasting Corporation, the British Broadcasting Corporation, the Public Broadcasting Service and dozens of other public broadcasters, a nation that saw the value in creating and operating a platformed online space for its

citizens could conceivably do so in an independent manner — although, of course, this takes an idealistic view of public service broadcasters, which are often formally independent but subject to much more subtle (often fiscal) levers of influence.. Alternatively, a handful of open-source platforms like Diaspora and Mastodon show a possibility raised by federated and open-source platform software — the ability to distribute governance by distributing infrastructure and operations. This approach is not without issue, as power can still concentrate in the hands of the individuals or organisations with the resources to host and operate popular federated instances, and different instances may be governed in dramatically different ways³¹. Moreover, network effects mean that the challenge of financial security, or unseating other dominant platforms in user’s minds will be particularly difficult for platform newcomers. Co-operative platforms have been proposed as alternatives to ‘Sharing Economy’ platforms like Airbnb and Uber (Scholz, 2014) and a solution that could allow users to more fairly and openly negotiate labour issues. Issues of user labour are particularly prominent in ‘Sharing Economy’ platforms — but are frankly, still at play in through unremunerated uses of emotive, affective and traditional labour in more overtly social platforms, and video gaming platforms (Joseph & Williams, 2015). Similarly, we may yet see the rise of not-for-profit platforms, operated by a trust, or organisation with an ambit that seeks to provide and improve on the platform itself. Although there are not yet substantial examples of this style of operation for a social networking platform, both the ‘Ghost’ website publishing platform and The

³¹ Mastodon provides a clear example of this, as different Mastodon instances have different rules in place. At the time of writing, instances may set different rules around norms relating to nudity (if it is allowed on the platform with/without a NSFW tag, or if it is allowed at all), to hate speech, to sexism, to racism, to advertisers, to media spoilers, and to abuse and harassment.

Guardian's operation under the Scott Media Trust offer notable examples of this style of non-profit development and operation.

Transparent Governance

Stepping back from the more drastic, or perhaps radical, changes suggested above — one of the simplest policy changes that could be managed by platforms in order to more meaningfully interact with users is a reform towards policies of transparency in governance. Indeed, Blumler and Coleman (2001 16-18) considered Transparency a key issue for a proposed governing organisation in their consideration of a Civic Commons. Transparency can be something of a two-sided virtue for social platforms. Whilst opening the deliberations and governance decisions of the platform — particularly in regard to moderation— to wider public scrutiny may help contextualise the decisions that platforms make, the scrutiny may also serve to highlight weaknesses in their governing processes. Could deliberately belligerent users and troublemakers be conceivably more dangerous to other users once armed with a more thorough knowledge of the platform and its governance capabilities? Or could they stage a malicious takeover — gaming representative processes to gain disproportionate power or influence? In the Twitter Case Study of this thesis, I highlighted an example whereby belligerent users discovered a weakness in Twitter's moderation — chiefly, that they could skirt blocks, and use niche targeting systems to direct abusive messages at their chosen targets. These users found these weaknesses through experimentation, not through formal disclosure — and highlighted the weaknesses by exploiting them. It may well be the case that transparency offers more risks because of this — but this should also spur platforms to taking a proactive and thorough approach to designing systems with user safety in mind — and a more

quickly reactive approach to policing exploitations. Of course, this sort of work is research-intensive and labour-intensive. It will not necessarily be immediately more cost-effective for platforms to be transparent with their users. It may not ever be cost-effective for platforms — but it is a principle of good governance. Koster (2017) argues that this sort of governance is essential to the hosting of online communities, and that it is negligent of online communities to operate without consideration of their responsibilities in governing users. As Koster argues, “If you host an online community, you are on the hook. If you don’t have the wherewithal to be on the hook, don’t host an online community.”

Platforms will have to start somewhere on the road to transparency. Perhaps one of the simplest avenues for transparency practices is the opening of their internal systems to external audit by independent or government-appointed regulators. Morozov (2013 185-186) suggests learning from the financial sector in this regard, noting that “algorithmic auditors” should be able to attest to the workings of platforms in a similar manner to the financial auditors appointed to inspect major financial organisations. Of course, the existing system of financial auditing and regulation leaves much to be desired, but the existence of these systems indicates the willingness and importance of regulation and observation for financial industries and may provide an opportunity for reform for other industries.

Close(r) to the Machine — Who audits the neural network?

The prospect of auditing these platforms and their component systems (from moderation processes to algorithms and designs) is complicated significantly by the advent of cognitive computing systems, deep learning and neural network-based

computing systems. Some of these systems represent new and fundamentally different methods of computing. The brief explanation is that where previous computing systems have been driven by the instruction sets (or ‘algorithms’) of software engineers, machine learning systems instead use a vast array of associative pattern systems to build understanding and associations based on data samples rather than instructions. Deep learning systems in particular embed these processes into simulated neurons, attempting to create and simulate brain-like reasoning processes. The problem these systems pose, is that they cannot yet be audited. Whilst the deductions these systems make may be perfect, correct or entirely plausible, the cognitive, and deep learning systems that are being used today are as yet unable to justify their reasoning (Knight, 2017). Whilst there may be developments in the future as to auditing the processes and reasoning of cognitive computing and learning systems, the reality of the present is that these systems are currently inscrutable to human oversight, and should likely be avoided in governance contexts, and especially in contexts where their use or deployment might have an adverse effect on the rights of a human user.

Chapter 8: Conclusions

This thesis has explored the complicated, detailed histories of three key conflicts on popular social media platforms. Through analysis and consideration of these three case studies, this work answers the three research questions this project was created to address:

1. What strategies do platforms use to influence and control user behaviours on their platform and the development of their governance?
2. What tactics do users deploy when in political conflict with a social media platform?
3. How do these strategies and tactics influence the interplay between users and platforms in political conflict?

As noted in the earlier discussion chapter, platforms use a range of strategies deployed either through organisational policies or the design of the platform. Each of these work by instrumentalising the platform's power in different ways — by creating and policing a *de facto* set of laws for the space for the former, or by shaping and influencing user behaviour through the design, development and deployment of prescriptive technologies for the latter.

Users, similarly wield power in disparate varied ways — like the notionally extrinsically-motivated usage of *organisational* tactics, which borrow significantly from historical and contemporary political movements with tactics that seek to raise

awareness, and articulate political cases for change, or the notionally intrinsically-motivated usage of *circumventive* tactics, in which users deploy their first-hand knowledge and experience of everyday life on the platform to subvert the platform's designed or intended uses. Of course, conflict in digital spaces is not simply a matter of the user's David standing against the platform's Goliath — users can similarly deploy tactics for ill, especially in belligerent conflict with other users. The interplay of these strategies and tactics is perhaps the most telling about prospects for reform and development in these existing spaces. Whilst users managed to eke out some small victories and concessions, platforms largely continued with little change to their core practices or principles. Considering this, I'd like to contextualise each of the case studies with some addenda, noting the issues that arose in relation to these conflicts after the end of each case study's research period, and illustrating what changed — and what did not.

What happened next? Addenda to the Case Studies

Facebook

The Continuing Conflict

One issue to consider, given the scope of this research, is what reform and change really look like in the context of the platform, and the broader media narratives about the platform. Whilst Facebook has initially received quite a significant amount of positive press and acclaim for the changes to their policies, the interceding time between the conclusion of these case studies and the writing of this thesis has seen a continuation of most of the conflicts studied. For Facebook, this has largely manifested

in continued, occasional reports of mothers who have had their breastfeeding photographs removed, and women who have had photographs of mastectomy scars removed. A growth in the popularity of ‘birth’ photography — that is, photographs taken during the birthing process, often with particular focus and detail on the delivery of a child — has also seen these images become a focal point of this censorship-based conflict over time. Whilst some birthing photos might more clearly include nudity in their subject matter, for many of the users who participate in this practice, the contextualisation of that nudity is inherently important — as the nudity is not sexualised, rather it is a representation of motherhood, birth and an attempt at an honest, personal account of the reproductive process. This indicates that for some users the diverse, and sometimes changing, social norms relating to nudity are in a contrast with Facebook’s comparatively conservative norms of nudity.

Facebook’s norms (or perhaps, the inelegant enforcement of Facebook’s policies) can particularly chafe against western social norms around art, documentary, ‘important’ and ‘worthy’ contextualisations of violence and/or nudity. In particular, late 2016 saw Facebook moderators censor the Norwegian *Aftenposten* Newspaper for a post that included Nick Ut’s photograph of Phan Thị Kim Phúc, and other children in the Vietnam war (Hansen, 2016). The Pulitzer award-winning photograph is informally known as the “Napalm girl” photo, as it depicts civilians running down a road near Trang Bang after a Napalm attack. In this photo, Kim Phúc (then, aged 9) is seen running towards the camera, badly burned and naked, having torn her burning clothes away. Facebook saw strong criticism from members and ministers of the *Høyre* Conservative Party, including Prime Minister Erna Solberg, over this and similar instances of censorship of the very same photo (Nordby, 2016).

Facebook, eventually conceding the significance of the image, noted in a statement (in Wong, 2016):

“While we recognize that this photo is iconic, it’s difficult to create a distinction between allowing a photograph of a nude child in one instance and not others. We try to find the right balance between enabling people to express themselves while maintaining a safe and respectful experience for our global community. Our solutions won’t always be perfect, but we will continue to try to improve our policies and the ways in which we apply them.”

This statement is largely devoid of the technocratic pose displayed in some of Facebook’s other communications — and does not minimise Facebook’s role as a moderator and intermediary. That said, Facebook’s characterisation of creating distinctions between instances of child nudity as ‘difficult’ underscores a temptation for the platform to eschew complexity and contextuality, and to favour simple, and easily applicable restrictions. This signals that future platform-user conflicts are likely to continue to occur, as users continue to engage with the platform and demand more careful and considered practices of moderation — that weigh user freedoms more carefully and weigh public interest and significant more critically.

‘Fake News’

Of course, conflicts surrounding Facebook’s role as a moderator and censor do not end with the issue of nudity. The 2016 election of Donald Trump as President of the United States of America has seen a rise in media criticism, particularly with reference to what is being described as “fake news”, a loosely defined concept that seems largely to focus on demonstrably and wilfully false and misleading information

being presented as news on social internet platforms. As the “fake news” phenomenon has grown in notoriety, it has begun to serve something of a scapegoat for the failures of the Democratic Party and Democratic Nominee, and for traditional news organisations struggling to understand why some voters would ignore journalistic evidence from mainstream news sources. Given significant pressure from political actors and the mainstream media, Facebook and its leadership have been recently become more vocal in their opposition to ‘fake news’ and in their desire to address the growth and delivery of ‘fake news’ on the Facebook platform by giving users more advanced fact-finding tools (Swisher & Wagner, 2017). Facebook intends to do this in two parts; firstly, by adding options that allow users to flag content as “fake” or “misleading” to the flagging and user-reporting mechanisms that already exists for users, and secondly, by partnering with non-profit fact-checking organisations like ‘Correctiv’ to label embedded links and off-platform content with labels distinguishing them as “disputed”, and algorithmically de-prioritising them in the individualised news feed that Facebook presents to each user. These approaches are a demonstrably light-touch approach to moderation, as Facebook does not directly censor or intervene in the posting of “fake news”, instead opting for an indirect approach. Without internal focus it’s difficult to know the motivations for this approach, although we might suspect that given the overt political nature of many “fake news” issues, Facebook and its operators are likely seeking to minimise the potential for them to be seen as a partisan intermediary in the digital life of their user base.

Amateur Custodians — Users demanding better governance

These conflicts can also be seen in the way users have assumed responsibility as proactive, amateur custodians of the digital spaces they inhabit. A particularly

interesting example of this is the work done by users in policing and reporting users who were flaunting or disregarding Facebook policies in regards to the sale of firearms (Schatz & Sammon, 2016).

Although Facebook formally banned the sale of firearms on the platform in January 2016, reports indicate that for a lot of users, this ban was quite easily circumvented by attempting to avoid overt mentions of firearm sales. To that end, groups like “Texas Pew Pew Pew Show & Tell”, sprouted up for “anything that goes pew pew pew,” admonishing users to “be smart” and “to take it to PM,” (private messages). Users reported that other users were circumventing Facebook’s policy with a wink, by taking photos of their firearms next to other, innocuous products that they’d list for sale. For example, a user photographed a .308 rifle next to a bottle of baby powder, advertising it for over \$800. Another took a photograph of an .223 scoped rifle next to a can of Hawaiian Punch drink, detailing calibre and gun options in the text of the post (Mele, 2016).

The users policing Facebook’s policy reported that they had seen over 1,000 pages and sales listings removed but expressed frustration at Facebook’s poor handling of the ban, asking why more significant tools like the image recognition tools Facebook reportedly uses to identify child abuse material are not also deployed in the enforcement of this policy. Monteiro (in Mele, 2016), a web designer involved in the user-led policing of this policy argues that “What all this boils down to is if Facebook did not want guns on the site, the guns would be gone tomorrow. They have the technology to do this.”

The issues raised by these users highlight the power and roles of platform policy documents. For these user-custodians, their frustration is in part, aimed at Facebook for not upholding the firearms sale policy well enough. This allows us to ask — what

is the point of a policy not carefully or diligently enforced? To what end are Facebook's policies documents of governance, and to what end are these policies simply public relations exercises?

Google+

The (Slow) Decline of Google+

In comparison to both Facebook and Twitter, Google+ has languished since its release in 2011. Whilst it has failed to achieve the comparative popular use and successes of Facebook, or the social and cultural relevance of Twitter, Google+ is by no means a 'dead' or 'abandoned' platform, as it seems to have found some relatively small niches of dedicated users — Google employees, YouTube users, bread makers, board game enthusiasts, and users of Google+'s relatively well-received suite of photography features (Bell, 2017).

Google+ still tends to receive some infrequent media attention, typically at the release or promise of new features (Dickey, 2017; and Wilson, 2017), however, Google+'s lack of significant growth or social import is likely compounded by the exit of key management figures in recent years, which seems to have left the platform comparatively rudderless. As noted in the Google+ Case Study, the recent hiring of 4chan founder Christopher '*moot*' Poole drew some industry speculation as to the potential for further development of the Google+ platform. As of the time of writing, Poole's employment at Google has not heralded radical changes to Google+ systems, principles or policies, but it may yet be too soon to tell what influence, if any, he has over the development of the Google+ platform. Whilst improvements in the last year

have been subtle, they have included features designed to draw attention to other communities and content on Google+, presumably to foster stronger social connections for the users already on the platform.

Facebook: A New 'Real Names' Policy and Struggles for Identity

Remarkably, Facebook has in recent years become embroiled in a similar conflict with its users to that outlined in the earlier Google+ Case study, in regards the names that they are allowed to use on the platform. During this research project, Facebook faced some criticism for its policy mandating the use of 'real names' on the platform. Although Facebook's development and growth as a platform significantly predates Google+'s, Facebook seemed to significantly avoid criticism of this policy in its early years — perhaps due to the platform's early focus as a social network that facilitated social connections with friends, and work, university and school colleagues. Nonetheless, Facebook's policies seemed to be treated as that platform's status quo for quite some time, perhaps because dissatisfied users at the time could look towards competing platforms on the horizon, including Google+, and the then burgeoning development of a handful of other social networking spaces like the open-source 'Diaspora', and the more intimately focused 'Path'.

Nonetheless, it seems as though Facebook's operators were either unaware or indifferent to the conflict surrounding Google's policies, as they eventually drew similar criticism and fostered a similar conflict in response to their policy and its enforcement on the Facebook platform. Many of the key issues at play in the Google+ conflict were similar points of frustration for users in conflict with Facebook. These included significant frustrations about the policing of pseudonyms and stage names, especially for lesbian, gay, bisexual, transgender, intersex and queer-identifying users

and drag performers whose self-identified names may not necessarily correspond with their birth certificate or any forms of government-sanctioned identification.

Facebook has similarly been criticised for implementing technological affordances that have been primarily designed based on the expectations of developers working with a basic understanding of names and naming norms in western nations like the United States of America and the United Kingdom. For this reason, the platform has been criticised for not accommodating names with multiple words or spaces. This includes the surnames of Native Americans, like 'Dana Lone Hill', 'Lance Brown Eyes', and 'Robin Kills The Enemy', all of whom have reported frustration with the platform's policy and policing (Holpuch, 2015).

Similarly, for people in the indigenous Chamoru culture of the Marianas islands, naming conventions have often resulted in children taking their mother's maiden name as a middle name, and so multi-word surnames like "De La Cruz" (which can be somewhat common, given the Marianas history as a colony of Spain) are being used as middle names. Facebook's technological affordances have precluded middle names from including 'space' characters, and as such these users are often required to input their names without spaces or in an abbreviated form — which then puts them in contravention of Facebook's policy, which requires them to use their legal name as it is presented on legal documents. Facebook also offers no support for users with mononyms, or users with Tamil names consisting of a mononym followed by maternal and/or paternal initials (moulee, 2009).

Facebook's policies relating to this area of concern have changed slightly during the course of this research project, but similarly to findings of my case study analysing Facebook's policy response to criticism of its policies regarding breastfeeding photographs, the overall policy and its effect seems largely unchanged, with only some minor concessions made. In particular, Facebook's response has been most notable to criticisms of the policy and its relation to pseudonyms and stage names, as made by the LGBT and ally communities. These communities drew mainstream media coverage and attention to the policy in 2014, after which Facebook made some concessions as to the potential for the use of pseudonyms on the platform under certain conditions, but largely kept its policy the same.

In a press release in 2015 (Osofsky & Gage, 2015), Facebook announced that it would be providing a compromise on its policy for people with "special circumstances", or a "unique situation" that could allow users the ability to use a pseudonym subject to a verification of their "true" identity with the platform. At the conclusion of writing this thesis, there is no evidence in Facebook's policy documents that this measure has even been implemented.

Twitter

Unfortunately, the Austengate and Gamergate conflicts seem to have served as a blueprint for some, as reports of abuse and harassment on online social platforms — especially Twitter— continued have continued even after the apparent conclusions of these conflicts.

Reactionaries, Abuse and the 'Alt Right'

With the benefit of time, it has become reasonably appropriate to look from Gamergate to the reactionary media and online campaigns in during the 2016 Presidential Election of the United States of America, and the ensuing presidency of Donald Trump. This is of course, not to say that President Trump is a Gamergate proponent or advocate, but rather to highlight some continuities between the communities that formed in and around Gamergate and the communities supporting President Trump both during and after his candidacy for President. Observing the presidential campaign, it has been peculiar to see a number of the key actors and thought-leaders in the Gamergate conflict notionally associated with the Trump campaign through employment, advocacy and other means. A key example of this is the Trump campaign and presidency's associations with 'alt-right' evangelists like Mike Cernovich, and the political website Breitbart, and the columnist formerly associated with that website, Milo Yiannopoulos. Cross (2017) correctly notes that a number of the political actors associated with Trump, including Yiannopoulos and Cernovich, had seemingly cut their teeth in online reactionary conflicts, including Gamergate — indeed, some of these actors have essentially built their personal brands during these conflicts with strident criticisms of "leftists", "feminists" and seemingly anyone who dared to question the status quo of the video games industry.

The 'alt-right' movement is complex and difficult to fully categorise, as its advocates seem to espouse a range of views and beliefs with varying degrees of coherence — ranging from the white nationalism and ethno-nationalism of overt anti-semites, to the economic nationalism espoused by Breitbart and the *Front National* political party, to the esotericism espoused by writers Savitri Devi and Vox Day, or

the more reactionary accelerationism and libertarianism espoused by Nick Land and Milo Yiannopoulos (Ambedkar, 2017). Whilst the philosophical and rhetorical underpinnings of the alt-right movement have roots in the writings and espoused philosophies of philosopher Nick Land, and computer scientist Curtis Yarvin, its movement to the fringes of the political mainstream might have less to do with the underlying philosophy outright, and more to do with the social groups that have formed in notionally counter-cultural, reactionary online spaces. Part and parcel of the development of the alt-right into a more mainstream subculture is the use and deployment of irony — especially as a technique for deflecting criticism and introspection.

To this end, the overt belligerence of people who identify with or who are identified as 'alt-right' is often presented as in jest, or "not serious", which tends to serve to minimise the perception of their malfeasance or behaviour. To wit, practices of anti-semitism, racism, and sexism are often characterised as "just trolling", or "baiting", a tactic that also functions to portray their critics as poorly informed and hyper-sensitive. The intent of the people using tactics of belligerence aside, these tactics of irony and semantic mischief are remarkably similar to those discussed by Sartre in his analysis of anti-semites (1965 17-22):

They know that their remarks are frivolous, open to challenge. But they are amusing themselves, for it is their adversary who is obliged to use words responsibly, since he believes in words. The anti-Semites have the right to play. They even like to play with discourse for, by giving ridiculous reasons, they discredit the seriousness of their interlocutors. They delight in acting in

bad faith, since they seek not to persuade by sound argument but to intimidate and disconcert.

This kind of frivolous debate was rife in Gamergate and continues to be a theme amongst the activities (both online and offline) of people who associate with the alt-right movement. To wit, the interrelation of both overt and covert bigotry with memes (like Matt Furie’s Pepe the Frog), gaming imagery (using brand and genre conventions, such as “Simulator” titles), popular cultural iconography (like anime characters, or Marvel’s Captain America and Red Skull characters), and religious iconography (like Hindu deity Kali Yuga, and Ancient Egyptian deity Kek³²), and modern re-creations and re-interpretations of historical bigotry (like the ‘Happy Merchant’ caricature), to present statements and arguments that are ludicrous at face value. By doing this, the deflection tactic of irony is used to avoid criticism and focus on their accuser: “Can’t you take a joke?”, “I’m just trolling.” Hess (2016) argues that the sardonic nature of these communications is moot, as the effect and influence of this communication tactic is largely the same, regardless of the earnestness, seriousness, or lack thereof of its proponent. She emphasises “Who cares if the Klansman is smirking beneath his hood?”

³² The ancient Egyptian deity *Kek* (a deification of primordial darkness — also transliterated as *Kuk*, *Keku*, and *Kekui*), is frequently spelled identically to “kek” — a fictional Orcish translation of the “lol” acronym, as it appears in *World of Warcraft*. Because the game features two opposing and notionally hostile factions, inter-faction communication is hindered by the presentation of a non-common language.

To wit, when players of the Horde or Alliance players type “lol” near each other, their speech bubbles appear to the opposing faction as “kek” or “bur” respectively. ‘Kek’ has been a synonymic term for ‘lol’, especially in internet subcultures, for years — likely since as early as *World of Warcraft*’s 2004 release. It is likely an amusing coincidence that users associated with online belligerence have found this association with an Ancient Egyptian deity.

Agenda-setting: Twitter, the Pugilistic Pulpit

Whilst this research project engaged more fully with the tactics used by those targeted by abuse and harassment on Twitter, the tactics of abuse and harassment are worthy of further investigation.

In particular, the Gamergate conflict demonstrated the efficacy of Twitter as a platform for agenda-setting, where the issues disputed, raised, and sometimes fabricated for discussion on Twitter were often seen and reported on by journalists and opinion writers for the industry and mainstream press. Whilst this may be partially due to the popularity of Twitter amongst journalists and professional writers, it may also be indicative of organised attempts to influence discourse. After all, the Gamergate conflict was not isolated to Twitter, and was a significantly multi-platform conflict that saw staging grounds in “offline” venues like conventions and meet-ups, in social platforms like Twitter, in blogs, in forums like Reddit, and highly anonymised imageboards like 4chan and 8chan. To that end, campaign and communications strategies could be discussed and shaped away from the eyes of those who would eventually see them on another platform. For some, this platformed distance facilitated enough of a distance to the overtly belligerent users to allow a belief that the campaign was not about the harassment that was so clearly and frequently associated with it (Danskin, 2015).

Of course, researchers tackling this problem should take particular care to ensure that their research systematically analyses these tactics, and the structures that enable them, without providing an *de facto* instruction manual that allows readers to potentially continue their perpetration.

The dominant narrative

From Gamergate onwards, the popular narrative about Twitter in both industry and popular press has focused on the failures of Twitter as a platform and as a social space to deal with abusive and belligerent behaviour on the platform. Twitter has in recent months introduced a bevy new features and filters to combat harassment (Burgess, 2017) —many of which were similar in design and purpose to the user-created solutions— but still struggled to fully engage with the desires of users in this area. The rise of the ‘Alt-Right’, and their continued usage of Twitter is a sticking point for users. Given the use of the platform by organisations and figures like the American Nazi Party and former Imperial Wizard of the Ku Klux Klan David Duke, the overt anti-semitism and white nationalism espoused by the people who identify or associate with the Alt Right, and the Nazi imagery used in display pictures and avatars of many anonymous troll accounts, many users and writers have come to term this issues as Twitter’s “Nazi”³³ problem. This is particularly exemplified in media coverage, where a news story about a federated, decentralised and open-source competitor to Twitter opened with the headline “Mastodon Is Like Twitter Without Nazis, So Why Are We Not Using It?” (Jeong, 2017), and an article commenting on a

³³ Given the discussion of anti-semitism and Nazis, references to Godwin’s law are almost inevitable in discussions about Twitter’s problems with the policing of abusive and belligerent users. Godwin’s law an adage created by lawyer Mike Godwin, who asserted “As an online discussion grows longer, the probability of a comparison involving Hitler approaches 1.”

Godwin argues that whilst the ‘law’ is phrased like a law of nature or mathematics, he had considered it as a pedagogical tool — hoping that people who were inclined to glibly compare other to Hitler would be more reflective about the Holocaust.

In 2015, he urged that comparisons to Hitler were not always inappropriate (Godwin, 2015). As former Electronic Frontiers Foundation office director Gerard Van der Leun argued with a corollary to Godwin’s Law (Wired Staff, 1994): “As global connectivity improves, the probability of actual Nazis being on the web approaches one.”

design change in light of Twitter’s corporate priorities leads with the headline “Twitter’s Nazi problem is still out of control” (Molitch-Hou, 2017).

The Limitations of this Research, and the Timeliness of this Investigation

As a project using case study methodology, there are some clear limitations to this research — case studies represent an opportunity to thoroughly analyse events or phenomena in the pursuit of detail and understanding. The cases I have analysed in this thesis show what has happened in platformed spaces, histories of conflict and governance in these platforms. However, as I have illustrated in this chapter, these conflicts are not strictly historical in scope — they have largely continued on, through the case boundaries established for this research and will likely continue for the immediately foreseeable future. These cases should serve, then, as detailed snapshots of platformed conflicts relating to user expression, identity and safety on Facebook, Google+ and Twitter. The observations from this research may not necessarily map to every future conflict on these platforms, or to similar conflicts on other platforms, but should provide opportunities for further research and investigation.

These sorts of conflicts are relatively prevalent on online platformed spaces, as tensions amongst users, and tensions between users and platform operators work to negotiate power in these spaces. Similar issues to those observed in this thesis have occurred on the Reddit platform, where users and platformed operators clashed over the boundaries of offensive, abusive and hateful speech acts on the platform, eventually leading to the controversial ouster of Reddit’s then-CEO Ellen Pao, after a grassroots user campaign against her that accused her of censorious overreach, often

pairing these accusations with racist or misogynistic slurs and imagery (West, 2015). Other platforms, like Tumblr, have grappled with user expectations regarding sexuality and free expression on the platform, as the platform's moves towards monetisation and profitability under parent company Yahoo! has led it to design affordances that worked to hide 'NSFW' content from users (C. Warren, 2013), and in doing so blocked users from accessing content tagged with LGBT hashtags like #gay on mobile devices (Tumblr Staff, 2013). These conflicts suggest that issues of expression, identity, and safety are particularly important to users, and that these areas are thematic fault-lines on which platform conflicts are likely form in the future.

The continued growth of major platforms like Facebook, and the market conditions and desires spurring the development of alternative platforms points to a need the relevance of this field of inquiry moving forward. Clearly, the politics and governance of social platforms is a timely issue, which will require further research to more clearly understand the interplay of power between users and platform operators.

Shocks & Exceptions

The pattern of shock and exception raised earlier in this thesis has become apparent in each of the case studies observed for this research project. Longer conflicts tend to comprise a multitude of shocks, followed by eventual exceptions. This is particularly apparent in the Facebook case study, where the tone of Facebook's responses changed significantly over the course of several years of user complaint, frustration and protest. Similarly, the Twitter case study saw dozens of shocks congeal

into an ongoing conflict and animosity towards Twitter's lack of meaningful exceptions or reforms.

Indeed, each of the case studies are essentially focused around platform shocks and observing platform exceptions and user responses. Ananny and Gillespie note exceptions as a particularly poor end-point for users in these "shock" conflicts, noting that platform operators could essentially "make an exception, something that breaks from business-as-usual under specific, extenuating circumstances". This is not dissimilar to the strategy of 'sufficing' that I identified platforms using earlier in this thesis, whereby platform operators attempt to continue on, "business as usual", following the path of least resistance for the platform's continued business operations.

As such, these shock and conflicts may well prove to be terrible opportunities for platform reform. As Ananny and Gillespie (2016, 20) note:

"Platforms can weather the public anger and introduce exceptions to their normal practices, in ways that defuse the public outcry and keep it from extending into a more involved consideration of ongoing harm to the public or regulatory intervention. Shocks become things that platforms might anticipate, plan for, and remember."

The lack of efficacy for reform should not discourage researchers from observing 'shocks' or 'conflicts' as part of research projects. As noted earlier in this thesis, conflicts present a valuable opportunity to observe platforms in-action, and under-stress, allowing researchers to identify structures and components of platform

operation and platform governance that may not necessarily be obvious to even the operators of a platform. This opportunity for research is similarly espoused by Ananny and Gillespie (Ibid.), who note “it can be a diagnostic that can help us, and regulators, better understand the implicit compact between platform and public, as well as its limits.”

The conflicts and shocks seen in this thesis have often re-tread historical tensions and previously seen conflicts. As highlighted in the literature review of this thesis, platform operators seem to operate without a strong understanding of the historical context in which they seek to exert governance. To that end many similar conflicts are fought over similar themes and issues. Indeed, the issues at the heart of the conflicts in this research—expression, identity and safety—are much broader than the conflicts engaged on these platforms. It is critical that future research in this area seeks to not only create a critical framework for understanding these conflicts, but also contributes to the historiographical work of chronicling these conflicts for future interrogation. This important work must not be left only to the platform operators, given not only their obvious conflict of interest, but also the lack of commercial pressures ensuring platform operators do this work in assisting or attempting at any form of preservation of the platforms. We have seen this is in the cavalier attitudes that platform owners had for the former contents of Geocities webpages (as discussed earlier in this thesis), and in the efforts that users took to try and preserve those pages. Similar efforts have been prompted recently by the news of the potential closure of the popular audio-sharing platform Soundcloud (Franzen, 2017), with users rushing to archive the platform ahead of a possible closure due to insolvency.

Whilst historical reference works to platforms, online communities and online conflicts could serve as helpful for platform operators, and developers in learning from and circumnavigating issues of yore, they could also be of significant use to the users in platformed conflicts. Historical records of platforms and the conflicts that occur on them are increasingly important as we continue to see both more usage and more turnover of intermediary internet platforms like Facebook, Twitter and Google+. As we see more both more usage and more turnover of social media, user generated content and other intermediary platforms, the creation of historical records of these platforms, the conflicts that shaped them, and the lived experiences of users becomes increasingly urgent. It is the task of researcher to create and preserve the memories and the lived experience of these digital spaces as artefacts and as historical records. Researchers must endeavour to create historical records of internet platforms now, and archaeological records of the platforms that have already come and gone. This can be done with reference to similar media archaeology (Parikka, 2012) research practices. Creating a wealth of historical resources for researchers, users and platforms to rely on may well save researchers, users and platform-operators from ‘re-inventing the wheel’ (with regards to policy or technology or social practices) with each new platform. Without historical practices and historical knowledge, the people who participate in these conflicts are currently left to their own devices — to create the strategies of governance or tactics of resistance and subversion as they go, without understanding the lessons already learned through prior experience. As Dash (in Hudson, 2015b) noted in a criticism of Twitter’s response to abuse and harassment, the problems at hand were not new or exclusive to Twitter — organisations in Silicon Valley had experience similar issues long before Twitter had even existed. Creating public records of platformed conflicts future researchers can assist in the creation of

theory and praxis for users in platformed spaces. This could be particularly critical in the formation of user-based or user-led movements campaigning for reform in platformed spaces. A user-led movement like that of a user's union might start to see users benefiting from the accrual of institutional knowledge, which may serve to even the disparities of power between users and platform operators. In time, further research in this area, or user-led organisations may yet investigate the potential of a rights-based framework for user and platform expectations — seeking to enshrine and safeguard certain expectations that users have of platforms, and vice versa. Researchers will need to consider the ethics implications of historical research that will likely still have relevance to living participants of either the research or the platforms observed.

Whilst detailed findings and supporting documentation can provide a wealth of knowledge for researchers and stakeholders, it is imperative that researchers take steps to avoid the instrumentalisation and weaponisation of research. For example, detailed discussion and interrogation of tactics of abuse could be accessed and used as instructional material by people looking to cause trouble on social platforms. In this thesis, I go into some detail about a tactic of abuse that exploits the weaknesses of governance in Twitter's self-service advertising platform. Similarly, the discussion of the counter-tactics deployed by users who are targeted by harassing users could be used to weaken the efficacy of these counter-tactics, or to devise new tactics of abuse. Researchers must engage critically with institutional and personal research ethics frameworks in reference to social media users, and to conflicts on social media. This is of particular importance in issues relating to user safety and the protection of user identity. I highlighted in the methodology chapter of this thesis that this can be a complex and frustrating problem for researchers. Indeed, attempts to instrumentalise

research, and use it as an opportunity to find targets for harassment was even observed during the course of the ‘Gamergate’ conflict. Of course, many of the risks present in research — for example, the re-identification of anonymous participants, or the weaponisation of research data — are not necessarily new issues for researchers, they are issues that have become exacerbated by the technological affordances available to researchers and universities. Open-Access journals and ePrints databases only make easier what could just as easily have been accomplished by walking into a University Library to find Closed-Access articles or physical printings of these.

Although these risks are significant, the growth of existing platforms, the rise of new platforms, and the churn of unsuccessful platforms shows that it is imperative that independent researchers work to provide a body of research that can discuss the roles, powers, influences, and lived experiences of platforms and users as a civic resource. This can provide resources for evidence-based civic discussion of platforms and users and provide a groundwork for civil bodies and/or independent regulators. Beyond the public good associated with these goals, further research will also help to contextualise and understand trends and trajectories in platforms, their power, governance strategies, and the political influence and implications of platforms in everyday life.

Bibliography

- /pol/ - Politically Incorrect. (2015). /pol/ - Politically Incorrect. Retrieved May 21, 2015, from <http://archive.is/7JhXR>
- Acker, M. (2009). Breast is Best...But Not Everywhere: Ambivalent Sexism and Attitudes Toward Private and Public Breastfeeding. *Sex Roles*, 61(7-8), 476–490. <http://doi.org/10.1007/s11199-009-9655-z>
- Alder, K. (1999). *Engineering the Revolution: Arms & Enlightenment in France, 1763-1815* (2nd ed.). Princeton, NJ: Princeton University Press.
- Alexander, L. (2014). Gamasutra - 'Gamers' don't have to be your audience. 'Gamers' are over. Retrieved August 28, 2014, from http://www.gamasutra.com/view/news/224400/Gamers_dont_have_to_be_your_audience_Gamers_are_over.php
- Alexandra, H. (2016). Star Wars Galaxies Is Dead, But These People Are Keeping It Alive. Retrieved September 15, 2016, from <https://www.kotaku.com.au/2016/09/star-wars-galaxies-is-dead-but-these-people-are-keeping-it-alive/>
- Allan, R. (2017). Hard Questions: Hate Speech | Facebook Newsroom. Retrieved June 27, 2017, from <https://newsroom.fb.com/news/2017/06/hard-questions-hate-speech/>
- Allen, J. (2014). GamerGate and the New Misogyny. Retrieved February 12, 2015, from https://medium.com/@a_man_in_black/gamergate-and-the-new-misogyny-284bea6a8bb3
- Allen, J. (2015a). How crowdfunding helps haters profit from harassment. Retrieved January 15, 2015, from <http://boingboing.net/2015/01/14/how-crowdfunding-helps-haters.html>
- Allen, J. (2015b). The invasion boards that set out to ruin lives. Retrieved January 20, 2015, from <http://boingboing.net/2015/01/19/invasion-boards-set-out-to-ruin.html>
- Allen, J. (2015c). A beginner's guide to the Redpill Right. Retrieved January 29, 2015, from <http://boingboing.net/2015/01/28/a-beginners-guide-to-the-red.html>
- Allen, J. (2015d). How imageboard culture shaped Gamergate. Retrieved January 1, 2015, from <http://boingboing.net/2014/12/31/how-imageboard-culture-shaped.html>
- Allen, K. (2013). New £5 note replaces Elizabeth Fry with Sir Winston Churchill. Retrieved February 12, 2014, from <http://www.theguardian.com/business/2013/apr/26/winston-churchill-new-five-pound-note>
- Ambedkar, M. (2017). The Aesthetics of the Alt-Right – Post-Office Arts Journal. Retrieved February 13, 2017, from <http://baltimore-art.com/2017/02/11/the-aesthetics-of-the-alt-right/>

- American Civil Liberties Union. (1994). Hate Speech on Campus. Retrieved December 4, 2013, from <https://www.aclu.org/free-speech/hate-speech-campus>
- Ananny, M., & Gillespie, T. (2016). Public Platforms: Beyond the Cycle of Shocks and Exceptions. Presented at The Internet Politics and Policy Conferences. [http://doi.org/10.1002/\(ISSN\)1944-2866](http://doi.org/10.1002/(ISSN)1944-2866)
- Anti-Defamation League. (2012). Hate Crime Laws — The ADL Approach. Retrieved December 1, 2013, from <http://www.adl.org/assets/pdf/combating-hate/Hate-Crimes-Law-The-ADL-Approach.pdf>
- Arango, T. (2011). Hot Social Networking Site Cools as Facebook Grows. Retrieved December 12, 2013, from http://www.nytimes.com/2011/01/12/technology/internet/12myspace.html?page_wanted=all&_r=0
- ArchiveTeam. (2009). GeoCities - Archiveteam. Retrieved December 9, 2013, from <http://www.archiveteam.org/index.php?title=GeoCities>
- Arthur, C. (2012). Facebook's nudity and violence guidelines are laid bare. Retrieved October 13, 2013, from <http://www.theguardian.com/technology/2012/feb/21/facebook-nudity-violence-censorship-guidelines>
- Ask Myspace. (2013). How do I get my Classic Myspace Blogs? - Ask Myspace. Retrieved December 9, 2013, from <https://www.askmyspace.com/t5/Guides/How-do-I-get-my-Classic-Myspace-Blogs/ba-p/74871>
- Aspan, M. (2008). On Facebook, leaving is hard to do. Retrieved December 9, 2013, from http://www.nytimes.com/2008/02/11/business/worldbusiness/11iht-11facebook.9919316.html?_r=0
- Auernheimer, A. (2014). What I Learned from My Time in Prison. Retrieved October 10, 2014, from <http://www.dailystormer.com/what-i-learned-from-my-time-in-prison/>
- Auernheimer, A. (2015a). Fun with Twitter's ad console. (with images, tweets) · weev. Retrieved September 30, 2015, from <https://storify.com/weev/fun-with-twitter-ads-day-1>
- Auernheimer, A. (2015b). So I'm banned from Twitter's ad console (with images, tweets) · weev. Retrieved September 30, 2015, from <https://storify.com/weev/twitter-ads-copypcat-roundup>
- Ax, J. (2012). Twitter gives Occupy protester's tweets to U.S. judge. Retrieved March 6, 2015, from <http://www.reuters.com/article/net-us-twitter-occupy-idUSBRE88D01S20120914>
- Baio, A. (2015). Stupid Tricks with Promoted Tweets — The Message. Retrieved February 4, 2015, from <https://medium.com/message/stupid-tricks-with-promoted-tweets-57325552109d>
- Baker, K. J. M. (2013). How to Shut Down Reddit's CreepShots Once and for All: Name Names. Retrieved October 15, 2013, from

<http://jezebel.com/5949379/naming-names-is-this-the-solution-to-combat-reddits-creepshots>

- Ball-Rokeach, S. J., & DeFleur, M. L. (1976). A Dependency Model of Mass-Media Effects. *Communication Research*, 3(1), 3–21.
<http://doi.org/10.1177/009365027600300101>
- Bank of England. (2015). Banknote Characters | Bank of England. Retrieved October 6, 2015, from
<http://www.bankofengland.co.uk/banknotes/Pages/characters/default.aspx>
- Bank of England. (2013). Moving to Polymer Banknotes | Bank of England. Retrieved February 12, 2014, from
<http://www.bankofengland.co.uk/banknotes/polymer/Pages/default.aspx>
- Banks, J. and S. Humphreys (2008). "The Labour of User Co-Creation: Emerging Social Network Markets?" *Convergence* 14(4): 401-418.
- Barnett, E. (2011). Twitter chief: We will protect our users from Government - Telegraph. *The Telegraph*.
- Bates, L. (2013). I Helped Start the #FBRape Campaign Because We Must Not Tolerate Facebook's Allowance of Horrifying Pictures Showing Violence Against Women. Retrieved May 31, 2013, from
<http://www.xojane.com/issues/laura-bates-fbrape>
- BBC News. (2009). Thai website to protect the king. Retrieved July 24, 2013, from
<http://news.bbc.co.uk/2/hi/asia-pacific/7871748.stm>
- Beardsley, E. (2013). French Twitter Lawsuit Pits Free Speech Against Hate Speech : NPR. Retrieved November 12, 2013, from
<http://www.npr.org/2013/01/22/169998834/french-twitter-lawsuit-pits-free-speech-against-hate-speech>
- Behrent, M. C. (2013). Foucault and Technology. *History and Technology*, 29(1), 54–104. <http://doi.org/10.1080/07341512.2013.780351>
- Bell, K. (2017). Google+ isn't dead and these are the people still using it the most. Retrieved January 22, 2017, from <http://mashable.com/2017/01/18/who-is-using-google-plus-anyway/>
- Bernstein, A. (2014). "Abuse and Harassment Diminish Free Speech" by Anita Bernstein. *Pace Law Review*, 35(1).
- Biddle, S. (2014a). Chart Brut: How the MS-Paint Graphics of Conspiracy Took Over the Web. Retrieved August 12, 2015, from <http://gawker.com/chart-brut-how-the-ms-paint-graphics-of-conspiracy-too-1651851261>
- Biddle, S. (2014b). iPad Hacker and "Troll" Weev Is Now a Straight-Up White Supremacist. Retrieved October 10, 2014, from <http://gawker.com/ipad-hacker-and-troll-weev-is-now-a-straight-up-white-1641763761>
- Blau, A. (2013). When Foucault says foc-all: Part 1 | BlauBlog. Retrieved December 6, 2013, from <http://adrianblau.wordpress.com/2013/06/07/when-foucault-says-eff-all-part-1/>

- Block, J. (2010). Breaking News, Analysis, Politics, Blogs, News Photos, Video, Tech Reviews - TIME.com. Retrieved October 13, 2013, from <http://content.time.com/time/health/article/0,8599,2032363,00.html>
- Blue, V. (2011). Google Plus Deleting Accounts En Masse: No Clear Answers | ZDNet. Retrieved November 24, 2014, from <http://www.zdnet.com/article/google-plus-deleting-accounts-en-masse-no-clear-answers/>
- Blue, V. (2013a). Anger mounts after Facebook's 'shadow profiles' leak in bug. Retrieved December 9, 2013, from <http://www.zdnet.com/anger-mounts-after-facebooks-shadow-profiles-leak-in-bug-7000017167/>
- Blue, V. (2013b). Firm: Facebook's shadow profiles are 'frightening' dossiers on everyone | ZDNet. Retrieved December 9, 2013, from <http://www.zdnet.com/firm-facebooks-shadow-profiles-are-frightening-dossiers-on-everyone-7000017199/>
- Blue, V. (2014). Thanks for nothing, jerkface. Retrieved June 12, 2014, from <http://www.zdnet.com/article/thanks-for-nothing-jerkface/>
- Blumler, J. G., & Coleman, S. (2001). *Realising Democracy Online: A Civic Commons in Cyberspace*. (pp. 1–28). London: Institute for Public Policy Research.
- Bogost, I. (2011). *How To Do Things With Videogames* (Kindle eBook). Minneapolis: University of Minnesota Press.
- Bolter, J. D., & Grusin, R. (2000). *Remediation*. Cambridge, MA: MIT Press.
- Bornstein, J. (2015). White supremacists stole my identity to spew hatred on the Times of Israel. Retrieved May 6, 2015, from <http://www.theguardian.com/world/2015/may/05/identity-stolen-white-supremacists-times-of-israel>
- Bosker, B. (2013). Yahoo CEO Hints She'll Let Tumblr Keep Pumping Out Porn. Retrieved October 15, 2015, from http://www.huffingtonpost.com/2013/05/20/yahoo-tumblr-porn-marissa-mayer_n_3306382.html
- boyd, D. (2009). Implications of User Choice: The Cultural Logic of "MySpace or Facebook?" Retrieved December 12, 2013, from <http://dl.acm.org/citation.cfm?doid=1620693.1620701>
- boyd, D. (2012). The politics of "real names." *Communications of the ACM*, 55(8), 29–31. <http://doi.org/10.1145/2240236.2240247>
- boyd, D. (2006). Friendster lost steam. Is MySpace just a fad? Retrieved December 12, 2013, from <http://www.danah.org/papers/FriendsterMySpaceEssay.html>
- boyd, D. (2011). "Real Names" Policies Are an Abuse of Power. Retrieved May 12, 2013, from http://www.zephoria.org/thoughts/archives/2011/08/04/real-names.html?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+zephoria%2Fthoughts+%28apophenia%29

- Brogan, J. (2015). Twitter Allows Promoted Tweets With Racist Content. Retrieved May 7, 2015, from http://www.slate.com/blogs/future_tense/2015/05/06/andrew_auernheimer_aka_weev_targets_activists_with_white_supremacist_promoted.html
- Broomfield, M. (2016). How Men's Rights Activists Make Money Off White Dudes Who Want to Feel Oppressed | VICE | United States. Retrieved February 3, 2016, from <http://www.vice.com/read/roosh-mens-rights-activists-white-men-false-oppression>
- Brown, K. V. (2015). The Ultimate Troll: The Terrifying Allure of Gamergate Icon Milo Yiannopoulos. Retrieved October 29, 2015, from <http://fusion.net/story/220646/the-terrifying-allure-of-gamergate-icon-milo-yiannopoulos/>
- Bruns, A. (2008). *Blogs, Wikipedia, Second Life, and Beyond: From Production to Produsage* (1st ed.). New York: Peter Lang.
- Brustein, J. (2012). Anger at Changes on Instagram. Retrieved January 7, 2014, from <http://bits.blogs.nytimes.com/2012/12/18/anger-at-changes-on-instagram/?src=xps>
- Buni, C., & Chemaly, S. (2016). The secret rules of the internet. Retrieved April 14, 2016, from <http://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>
- Burgess, M. (2017). Twitter now lets you mute “eggs” and will tell you when action is taken against trolls. Retrieved March 2, 2017, from <http://www.wired.co.uk/article/twitter-tools-harassment>
- Cameron, D. (2013). @David_Cameron: “It's irresponsible of Facebook to post beheading videos, especially without a warning. They must explain their actions to worried parents..” Retrieved October 22, 2013, from https://twitter.com/David_Cameron/statuses/392573625750265857
- Carmody, T. (2011). Google+ Identity Crisis: What's at Stake With Real Names and Privacy. Retrieved September 28, 2015, from <https://www.wired.com/2011/07/google-plus-user-names/>
- Carter, C. (2014). Twitter troll jailed for “campaign of hatred” against Stella Creasy. Retrieved October 3, 2014, from <http://www.telegraph.co.uk/news/uknews/crime/11127808/Twitter-troll-jailed-for-campaign-of-hatred-against-Stella-Creasy.html>
- Centre for Independent Studies. (2011). Australia's Battle of the Think Tanks. Retrieved December 15, 2013, from
- Charmaz, K. (2014). *Constructing Grounded Theory* (2nd ed.). London: SAGE Publications.
- Chen, A. (2012a). Inside Facebook's Outsourced Anti-Porn and Gore Brigade, Where 'Camel Toes' are More Offensive Than 'Crushed Heads'. Retrieved March 12, 2013, from <http://gawker.com/5885714/inside-facebooks-outsourced-anti+porn-and-gore-brigade-where-camel-toes-are-more-offensive-than-crushed-heads>

- Chen, A. (2012b). Unmasking Reddit's Violentacrez, The Biggest Troll on the Web. Retrieved October 13, 2012, from <http://gawker.com/5950981/unmasking-reddits-violentacrez-the-biggest-troll-on-the-web>
- Chen, A. (2014). The Laborers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed. Retrieved November 30, 2014, from <https://www.wired.com/2014/10/content-moderation/>
- Cheney-Lippold, J. (2011). "A new algorithmic identity. Soft biopolitics and the modulation of control." *Theory, Culture & Society* 28(6): 164-181.
- Chircu, A. M., & Kauffman, R. J. (1999). Strategies for Internet Middlemen in the Intermediation/Disintermediation/Reintermediation Cycle. *Electronic Markets*, 9(1-2), 109–117. <http://doi.org/10.1080/101967899359337>
- Chituc, V. (2015). GamerGate: A Culture War for People Who Don't Play Video Games. Retrieved October 17, 2015, from <http://www.newrepublic.com/article/122785/gamergate-culture-war-people-who-dont-play-videogames>
- Clark, K. (2015). Clark calls for investigation and prosecution of online threats against women - Press Releases - Congresswoman Katherine Clark. Retrieved March 11, 2015, from <http://katherineclark.house.gov/index.cfm/press-releases?ID=2CB60BD7-D763-4464-96EB-1D113725559D>
- Colebrook, C. (2001). Certeau and Foucault: Tactics and Strategic Essentialism. *South Atlantic Quarterly*, 100(2), 543–574.
- Consalvo, M. (2012). Confronting toxic gamer culture: A challenge for feminist game studies scholars. *Ada: A Journal of Gender, New Media, and Technology*, 1. doi:10.7264/N33X84KH
- Constantine, J. (2013). Facebook Is Done Giving Its Precious Social Graph To Competitors. Retrieved August 14, 2016, from <http://social.techcrunch.com/2013/01/24/my-precious-social-graph/>
- Crawford, K., & Gillespie, T. (2014). What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society*, 1461444814543163. <http://doi.org/10.1177/1461444814543163>
- Creasy, S. (2014). Twitter troll hell: 'I can't get the last year of my life back'. Retrieved October 1, 2014, from <http://www.telegraph.co.uk/women/womens-life/11127782/Stella-Creasy-Twitter-troll-hell-I-cant-get-the-last-year-of-my-life-back.html>
- Criado-Perez, C. (2013a). We need women on British banknotes. Retrieved February 12, 2014, from <https://www.change.org/p/we-need-women-on-british-banknotes>
- Criado-Perez, C. (2013b). Women on bank notes: is the Bank of England finally listening? Retrieved February 12, 2014, from <http://www.theguardian.com/lifeandstyle/the-womens-blog-with-jane-martinson/2013/jul/04/women-bank-notes-bank-england>
- Criado-Perez, C. (2013c). After the Jane Austen announcement I suffered rape threats for 48 hours, but I'm still confident the trolls won't win. Retrieved

- August 10, 2013, from <http://www.newstatesman.com/media/2013/07/after-jane-austen-announcement-i-suffered-rape-threats-48-hours-im-still-confident-tro>
- Cristina, M. (2015). Fighting abuse to protect freedom of expression | Twitter Blogs. Retrieved December 30, 2015, from <https://blog.twitter.com/2015/fighting-abuse-to-protect-freedom-of-expression-au>
- Cross, K. (2015). Fictive Ethnicity and Nerds. In *Lean Out*. New York: OR Books.
- Cross, K. (2017). The Art of the Real. Retrieved June 9, 2017, from <https://thebaffler.com/salvos/art-of-the-real-cross>
- Cross, K. (2014). The nightmare is over: They're not coming for your games. Retrieved November 27, 2014, from <http://www.polygon.com/2014/7/29/5947661/the-nightmare-is-over-theyre-not-coming-for-your-games>
- Cush, A. (2014). These Are the Creepy 4Chan Successors Behind Gamergate's Pathetic War. Retrieved January 15, 2015, from <http://gawker.com/these-are-the-creepy-4chan-successors-behind-gamergates-1648966614>
- Danskin, I. (2015). Why Are You So Angry? Part 4: An Autopsy on GamerGate. Retrieved July 22, 2015, from <https://www.youtube.com/watch?v=c6TrKkkVEhs>
- Dash, A. (2013). The Web We Lost - Anil Dash. Retrieved December 9, 2013, from <http://dashes.com/anil/2012/12/the-web-we-lost.html>
- Davidson, H. (2013). Blocked Twitter users now able to interact with and follow accounts. Retrieved December 13, 2013, from <http://www.theguardian.com/media/2013/dec/13/blocked-twiters-users-can-still-interact-follow>
- Davies, R. (2012). Let's put Facebook's "no nipples" rule to test with your breastfeeding photos. Retrieved October 10, 2013, from <http://www.theguardian.com/commentisfree/2012/feb/22/facebook-no-nipples-rule-breastfeeding>
- de Certeau, M. (2002). *The Practice of Everyday Life*. (S. Rendall, Trans.) (2nd ed.). Berkeley, California: University of California Press.
- Dean, M. (1999). *Governmentality: Power and rule in modern society* (1st ed.). London: SAGE Publications.
- Department of Homeland Security. (2010). "If You See Something, Say Something™" Campaign. Retrieved December 14, 2013, from <http://www.dhs.gov/if-you-see-something-say-something%E2%84%A2-campaign>
- Dibbell, J. (1993). Julian Dibbell » A Rape in Cyberspace. Retrieved September 25, 2015, from <http://www.juliandibbell.com/articles/a-rape-in-cyberspace/>
- Dickey, M. R. (2017). Google+ won't go away. Retrieved January 22, 2017, from <http://social.techcrunch.com/2017/01/17/google-plus-updates/>

- Doctorow, C. (2011). Getting people's names right in software design: a LOT harder than it looks. Retrieved May 5, 2016, from <http://boingboing.net/2011/08/23/getting-peoples-names-right-in-software-design-a-lot-harder-than-it-looks.html>
- Doctorow, C. (2014). The Terrible Sea Lion: a social media parable. Retrieved October 30, 2014, from <http://boingboing.net/2014/10/28/the-terrible-sea-lion-a-socia.html>
- Duckett, C. (2013). Australia's chief law officer brands metadata a 'contestable concept'. Retrieved January 7, 2014, from <http://www.zdnet.com/australias-chief-law-officer-brands-metadata-a-contestable-concept-7000023859/>
- Dunleavy, P. (2003). *Authoring a PhD (Kindle Edition)*. New York: Palgrave Macmillan.
- Electronic Frontier Foundation. (n.d.). Free Speech is Only As Strong As the Weakest Link. Retrieved November 12, 2013, from <https://www.eff.org/free-speech-weak-link>
- Eleveth, R. (2014). A New Harassment Policy for Twitter. Retrieved December 3, 2014, from <http://www.theatlantic.com/technology/archive/2014/12/new-harassment-policy-for-twitter/383344/>
- Elgot, J. (2013). Twitter Rape Abuse Of Caroline Criado-Perez Leads To Boycott Threat. Retrieved August 28, 2013, from http://www.huffingtonpost.co.uk/2013/07/27/twitter-rape-abuse_n_3663904.html
- Ernst, A. (2014). GamerGate: How the video game industry's culture war began. Retrieved January 15, 2014, from <http://america.aljazeera.com/watch/shows/america-tonight/articles/2014/12/10/gamergate-women-gamers.html>
- Esco, L. (2013). Why I Made a Film Called Free the Nipple and Why I'm Being Censored in America. Retrieved December 15, 2013, from http://www.huffingtonpost.com/lina-esco/free-the-nipple_b_4415859.html
- Esguerra, R. (2009). Google CEO Eric Schmidt Dismisses the Importance of Privacy. Retrieved November 12, 2013, from <https://www.eff.org/deeplinks/2009/12/google-ceo-eric-schmidt-dismisses-privacy>
- Europe Versus Facebook. (n.d.). Legal Procedure against "Facebook Ireland Limited." Retrieved December 9, 2013, from <http://www.europe-v-facebook.org/EN/Complaints/complaints.html>
- Facebook. (2013a). What does Facebook consider to be hate speech? | Facebook Help Center. Retrieved December 9, 2013, from <https://www.facebook.com/help/135402139904490/>
- Facebook. (2013b). Statement of Rights and Responsibilities. Retrieved December 15, 2013, from <https://www.facebook.com/legal/terms>

- Facebook. (n.d.). Data Use Policy. Retrieved December 20, 2013, from <https://www.facebook.com/about/privacy/your-info>
- Facebook Safety. (2012). What Happens After You Click “Report.” Retrieved June 30, 2013, from <https://www.facebook.com/notes/facebook-safety/what-happens-after-you-click-report/432670926753695>
- Fasnacht, K. (2015). On shadowbans. • /r/self. Retrieved September 24, 2015, from https://np.reddit.com/r/self/comments/3ey0fv/on_shadowbans/
- Feminist Frequency. (2015). Speak Up & Stay Safe(r): – A Guide to Protecting Yourself From Online Harassment. Retrieved December 8, 2015, from <https://onlinesafety.feministfrequency.com/en/>
- Fiss, O. M. (1996). *The Irony of Free Speech*. Cambridge, MA: Harvard University Press.
- Fleishman, G. (2014). The Problem with Patreon: Nazis, Pedophilia Pushers, Harassers, and More. Retrieved October 28, 2014, from <http://glog.glennf.com/blog/2014/11/26/the-problem-with-patreon-nazis-harassers-and-more>
- Fletcher, D. (2009). Yahoo! Pulls Plug on GeoCities, Erases Internet History - TIME. Retrieved December 9, 2013, from <http://content.time.com/time/business/article/0,8599,1936645,00.html>
- Ford, C. (2015). Clementine Ford: Why I reported a man's abusive comment to his employer. Retrieved December 1, 2015, from <http://www.dailylife.com.au/news-and-views/dl-opinion/clementine-ford-why-i-reported-hotel-supervisor-michael-nolans-abusive-comment-to-his-employer-20151201-glcf96.html>
- Ford, P. (2011,). *The Web Is a Customer Service Medium*. Retrieved July 15, 2013, from <http://www.ftrain.com/wwic.html>
- Foucault, M. (1995). *Discipline and Punish: The Birth of the Prison*. (A. Sheridan, Trans.). New York: Vintage Books.
- Fox, J., & Tang, W. Y. (2016). Women’s experiences with general and sexual harassment in online video games: Rumination, organizational responsiveness, withdrawal, and coping strategies. *New Media & Society*. doi:10.1177/1461444816635778
- Foxman, A. H., & Wolf, C. (2013). *Viral Hate: Containing Its Spread on the Internet* (Kindle Edition). New York: Palgrave Macmillan. Retrieved from <http://www.amazon.com/Viral-Hate-Containing-Spread-Internet-ebook/dp/B00BRALE20/>
- Franklin, U. (1999). *The Real World of Technology* (2nd ed.). Toronto: House of Anansi Press.
- Franzen, C. (2017). A Team of Volunteers Is Archiving SoundCloud in Case It Dies. Retrieved July 16, 2017, from https://motherboard.vice.com/en_us/article/a3d4x8/archive-team-soundcloud

- Fried, L. (2011). UPDATE: I am back. Retrieved May 5, 2016, from <https://blog.adafruit.com/2011/07/23/google-suspended-limor-fried-ladyada-google-profile-no-show-and-tell-tonight/>
- Funnell, A. (2013). *A healthy digital ecology*. *abc.net.au*. ABC Radio National.
- Gadde, V. (2015). Twitter executive: Here's how we're trying to stop abuse while preserving free speech. Retrieved April 16, 2015, from <http://www.washingtonpost.com/posteverything/wp/2015/04/16/twitter-executive-heres-how-were-trying-to-stop-abuse-while-preserving-free-speech/>
- Galinsky, A. D., Wang, C. S., Whitson, J. A., Anicich, E. M., Hugenberg, K., & Bodenhausen, G. V. (2013). The reappropriation of stigmatizing labels: the reciprocal relationship between power and self-labeling. - PubMed - NCBI. *Psychological Science*, 24(10), 2020–2029.
- Gallegos, D., Vicca, N., & Streiner, S. (2013). Breastfeeding beliefs and practices of African women living in Brisbane and Perth, Australia. - PubMed - NCBI. *Maternal & Child Nutrition*, 11(4), n/a–n/a. <http://doi.org/10.1111/mcn.12034>
- Galperin, E. (2011a). Randi Zuckerberg Runs in the Wrong Direction on Pseudonymity Online. Retrieved November 24, 2014, from <https://www.eff.org/deeplinks/2011/08/randi-zuckerberg-runs-wrong-direction-pseudonymity>
- Galperin, E. (2011b). 2011 in Review: Nymwars | Electronic Frontier Foundation. Retrieved March 3, 2013, from <https://www.eff.org/deeplinks/2011/12/2011-review-nymwars>
- Galperin, H. (2004). Beyond Interests, Ideas, and Technology: An Institutional Approach to Communication and Information Policy. *The Information Society*, 20(3), 159–168. <http://doi.org/10.1080/01972240490456818>
- Gannes, L. (2011). Hey Facebook, Where's That Timeline and Open Graph You Promised? Retrieved January 7, 2014, from <http://allthingsd.com/20111205/hey-facebook-wheres-that-timeline-and-open-graph-you-promised/>
- Garcia, D., Mavrodiev, P., & Schweitzer, F. (2013). Social Resilience in Online Communities: The Autopsy of Friendster. *Cornell University Library*.
- Garside, J. (2013). Facebook's Graph Search tool causes increasing privacy concerns. Retrieved February 28, 2013, from <http://www.theguardian.com/technology/2013/jan/23/facebook-graph-search-privacy-concerns>
- Gehl, R. W. (2013). What's on your mind? Social media monopolies and noopower. *First Monday*, 18(3). <http://doi.org/10.5210/fm.v18i3.4618>
- Geller, J. (2013). Users = Zombies — on startups. Retrieved December 13, 2013, from <https://medium.com/on-startups/4575d5430061>
- George, C. (2013). The Internet as a Platform for Civil Disobedience. In J. Hartley, J. Burgess, & A. Bruns (Eds.), *A companion to new media dynamics* (1st ed., pp. 385–395). Chichester: Wiley-Blackwell.

- Gerring, J. (2004). What Is a Case Study and What Is It Good for? *American Political Science Review*, 98(02), 341–354.
<http://doi.org/10.1017/S0003055404001182>
- Gertz, G. (2013). Breastfeeding Stigma in America: Nursing Moms Lack Social Support. Retrieved June 1, 2015, from
<http://decodedpregnancy.com/breastfeeding-stigma-in-america/3098/>
- Gibbs, S. (2015). Troll uses Twitter adverts to urge transgender people to kill themselves. Retrieved May 22, 2015, from
<http://www.theguardian.com/technology/2015/may/20/troll-uses-twitter-adverts-to-urge-trans-people-to-kill-themselves>
- Gillespie, T. (2010). The politics of “platforms.” *New Media & Society*, 12(3), 347–364. <http://doi.org/10.1177/1461444809342738>
- Gillespie, T. (2011). Can an Algorithm be Wrong? Retrieved December 6, 2013, from <http://limn.it/can-an-algorithm-be-wrong/>
- Gillespie, T. (2017). Governance of and by platforms. (J. Burgess, T. Poell, & A. E. Marwick, Eds.) *Culture Digitally*.
- Gjoni, E. (2014). TL;DR. Retrieved August 24, 2014, from
<https://thezoepost.wordpress.com/2014/08/16/tldr-2/>
- Gjoni, E., Anonymous /r9k/ Users. (2014). FIVE GUYS - BURGER AND FRISE - NAKED EDITION — /r9k/ - Misogyny and Greentext Stories. Retrieved November 5, 2014, from <http://archive.is/m2fmN>
- Glazer, R. N. (1993). Women's Body Image and the Law. *Duke Law Journal*, 43(1), 113. <http://doi.org/10.2307/1372748>
- Google+. (2014). "When we launched Google+ over three years ago, we had a lot of restrictions on what name you could use on your profile...". Retrieved 24 July 2014 from <https://plus.google.com/+googleplus/posts/V5XkYQYYJqy>
- Go Topless Inc. (n.d.). What is GoTopless.org? Retrieved June 15, 2015, from <http://gotopless.org/>
- Godwin, M. (2015). Sure, call Trump a Nazi. Just make sure you know what you're talking about. Retrieved December 19, 2015, from
<https://www.washingtonpost.com/posteverything/wp/2015/12/14/sure-call-trump-a-nazi-just-make-sure-you-know-what-youre-talking-about/>
- Golumbia, D. (2016). *The Politics of Bitcoin : Software as Right-Wing Extremism* (Kindle Edition). Minnesota: University of Minnesota Press.
- Gonzales, C. (2014). Quick Tip: Use Tweetbot to mute unwanted tweets - The Sweet Setup. Retrieved April 30, 2015, from <http://thesweetsetup.com/quick-tip-use-tweetbot-mute-unwanted-tweets/>
- Gorshenin, V. (2013). Good bye, totalitarian Facebook. Retrieved December 25, 2013, from <http://english.pravda.ru/news/society/31-07-2013/125294-Facebook-0/>

- Goulston, M. (2011). Why Did Myspace Fail?... Are You Listening Microsoft, Yahoo, AOL, And Cisco? | Fast Company | Business + Innovation. Retrieved December 12, 2013, from <http://www.fastcompany.com/1764384/why-did-myspace-fail-are-you-listening-microsoft-yahoo-aol-and-cisco>
- Graham, K. (2013). Twitter: Add A Report Abuse Button To Tweets. Retrieved February 12, 2014, from <https://www.change.org/p/twitter-add-a-report-abuse-button-to-tweets>
- Greene, D. M., & Joseph, D. (2015). The Digital Spatial Fix. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 13(2), 223–247.
- Grimmelmann, J. (2009). Virtual World Feudalism. *The Yale Law Journal Pocket Part*, 118, 126.
- Gross, M. A., & Guerrero, L. K. (2000). Managing Conflict Appropriately and Effectively: an Application of the Competence Model to Rahim's Organizational Conflict Styles. *International Journal of Conflict Management*, 11(3), 200–226. <http://doi.org/10.1108/eb022840>
- GrrlScientist. (2011). Google's gormless 'no pseudonym' policy | GrrlScientist. Retrieved October 15, 2015, from <https://www.theguardian.com/science/punctuated-equilibrium/2011/jul/25/1>
- Grubb, J. (2014). Hackers post Fez developer Phil Fish's personal information and take control of his studio's website. Retrieved August 31, 2014, from <http://venturebeat.com/2014/08/22/hackers-post-fez-developer-phil-fishs-personal-information-and-take-control-of-his-studios-website/>
- Gutelle, S. (2013). YouTube Is Suddenly Flagging A LOT Of Videos For Copyright Claims. Retrieved December 11, 2013, from <http://www.tubefilter.com/2013/12/10/youtube-contentid-flag-gaming-videos/>
- Halliday, J. (2011). Justin Timberlake buys his own social network with Myspace investment | Technology | theguardian.com. Retrieved December 9, 2013, from <http://www.theguardian.com/technology/2011/jun/30/myspace-internet>
- Hankes, K. (2016). How the extremist right hijacked "Star Wars," Taylor Swift and the Mizzou student protests to promote racism. Retrieved January 5, 2016, from <https://www.splcenter.org/hatewatch/2016/01/05/how-extremist-right-hijacked-%E2%80%98star-wars%E2%80%99-taylor-swift-and-mizzou-student-protests-promote>
- Hansen, E. E. (2016). Dear Mark Zuckerberg. I shall not comply with your requirement to remove this picture. Retrieved September 9, 2016, from <http://www.aftenposten.no/article/ap-604156b.html>
- Harper, R. (2014a). ggautoblocker/sourcelist.txt at master • freebsdgirl/ggautoblocker. Retrieved October 5, 2015, from <https://github.com/freebsdgirl/ggautoblocker>
- Harper, R. (2014b). Good Game Auto Blocker. Retrieved October 5, 2015, from <http://blog.randi.io/good-game-auto-blocker/>

- Hathaway, J. (2015). 4chan Troll Claims He Framed Feminist for Transphobic Promoted Tweet. Retrieved May 20, 2015, from <http://gawker.com/4chan-troll-claims-he-framed-feminist-for-transphobic-p-1705889031>
- Hegel, G. W. F. (2012). *Hegel: Lectures on Natural Right and Political Science: The First Philosophy of Right*. (O. Poggeler, Ed., J. M. Stewart & P. C. Hodgson, Trans.). Oxford University Press.
- Helft, M. (2011). Facebook, Foe of Anonymity, Is Forced to Explain a Secret. Retrieved October 15, 2015, from <http://www.nytimes.com/2011/05/14/technology/14facebook.html>
- Herold, E., Corbesi, B., & Collins, J. (1994). Psychosocial aspects of female topless behavior on Australian beaches. *Journal of Sex Research*, 31(2), 133–142. <http://doi.org/10.1080/00224499409551740>
- Hess, A. (2014). The Next Civil Rights Issue: Why Women Aren't Welcome on the Internet. Retrieved January 7, 2014, from <http://www.psmag.com/navigation/health-and-behavior/women-arent-welcome-internet-72170/>
- Hess, A. (2016). For the Alt-Right, the Message Is in the Punctuation. Retrieved June 13, 2016, from <https://www.nytimes.com/2016/06/11/arts/for-the-alt-right-the-message-is-in-the-punctuation.html>
- Hill, K. (2013). The Death And Rebirth Of @ProfJeffJarvis, The Best Parody Account On Twitter. Retrieved December 11, 2013, from <http://www.forbes.com/sites/kashmirhill/2013/11/04/the-death-and-rebirth-of-profjeffjarvis-the-best-parody-account-on-twitter/>
- Hill, K. (2014). Meet Del Harvey, Twitter's Troll Patrol - Forbes. (2014). Meet Del Harvey, Twitter's Troll Patrol - Forbes. Retrieved August 20, 2014, from <https://www.forbes.com/sites/kashmirhill/2014/07/02/meet-del-harvey-twiters-troll-patrol/>
- Hinge, D. (2013). Bank of England faces legal challenge over all-male banknote roster. Retrieved February 12, 2014, from <http://www.centralbanking.com/central-banking/news/2274214/bank-of-england-faces-legal-challenge-over-allmale-banknote-roster>
- Hoffman-Andrews, J. (n.d.). Block Together. Retrieved April 30, 2015, from <https://blocktogether.org/>
- Holpuch, A. (2015). Instagram ban on emoji has sexters searching for fruity alternatives.
- Hood, J. C. (2007). Orthodoxy vs Power: The defining traits of grounded theory. In *The Sage Handbook of Grounded Theory* (pp. 151–164). Thousand Oaks, CA: SAGE.
- Hopkins, N. (2017). Revealed: Facebook's internal rulebook on sex, terrorism and violence. Retrieved May 23, 2017, from <http://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence>

- Horowitz, B. (2012). Toward a more inclusive naming policy for Google+ With Google+, we aspire to... Retrieved November 24, 2014, from <https://plus.google.com/+BradleyHorowitz/posts/SM5RjubMmV>
- Hudson, L. (2015a). Gamergate Target Zoe Quinn Launches Anti-Harassment Support Network. Retrieved January 20, 2015, from <http://www.wired.com/2015/01/gamergate-anti-harassment-network/>
- Hudson, L. (2015b). 6 Experts on How Silicon Valley Can Solve Online Harassment. Retrieved October 23, 2015, from <http://www.wired.com/2015/10/how-silicon-valley-can-solve-online-diversity-and-harassment/>
- Humphreys, A. M. (2005). *Massively Multiplayer Online Games Productive players and their disruptions to conventional media practices*. eprints.qut.edu.au.
- Humphreys, S. (2008). "Ruling the virtual world. Governance in massively multiplayer online games." *European Journal of Cultural Studies* 11(2): 149-171.
- Humphreys, S. (2015). Virtual World Governance. *The International Encyclopedia of Digital Communication and Society*. R. Mansell and P. H. Ang. Online, Wiley Online.
- Hussein, S. (2015). Did FBI Islamophobia entrap an American Jew? Retrieved September 20, 2015, from <http://www.crikey.com.au/2015/09/18/did-fbi-islamophobia-entrap-an-american-jew/>
- Ingram, M. (2012). Careful, Twitter — remember what happened to MySpace and Digg. *Giga Om*.
- Ingram, M. (2011). The downside of Facebook as a public space: Censorship. Retrieved October 10, 2013, from <http://gigaom.com/2011/06/21/the-downside-of-facebook-as-a-public-space-censorship/>
- Ingram, M. (2013). The flip-flop over beheadings: Facebook finds being a media entity isn't as easy as it looks. Retrieved October 24, 2013, from <http://gigaom.com/2013/10/23/the-flip-flop-over-beheadings-facebook-finds-being-a-media-entity-isnt-as-easy-as-it-looks/>
- Jason, Z. (2015,). Game of Fear: The Story Behind GamerGate. Retrieved June 3, 2015, from <http://www.bostonmagazine.com/news/article/2015/04/28/gamergate/>
- Jeong, S. (2013.) "Revenge Porn is Bad. Criminalizing It Is Worse." *Wired*. Retrieved January 4, 2019 from <https://www.wired.com/2013/10/why-criminalizing-revenge-porn-is-a-bad-idea/>
- Jeong, S. (2015). *The Internet Of Garbage* (Kindle Edition). Forbes Media. Retrieved from <http://www.amazon.com/The-Internet-Garbage-Sarah-Jeong-ebook/dp/B011JAV030>
- Jeong, S. (2017). Mastodon Is Like Twitter Without Nazis, So Why Are We Not Using It? Retrieved April 6, 2017, from https://motherboard.vice.com/en_us/article/mastodon-is-like-twitter-without-nazis-so-why-are-we-not-using-it

- Johnson, L. (2016). Blizzard Erases Gaming History By Axing a Fan-Made “World of Warcraft” Server. Retrieved April 20, 2016, from https://motherboard.vice.com/en_us/article/world-of-warcraft-legacy-server-shut-down
- Johnson, N. A., Cooper, R. B., & Chin, W. W. (2008). Anger and flaming in computer-mediated negotiation among strangers. Retrieved January 14, 2014, from <http://www.sciencedirect.com/science/article/pii/S0167923608001851>
- Johnston, C. (2014). The death of the “gamers” and the women who ‘killed’ them. Retrieved August 30, 2014, from <http://arstechnica.com/gaming/2014/08/the-death-of-the-gamers-and-the-women-who-killed-them/>
- Johnston-Robledo, I., Wares, S., Fricker, J., & Pasek, L. (2007). Indecent Exposure: Self-objectification and Young Women’s Attitudes Toward Breastfeeding. *Sex Roles, 56*(7-8), 429–437. <http://doi.org/10.1007/s11199-007-9194-4>
- Jones, M. (2011). Myspace: Five lessons learned from the front - Fortune Tech. Retrieved December 12, 2013, from <http://tech.fortune.cnn.com/2011/10/24/mike-jones-myspace/>
- Jones, R. (2014). Wizardchan Is A Place For Male Virgins Over 30 To Make Fun Of Themselves - ANIMAL. Retrieved December 12, 2014, from <http://animalnewyork.com/2014/wizardchan-place-male-virgins-30-make-fun/>
- Joseph, D., & Williams, I. (2015). Games Without Wages. Retrieved July 28, 2015, from <http://www.tandfonline.com/doi/abs/10.1080/1369118032000163277>
- Jozuka, E. (2014). Twitter troll jailed for abusive tweets against UK MP (Wired UK). Retrieved October 5, 2014, from <http://www.wired.co.uk/news/archive/2014-09/29/twitter-troll-nunn-jailed-mp-creasy>
- Kain, E. (2012). All The Pretty Doritos: How Video Game Journalism Went Off The Rails - Forbes. Retrieved October 17, 2015, from <http://www.forbes.com/sites/erikkain/2012/10/26/all-the-pretty-doritos-how-video-game-journalism-went-off-the-rails/#6731c0905c01>
- Kaliya. (2011). Open Letter to Google+ Profile Support. Retrieved March 15, 2015, from <https://identitywoman.net/open-letter-to-google-profile-support/>
- Katz, E. (1988). Disintermediation: Cutting Out the Middle Man. Retrieved December 21, 2013, from http://repository.upenn.edu/cgi/viewcontent.cgi?article=1165&context=asc_papers
- Kelly, H. (2012). Voting closes on Facebook policy changes, only 299 million votes short. Retrieved December 5, 2013, from <http://edition.cnn.com/2012/12/10/tech/social-media/facebook-policy-vote/>
- Kelly, H. (2013). U.S. court says “Liking” something on Facebook is free speech - CNN.com. Retrieved December 4, 2013, from <http://edition.cnn.com/2013/09/18/tech/social-media/facebook-likes-free-speech/>

- Kessler, S. (2013). 4 Persuasion Tricks Facebook Uses To Keep You From Quitting. Retrieved December 9, 2013, from <http://www.fastcompany.com/3017530/4-persuasion-tricks-facebook-uses-to-keep-you-from-quitting#2>
- Kirkpatrick, M. (2010). Facebook's Zuckerberg Says The Age of Privacy is Over. Retrieved November 12, 2013, from http://readwrite.com/2010/01/09/facebooks_zuckerberg_says_the_age_of_privacy_is_ov#awesm=~omWmQ8YLMr9nHa
- Klepek, P. (2012). When Passions Flare, Lines Are Crossed [UPDATED]. Retrieved March 18, 2015, from <http://www.giantbomb.com/articles/when-passions-flare-lines-are-crossed-updated/1100-4006/>
- Knibbs, K. (2015). Twitter Got Paid to Promote a Tweet About "Killing Arabs". Retrieved May 8, 2015, from <http://gizmodo.com/twitter-got-paid-to-promote-a-tweet-about-killing-arab-1702883517>
- Knight, W. (2017). There's a big problem with AI: even its creators can't explain how it works. Retrieved April 13, 2017, from <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>
- Know Your Meme. (2014a). Sea-Lioning. Retrieved November 12, 2014, from <http://knowyourmeme.com/memes/sea-lioning>
- Know Your Meme. (2014b). Actually It's About Ethics. Retrieved November 27, 2014, from <http://knowyourmeme.com/memes/actually-its-about-ethics>
- Koh, Y. (2015). Boss Talk: Twitter Safety Officer Talks Building Trust. Retrieved February 17, 2015, from <http://www.wsj.com/articles/boss-talk-twitter-safety-officer-talks-building-trust-1424221141>
- Komodiki E, Kontogeorgou A, Papastavrou M, Volaki P, RMidw, et al. (2014). Breastfeeding in Public: A Global Review of Different Attitudes towards It. *J Pediatr Neonatal Care* 1(6): 00040. DOI: 10.15406/jpnc.2014.01.00040
- Konstantinidis, C. (2003). Open Networks, Closed Regimes: The Impact of the Internet on Authoritarian Rule, and: Democracy and the Internet: Allies or Adversaries? (review). *SAIS Review*, 23(2), 245–248. <http://doi.org/10.1353/sais.2003.0042>
- Koster, R. (2017). Still Logged In: What AR and VR Can Learn from MMOs. Retrieved March 18, 2017, from <https://vimeo.com/208372546>
- Kou, Y., & Nardi, B. (2014). Governance in League of Legends: A Hybrid System. Presented at the Foundation of Digital Games.
- Kozyra, J., & Muir, S. (2009). Breastfeeding uncut, with Stephanie Muir of Facebook petition fame. Retrieved October 15, 2013, from <http://blog.oncloudmom.com/2009/03/breastfeeding-uncut-with-stephanie-muir.html>
- Kunyosying, K., & Soles, C. (2012). Postmodern geekdom as simulated ethnicity by Kom Kunyosying and Carter Soles. *Jump Cut - a Review of Contemporary Media*. Retrieved June 31, 2016, from <https://www.ejumpcut.org/archive/jc54.2012/SolesKunyoGeedom/>

- Kushner, D. (2015). 4chan's Overlord Christopher Poole Reveals Why He Walked Away | Rolling Stone. Retrieved March 15, 2015, from <http://www.rollingstone.com/culture/features/4chans-overlord-christopher-poole-reveals-why-he-walked-away-20150313>
- Kwasnica, E. (2009). One small step for breastfeeding..... Retrieved February 15, 2014, from <http://bliss-breastfeeding.blogspot.com/2009/01/hello-my-name-is-emma-kwasnica.html>
- Lanier, J. (2013). *Who Owns The Future?* (Kindle Edition). Penguin. Retrieved from <http://www.amazon.com/Who-Owns-Future-Jaron-Lanier-ebook/dp/B00ADNP2ZM/>
- Lapowsky, I. (2015). It's Too Easy for Trolls to Game Twitter's Anti-Abuse Tools. Retrieved May 15, 2015, from <http://www.wired.com/2015/05/wam-twitter-harassment/>
- Lastowka, G. (2010). *Virtual Justice: The New Laws of Online Worlds*. London: Yale University Press.
- Lee, A. (2013). Myspace Collapse: How The Social Network Fell Apart. Retrieved December 12, 2013, from http://www.huffingtonpost.com/2011/06/30/how-myspace-fell-apart_n_887853.html
- Lessig, L. (2006). *Code: Version 2.0*. New York: Basic Books.
- Lesteshen, R. (2013). Banned from Facebook: Did you know a picture of an egg is considered nudity/pornography? | unlatched. Retrieved October 10, 2013, from <http://unlatched.wordpress.com/2013/08/27/banned-from-facebook-did-you-know-a-picture-of-an-egg-is-considered-nuditypornography/>
- Levine, M. (2013). Controversial, Harmful and Hateful Speech on Facebook. Retrieved May 29, 2013, from <https://www.facebook.com/notes/facebook-safety/controversial-harmful-and-hateful-speech-on-facebook/574430655911054>
- Light, B. (2014). *Disconnecting with social networking sites*. London: Palgrave Macmillan.
- Livingston, R., McGinley, J. C., & Wilson, P. (1999). *Office Space*. (M. Judge, Ed.) (Vol. 89). 20th Century Fox.
- Lomas, N. (2014). #Gamergate Shows Tech Needs Far Better Algorithms. Retrieved October 26, 2014, from <http://techcrunch.com/2014/10/18/gamergate-tactics/>
- Love, D. (2013). THINKFLUENCER: Meet The Guy Who Runs Our favourite Twitter Parody Account. Retrieved November 30, 2013, from <http://www.businessinsider.com.au/fake-jeff-jarvis-interview-2013-4>
- Ludlam, S. (2012). Australian Federal Police | Scott Ludlam. Retrieved January 7, 2014, from <http://scott-ludlam.greensmps.org.au/content/estimates/australian-federal-police-3>
- Lunceford, B. (2012). *Naked Politics : Nudity, Political Action, and the Rhetoric of the Body* (eBook, pp. 35–82). Lexington Books.

- Machkovech, S. (2015a). Full transcript: Ars interviews 8chan founder Fredrick Brennan. Retrieved March 20, 2015, from <http://arstechnica.com/business/2015/03/full-transcript-ars-interviews-8chan-founder-fredrick-brennan/>
- Machkovech, S. (2015b). GamerGate critic posts death threat voicemail, requests subpoena [Updated]. Retrieved March 25, 2015, from <http://arstechnica.com/business/2015/05/gamergate-critic-posts-death-threat-voicemail-after-inaction-by-prosecutor/>
- Machkovech, S. (2015c). 8chan-hosted content disappears from Google searches [Updated]. Retrieved August 18, 2015, from <http://arstechnica.com/tech-policy/2015/08/8chan-hosted-content-disappears-from-google-searches/>
- MacKinnon, R. (2013). *Consent of the Networked: The Worldwide Struggle for Internet Freedom (Revised Paperback Edition)*. New York: Basic Books.
- Madrigal, A. C. (2011). Why Facebook and Google's Concept of 'Real Names' Is Revolutionary. Retrieved September 28, 2015, from <http://www.theatlantic.com/technology/archive/2011/08/why-facebook-and-googles-concept-of-real-names-is-revolutionary/243171/>
- Mantilla, K. (2015). *Gender trolling: How Misogyny Went Viral*. Santa Barbara, CA: Praeger.
- Maslow, A. H. (1966). *The Psychology of Science*. Joanna Cotler Books.
- Marwick, A. (2013). "Online Identity". Companion to New Media Dynamics J. Hartley, J. Burgess and A. Bruns, John Wiley & Sons 355-364.
- Massanari, A. (2015). #Gamergate and The Fapping: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society*, 1461444815608807. <http://doi.org/10.1177/1461444815608807>
- Matias, J. N., Johnson, A., Boesel, W. E., Keegan, B., Friedman, J., & DeTar, C. (2015). Reporting, Reviewing, and Responding to Harassment on Twitter. Retrieved May 14, 2015, from <http://womenactionmedia.org/cms/assets/uploads/2015/05/wam-twitter-abuse-report.pdf>
- McCullagh, D. (2011). Google+ will let users conceal gender. Retrieved October 15, 2015, from <http://www.cnet.com/news/google-will-let-users-conceal-gender/>
- McCullagh, D. (2012). Instagram says it now has the right to sell your photos | Politics and Law. Retrieved January 7, 2014, from http://news.cnet.com/8301-13578_3-57559710-38/instagram-says-it-now-has-the-right-to-sell-your-photos/
- Mele, C. (2016). Facebook Banned Gun Sales. So Why Is It Still "Full of Them?" Retrieved July 23, 2016, from <https://www.nytimes.com/2016/07/22/technology/facebook-banned-gun-sales-so-why-is-it-still-full-of-them.html>
- Melendez, S. (2014). The Secret Meaning Behind GamerGate's Branding. Retrieved November 3, 2014, from <http://www.fastcodesign.com/3037941/the-secret-meaning-behind-gamergates-branding>

- Meyer, E. (2014). Inadvertent Algorithmic Cruelty. Retrieved October 31, 2015, from <http://meyerweb.com/eric/thoughts/2014/12/24/inadvertent-algorithmic-cruelty/>
- Mio, J. S. (1997). Metaphor and Politics. *Metaphor and Symbol*, 12(2), 113–133. http://doi.org/10.1207/s15327868ms1202_2
- Molitch-Hou, M. (2017). Twitter's Nazi problem is still out of control. Retrieved June 17, 2017, from <https://theoutline.com/post/1743/twitter-s-nazi-problem-is-still-out-of-control>
- Moll, E. (2014). What's in a Nym? Gender, Race, Pseudonymity, and the Imagining of the Online Persona. *M/C Journal*, 17(3).
- Moore, M. (2012). China moves to control Sina Weibo social network with real names. Retrieved December 10, 2014, from <http://www.telegraph.co.uk/technology/news/9147767/China-moves-to-control-Sina-Weibo-social-network.html>
- Morozov, E. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism* (Kindle Edition). PublicAffairs.
- Morozov, E. (2012). Disintermediation: We aren't seeing fewer gatekeepers. We're seeing more. Retrieved October 16, 2013, from http://www.slate.com/articles/technology/future_tense/2012/10/disintermediation_we_aren_t_seeing_fewer_gatekeepers_we_re_seeing_more.single.html
- Morris, N. (2013). Woolwich aftermath: Home Secretary Theresa May threatens ban for online radical Islamists - even those who don't advocate violence. Retrieved December 1, 2013, from <http://www.independent.co.uk/news/uk/politics/woolwich-aftermath-home-secretary-theresa-may-threatens-ban-for-online-radical-islamists--even-those-who-dont-advocate-violence-8632652.html>
- moulee. (2009). Tamil name structure.... Retrieved May 25, 2017, from <https://bumpahead.net/tamil-name-structure-10dd63730563>
- Moyer, J. W. (2015). Florida man plotted Sept. 11 attack on Kansas City, Mo., FBI says. Retrieved September 11, 2015, from <https://www.washingtonpost.com/news/morning-mix/wp/2015/09/11/florida-man-plotted-sept-11-attack-in-kansas-city-fbi-says/>
- Mulready-Ward, C., & Hackett, M. (2014). Perception and attitudes: breastfeeding in public in New York City. *Journal of Human Lactation*, 30(2), 195–200. <http://doi.org/10.1177/0890334414524988>
- Myerscough, P. (2013). Short Cuts. *London Review of Books*, 35(1), 25.
- Nakamura, L. (2015). Afterword. Blaming, shaming and the feminization of social media. *Feminist Surveillance Studies*. R. Dubrofsky and S. A. Magnet. London, Duke University Press.
- Newman, J. (2009). You Own Your Tweets, But So Does Twitter? Retrieved January 7, 2014, from

- http://www.pcworld.com/article/171892/You_Own_Your_Tweets_But_So_Does_Twitter.html
- Nimmo, J. (2014). My experience as a convicted twitter bully. Retrieved March 5, 2015, from <https://trollslife.wordpress.com/2014/11/29/my-experience-2/>
- Noelle-Neumann, E. (1974). The Spiral of Silence A Theory of Public Opinion. *Journal of Communication*, 24(2), 43–51. <http://doi.org/10.1111/j.1460-2466.1974.tb00367.x>
- Noguchi, S. (2008). Protests mount over Facebook ban on breast-feeding photos; bigger turnout online than in Palo Alto. Retrieved April 1, 2015, from http://www.mercurynews.com/ci_11321458?source%253Dmost_viewed.20F88DA3D7D369F5BB70F372987EAE1F.html
- Nordby, K. J. (2016). Norway's prime minister and several government members censored by Facebook. Retrieved September 10, 2016, from <http://www.aftenposten.no/article/ap-604195b.html>
- O'Farrell, C. (2005). Michel Foucault. London: SAGE Publications.
- OMalley, H. (2014). What We Talk About When We Talk About GamerGate. Retrieved November 17, 2014, from <http://www.doctornerdlove.com/2014/10/when-we-talk-about-gamergate/>
- O'Neil, C (2016), Weapons of Math Destruction: How big data increases inequality and threatens democracy, Penguin Random House UK, UK.
- Orlean, S. (2013). Horse_ebooks Is Human After All - The New Yorker. Retrieved May 5, 2016, from http://www.newyorker.com/tech/elements/horse_ebooks-is-human-after-all
- Osofsky, J., & Gage, T. (2015). Community Support FYI: Improving the Names Process on Facebook | Facebook Newsroom. Retrieved December 22, 2015, from <https://newsroom.fb.com/news/2015/12/community-support-fyi-improving-the-names-process-on-facebook/>
- Otton, A. (2014). Understanding pro and anti-Gamergate |. Retrieved March 3, 2015, from <https://techraptor.net/content/understanding-pro-anti-gamergate>
- Oyler, L. (2015). The History of Toplessness | Broadly. Retrieved August 26, 2015, from https://broadly.vice.com/en_us/article/the-history-of-toplessness
- Packer, J. (2013). “Epistemology Not Ideology OR Why We Need New Germans.” *Communication and Critical/Cultural Studies*, 1–6. <http://doi.org/10.1080/14791420.2013.806154>
- Pantozzi, J. (2014). No, the Game Journalist Who Contacted the Moms of Threatening Trolls Is Not Being Sued. Retrieved February 2, 2015, from <http://www.themarysue.com/game-journalist-not-being-sued/>
- Panzarino, M. (2012). IFTTT forced to remove Twitter triggers on September 27th to comply with new API policies. Retrieved September 20, 2012, from <http://thenextweb.com/apps/2012/09/20/ifttt-removes-twitter-triggers-comply-new-api-policies/>

- Parikka, J. (2012). *What is Media Archaeology?* Cambridge: Polity.
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
- Pasquale, F., & Vaidhyanathan, S. (2016). Uber and the lawlessness of “sharing economy” corporates. Retrieved July 28, 2016, from <http://www.theguardian.com/technology/2015/jul/28/uber-lawlessness-sharing-economy-corporates-airbnb-google>
- Payton, N. (2015). Twitter users shown transphobic promoted tweets from account impersonating feminist writer. Retrieved May 22, 2015, from <http://www.pinknews.co.uk/2015/05/20/twitter-users-shown-transphobic-promoted-tweet/>
- Penenberg, A. L. (2015). The Troll’s Lawyer — Backchannel. Retrieved January 12, 2015, from <https://medium.com/backchannel/the-trolls-lawyer-8bf7b2283>
- Penny, L. (2013a). *Cybersexism: Sex, Gender and Power on the Internet* (Kindle edition). Bloomsbury. Retrieved from <http://www.amazon.com/Cybersexism-Sex-Gender-Power-Internet-ebook/dp/B00EO24J3O/>
- Penny, L. (2013b). Laurie Penny on online aggression: What do you do the day after a death threat? Retrieved August 10, 2013, from <http://www.newstatesman.com/lifestyle/2013/08/what-do-you-do-day-after-death-threat>
- Perez, S. (2013). Tumblr’s Adult Fare Accounts For 11.4% Of Site’s Top 200K Domains, Adult Sites Are Leading Category Of Referrals. Retrieved October 15, 2013, from <http://social.techcrunch.com/2013/05/20/tumblrs-adult-fare-accounts-for-11-4-of-sites-top-200k-domains-tumblrs-adult-fare-accounts-for-11-4-of-sites-top-200k-domains-adults-sites-are-leading-category-of-referrals/>
- Petit, C. (2015). GamerGate Gets One Right. Retrieved January 15, 2015, from <http://agameofme.tumblr.com/post/106975098317/gamergate-gets-one-right>
- Pfaffenberger, B. (2012). “A Standing Wave in the Web of Our Communities”: Usenet and the Socio-Technical Construction of Cyberspace Values. In C. Lueg & D. Fisher (Eds.), *From Usenet to CoWebs: Interacting with Social Information Spaces* (pp. 20–44). London: Springer-Verlag.
- Phipps, B. (2012). Sorry, Facebook – on breastfeeding you seem rather confused | Belinda Phipps. Retrieved October 10, 2013, from <http://www.theguardian.com/commentisfree/2012/feb/29/facebook-breastfeeding-photo-policy-confused>
- Phillips, W. (2015). *This is why we can’t have nice things. Mapping the relationship between online trolling and mainstream culture*. Cambridge, MA: MIT Press.
- Plenke, M. (2015). Trolls Are Paying Twitter to Promote Disturbing Hate Speech. Retrieved May 20, 2015, from <http://mic.com/articles/118862/trolls-are-paying-twitter-to-promote-hate-speech-and-there-s-nothing-stopping-them>
- Pless, M. (2015a). Sargon of Akkad and Thunderf00t: #Gamergate's Well-Paid Talking Heads. Retrieved September 28, 2015, from

- <http://www.dailykos.com/story/2015/3/17/1370280/-Sargon-of-Akkad-and-Thunderf00t-Gamergate-s-Well-Paid-Talking-Heads>
- Pless, M. (2015b). Is Milo Yiannopoulos Dependent Upon 8chan and Baphomet for Stories? Retrieved January 12, 2016, from <http://www.dailykos.com/story/2015/11/3/1443784/-Is-Milo-Yiannopoulos-Dependent-Upon-8chan-and-Baphomet-for-Stories>
- Pless, M. (2016a). Family Business Brigaded on Yelp by MRA YouTube Stars. Retrieved February 1, 2016, from <http://www.dailykos.com/story/2015/10/22/1437368/-Family-Business-Brigaded-on-Yelp-by-MRA-YouTube-Stars>
- Pless, M. (2016b). Thunderf00t's Footman - A Case Study in Online Intimidation. Retrieved January 14, 2016, from <http://www.dailykos.com/story/2016/1/12/1466042/-Thunderf00t-s-Footman-A-Case-Study-in-Online-Intimidation>
- Plunkett, L. (2012). Yes, a Games Writer was Fired Over Review Scores. Retrieved October 17, 2015, from <http://kotaku.com/5893785/yes-a-games-writer-was-fired-over-review-scores>
- Politi, D. (2013). Twitter U.K. Boss Apologizes to Abuse Victims. Retrieved August 28, 2013, from http://www.slate.com/blogs/the_slatest/2013/08/03/tony_wang_twitter_uk_boss_apologizes_to_abuse_victims.html
- Potaka, E., & McMahon, L. (2015). Unmasking a troll: Aussie “jihadist” Australi Witness a 20-year-old American nerd. Retrieved September 12, 2015, from <http://www.smh.com.au/national/unmasking-a-troll-aussie-jihadist-australi-witness-a-20yearold-american-nerd-20150909-gjil47.html>
- Poulsen, K. (2010). PayPal Freezes WikiLeaks Account. Retrieved March 4, 2016, from <https://www.wired.com/2010/12/paypal-wikileaks/>
- Quinn, B. (2013). Twitter bomb threats made against more women in public eye. Retrieved August 12, 2013, from <http://www.theguardian.com/technology/2013/aug/05/twitter-bomb-threats-women>
- Quinn, Z. (2015). Zoe Quinn, Crash Override - XOXO Festival (2015). Retrieved November 17, 2015, from <https://www.youtube.com/watch?v=vAcdKTXtx1k>
- Quinn, Z. (2016a). Zoë “Shitpost” Quinn on Twitter: ‘any time I create something or link to it, I'm promoting my abusers channels directly to my own audience. I raise their profile by existing.’. Retrieved January 6, 2016, from <https://twitter.com/UnburntWitch/status/684915954808037377>
- Quinn, Z. (2016b). Zoë “Shitpost” Quinn on Twitter: ‘if this shit involved pirated game of thrones episodes though, none of this would be an issue.’. Retrieved January 6, 2016, from <https://twitter.com/UnburntWitch/status/684919876834623488>

- Quinn, Z. (2016c). Zoë “Shitpost” Quinn on Twitter: ‘People aren't sending these videos *to* me in this example. The platforms themselves are doing it for them.’. Retrieved January 6, 2016, from <https://twitter.com/UnburntWitch/status/684917563747258369>
- Quinn, Z. (2016d). Zoë “Shitpost” Quinn on Twitter: ‘So my choice is disappear, or help abusers profit off of my own body of work.’. Retrieved January 6, 2016, from <https://twitter.com/UnburntWitch/status/684917003883130880>
- Quinn, Z. (2016e). Zoë “Shitpost” Quinn on Twitter: ‘So while platforms do little to nothing about perpetually abusive content/users, they help strengthen their connections to me like this.’. Retrieved January 6, 2016, from <https://twitter.com/UnburntWitch/status/684916457688305664>
- Quinn, Z. (2016f). Zoë “Shitpost” Quinn on Twitter: ‘The suggestions are content-neutral and only care about how many clicks they get. They turn me into an advertisement for my own hate.’. Retrieved January 6, 2016, from <https://twitter.com/UnburntWitch/status/684916301341429762>
- Quinn, Z. (2016g). Zoë “Shitpost” Quinn on Twitter: ‘This makes major platforms more than even just negligent in moderation. This makes them an active tool for unchecked abuse.’. Retrieved January 6, 2016, from <https://twitter.com/UnburntWitch/status/684917441873362944>
- Quinn, Z. (2016h). 2016. Retrieved January 12, 2016, from <http://blog.unburntwitch.com/post/137105243484/2016>
- Quodling, A. (2018). Shadow profiles - Facebook knows about you, even if you're not on Facebook. The Conversation. <https://theconversation.com/shadow-profiles-facebook-knows-about-you-even-if-youre-not-on-facebook-94804>
- Rapoport, P. (n.d.). Topfree Equal Rights Association. Retrieved June 15, 2015, from <http://www.tera.ca/>
- RationalWiki. (n.d.). Just asking questions - RationalWiki. Retrieved January 29, 2016, from http://rationalwiki.org/wiki/Just_asking_questions
- Rauch, J. (1995). In defense of prejudice. Retrieved December 10, 2013, from <http://harpers.org/archive/1995/05/in-defense-of-prejudice/>
- Read, M. (2012). That Facebook Copyright Thing Is Meaningless and You Should Stop Sharing It. Retrieved January 7, 2014, from <http://gawker.com/5963225/that-facebook-copyright-thing-is-meaningless-and-you-should-stop-sharing-it>
- Rebant. (2014). The Escapist : Forums : Off-topic Discussion : Zoe Quinn and the surrounding controversy. Retrieved November 5, 2014, from <http://www.escapistmagazine.com/forums/read/18.858347-Zoe-Quinn-and-the-surrounding-controversy#21276142>
- Reed, J., & Gardiner, B. (2012). The beautiful breastfeeding images Facebook is missing out on. Retrieved October 10, 2013, from <http://www.theguardian.com/commentisfree/2012/feb/23/breastfeeding-images-facebook-missing>

- Riddle, H. (2014). The Queen -v- John Raymond Nimmo and Isabella Kate Sorley. Retrieved June 30, 2014, from <https://www.judiciary.gov.uk/wp-content/uploads/JCO/Documents/Judgments/r-v-nimmo-and-sorley.pdf>
- Rifkind, H. (2013). Hugo Rifkind: What is Facebook? Retrieved December 9, 2013, from <http://www.spectator.co.uk/columnists/hugo-rifkind/9063931/what-is-facebook-thats-the-problem-nobody-really-knows/>
- Rigby, E. (2013). Proposed UK extremist broadcast and web ban hits resistance. Retrieved December 4, 2013, from <http://www.ft.com/intl/cms/s/0/bae5e470-c6e1-11e2-8a36-00144feab7de.html#axzz2mThvSWij>
- Robertson, A. (2015). Was Reddit always about free speech? Yes, and no | The Verge. Retrieved June 15, 2015, from <http://www.theverge.com/2015/7/15/8964995/reddit-free-speech-history>
- Robinson, J. (2011). Twitter and Facebook riot restrictions would be a mistake, says Google chief | Media | theguardian.com. Retrieved October 22, 2013, from <http://www.theguardian.com/media/2011/aug/27/twitter-facebook-riot-restrictions-eric-schmidt>
- Roiphe, K. (2013). Austengate. Retrieved August 10, 2013, from http://www.slate.com/articles/double_x/roiphe/2013/08/the_anger_over_jane_austen_on_a_10_pound_note_proves_people_can_rage_over.html
- Romano, A. (2013a). Anita Sarkeesian haters manipulate YouTube, take down new episode. Retrieved June 30, 2013, from <http://www.dailydot.com/entertainment/anita-sarkessian-mra-tropes-women-video/>
- Romano, A. (2013b). Twitter's 'report abuse' button takes out its first innocent victim. Retrieved September 10, 2014, from <http://www.dailydot.com/lifestyle/twitter-report-abuse-transphobia-transphobes-suspended/>
- Rosen, J. (2013). The Delete Squad. Retrieved May 3, 2013, from <http://www.newrepublic.com/article/113045/free-speech-internet-silicon-valley-making-rules>
- Rosenbloom, S. (2014). Dealing with digital cruelty. Retrieved December 12, 2014, from <http://www.nytimes.com/2014/08/24/sunday-review/dealing-with-digital-cruelty.html>
- Rugnetta, M. (2015). If There Are No Pics, Did It Happen? | Idea Channel | PBS Digital Studios. Retrieved August 12, 2015, from <http://www.youtube.com/watch?v=OQxEGLOFyto>
- Rushkoff, D. (2011). Does Facebook really care about you? - CNN.com. Retrieved April 12, 2017, from <http://www.cnn.com/2011/09/22/opinion/rushkoff-facebook-changes/index.html>
- Russell, B. (1961). From the archive: Bertrand Russell on civil disobedience. Retrieved November 14, 2013, from <http://www.newstatesman.com/2013/11/civil-disobedience>

- Sandoval, G. (2013). The end of kindness: weev and the cult of the angry young man. Retrieved September 13, 2013, from <http://www.theverge.com/2013/9/12/4693710/the-end-of-kindness-weev-and-the-cult-of-the-angry-young-man>
- Santayana, G. (1980). The Project Gutenberg eBook of Reason In Common Sense, by George Santayana. (1980 ed.). New York: Project Gutenberg / Dover Publications.
- Sarkeesian, A. (2014). Anita Sarkeesian speaking at XOXO Conference. Retrieved October 7, 2014, from <http://feministfrequency.com/2014/10/07/anita-sarkeesian-speaking-at-xoxo-conference/>
- Sartre, J.-P. (1965). Anti-Semite and Jew. New York: Schocken Books.
- Schatz, B., & Sammon, A. (2016). Facebook's ban on private gun sales is being enforced by a few dedicated users. Retrieved July 12, 2016, from <http://www.motherjones.com/politics/2016/06/gun-sales-facebook-flagged-reported/>
- Schiffman, B. (2008). Tweeterer Takes on Twitter Harassment Policy. Retrieved December 9, 2015, from <http://www.wired.com/2008/05/tweeter-takes-o/>
- Schneier, B. (2013). The Battle for Power on the Internet - Schneier on Security. Retrieved December 3, 2016, from https://www.schneier.com/blog/archives/2013/10/the_battle_for_1.html
- Scholz, T. (2014). Platform Cooperativism vs. the Sharing Economy – Trebor Scholz – Medium. Retrieved September 22, 2016, from <https://medium.com/@trebors/platform-cooperativism-vs-the-sharing-economy-2ea737f1b5ad>
- Schrage, E. (2012, June 1). The Facebook Site Governance Vote | Facebook Newsroom. Retrieved October 30, 2013, from <http://newsroom.fb.com/news/2012/06/the-facebook-site-governance-vote/>
- Schubert, D. (2014). How GamerGate's Right Wing Nutjob Heroes Are Betraying a Hidden Ideological Purge | Zen Of Design. Retrieved September 25, 2014, from <http://www.zenofdesign.com/what-their-right-wing-nutjob-heroes-say-about-gamergates-hidden-ideology/>
- Schwartz, M. (2008). Malwebolence - The World of Web Trolling - The New York Times. Retrieved October 10, 2015, from http://www.nytimes.com/2008/08/03/magazine/03trolls-t.html?pagewanted=all&_r=2
- Scimeca, D. (2013). Using science to reform toxic player behavior in League of Legends. Retrieved June 6, 2015, from <https://arstechnica.com/gaming/2013/05/using-science-to-reform-toxic-player-behavior-in-league-of-legends/>
- Scoble, R. (2011). MySpace's Death Spiral: Insiders Say It's Due To Bets On Los Angeles And Microsoft | Business Insider. Retrieved December 12, 2013, from

- <http://www.businessinsider.com.au/myspaces-death-spiral-insiders-say-its-due-to-bets-on-los-angeles-and-microsoft-2011-3>
- Scott, T. (2013). Actual Facebook Graph Searches. Retrieved July 31, 2013, from <http://actualfacebookgraphsearches.tumblr.com/>
- Seife, C. (2014). The Weird Reasons Why People Make Up False Identities on the Internet. Retrieved June 12, 2015, from <http://www.wired.com/2014/07/virtual-unreality-the-online-sockpuppets-that-trick-us-all/>
- SemperBanU. (2012). Trenches Labs : How TwitterGulag Works Part 1 – Setting Reply Traps. Retrieved November 30, 2013, from <http://thetrenches.us/2012/07/trenches-labs-how-twittergulag-works-part-1-setting-reply-traps/>
- Shackford, S. (2014). GamerGate: How Consumers Are Influencing the Culture Battle - Reason.com. Retrieved January 15, 2014, from <https://reason.com/archives/2014/11/20/gamergate-consumerism-as-culture-war>
- Shaw, A. (2014). The Internet is full of jerks, because the world is full of jerks: What feminist theory teaches us about the Internet. *Communication and Critical/Cultural Studies*, 11, 273–277. doi:10.1080/14791420.2014.926245
- Sippey, M. (2013). Reverting the changes to block functionality | Twitter Blogs. Retrieved December 13, 2013, from <https://blog.twitter.com/2013/reverting-the-changes-to-block-functionality>
- Smith, O. (2013). Facebook terms and conditions: why you don't own your online life. Retrieved December 12, 2013, from <http://www.telegraph.co.uk/technology/social-media/9780565/Facebook-terms-and-conditions-why-you-dont-own-your-online-life.html>
- Smith, P. (2015). This Is What It's Like To Go To Prison For Trolling. Retrieved March 5, 2015, from <http://www.buzzfeed.com/patricksmith/isabella-sorley-john-nimmo-interview>
- Sottek, T. C. (2015). If this is the beginning of the end of Reddit, then Reddit deserves to die. Retrieved June 11, 2015, from <http://www.theverge.com/2015/6/10/8762839/jerks>
- Speed, K. (2007a). Facebook STILL removing breastfeeding pictures. Retrieved March 24, 2015, from <http://bliss-breastfeeding.blogspot.com.au/2007/07/facebook-still-removing-breastfeeding.html>
- Speed, K. (2007b). The breastfeeding picture that Facebook removed from my account. Retrieved September 26, 2015, from <http://bliss-breastfeeding.blogspot.com.au/2007/07/breastfeeding-picture-that-facebook.html>
- Speed, K. (2007c). Facebook STILL harrassing me. Retrieved October 15, 2013, from <http://bliss-breastfeeding.blogspot.com.au/2007/08/facebook-still-harrassing-me.html>

- Speed, K. (2007d). Banned from Facebook permanently----Facebook will not tolerate breastfeeding pictures. Retrieved March 24, 2015, from <http://bliss-breastfeeding.blogspot.com.au/2007/08/banned-from-facebook-permanently.html>
- Spiegel Online. (2012, October 18). Twitter Blocks Neo-Nazi Account in Germany - SPIEGEL ONLINE. Retrieved November 12, 2013, from <http://www.spiegel.de/international/germany/twitter-blocks-neo-nazi-account-in-germany-a-862018.html>
- Stake, R. E. (1995). *The Art of Case Study Research*. Thousand Oaks, California: SAGE Publications.
- Stallman, R. M. (2010). *Free Software, Free Society* (2nd ed.). Boston, MA: Free Software Foundation.
- Stern, R. (2007). *The Gaslight Effect: How to Spot and Survive the Hidden Manipulation Others Use to Control Your Life* (eBook). Harmony Books. Retrieved from <http://www.penguinrandomhouse.com/books/173047/the-gaslight-effect-by-dr-robin-stern/>
- Stilgherrian. (2012, October 22). Data retention's underlying attitude problems. Retrieved January 7, 2014, from <http://www.zdnet.com/data-retentions-underlying-attitude-problems-7000006123/>
- Stilgherrian. (2013). Ignorant Oz politicians prevent meaningful metadata debate. Retrieved January 7, 2014, from <http://www.zdnet.com/au/ignorant-oz-politicians-prevent-meaningful-metadata-debate-7000017117/>
- Sullivan, A. (2009). *The Revolution Will Be Twittered - The Daily Dish*. Retrieved December 9, 2013, from <http://www.theatlantic.com/daily-dish/archive/2009/06/the-revolution-will-be-twittered/200478/>
- Sun-Higginson, S. (2015). *GTFO: Get the F&#% Out*. (S. Sun-Higginson, Ed.) *gtfothemovie.com* (Vol. 1).
- Suzor, N. (2010). *Digital constitutionalism and the role of the rule of law in the governance of virtual communities*. QUT Eprints.
- Sweney, M. (2008). Mums furious as Facebook removes breastfeeding photos. Retrieved October 10, 2013, from <http://www.theguardian.com/media/2008/dec/30/facebook-breastfeeding-ban>
- Swisher, K., & Wagner, K. (2017). Facebook's Mark Zuckerberg wrote a 6,000-word letter addressing fake news and saving the world. Retrieved February 18, 2017, from <https://www.recode.net/2017/2/16/14632726/mark-zuckerberg-facebook-manifesto-fake-news-terrorism>
- Taibi, C. (2013). MySpace Co-Founder Blames Rupert Murdoch For Failure Of Website. Retrieved December 12, 2013, from http://www.huffingtonpost.com/2013/11/18/myspace-cofounder-rupert-murdoch-failure-website_n_4295888.html
- Taylor, C. (2012). Is This Censorship? Facebook Stops Users From Posting "Irrelevant Or Inappropriate" Comments. Retrieved November 12, 2013, from

- <http://techcrunch.com/2012/05/05/facebooks-positive-comment-policy-irrelevant-inappropriate-censorship/>
- The Block Bot. (2013). The Block Bot. Retrieved April 30, 2015, from <http://www.theblockbot.com/>
- The Economist. (2011). Technology and disorder: The BlackBerry riots | The Economist. Retrieved June 5, 2013, from <http://www.economist.com/node/21525976>
- The Iconfactory. (2013). Twiterrific 5.1 - What's New Highlights. Retrieved April 30, 2015, from <https://vimeo.com/59750120>
- The Observer. (2003). Murdoch and China | Media | The Observer. Retrieved December 9, 2013, from <http://www.theguardian.com/media/2003/aug/24/chinathemedia.rupertmurdoch>
- Thomsen, M. (2014). Gamergate And The Unbearable Maleness Of Computers. Retrieved October 28, 2014, from <http://www.forbes.com/sites/michaelthomsen/2014/10/16/gamergate-and-the-unbearable-maleness-of-computers/>
- Tiffany, K. (2017). Verizon is killing Tumblr's fight for net neutrality. Retrieved June 22, 2017, from <https://www.theverge.com/2017/6/21/15816974/verizon-tumblr-net-neutrality-internet-politics-david-karp>
- Tiku, N., & Newton, C. (2015). Twitter CEO: "We suck at dealing with abuse." Retrieved February 4, 2015, from <http://www.theverge.com/2015/2/4/7982099/twitter-ceo-sent-memo-taking-personal-responsibility-for-the>
- Times of Israel Staff. (2015). Hoax blogger posts vile post. Retrieved April 12, 2015, from <http://www.timesofisrael.com/hoax-blogger-posts-vile-post/>
- Todd, A. (2016). GamerGaters Are Mad About Localization Because They Don't Understand It. Retrieved March 12, 2016, from <http://birthmoviesdeath.com/2016/03/04/gamergaters-are-mad-about-localization-because-they-dont-understand-it>
- Totilo, S. (2013). YouTube Channels Crippled By Copyright Claims. Retrieved December 10, 2013, from <http://kotaku.com/youtube-channels-crippled-by-copyright-claims-1480717110>
- Totilo, S. (2014). In recent days I've been asked several times about a possible breach of ethics... Retrieved August 20, 2014, from <http://kotaku.com/in-recent-days-ive-been-asked-several-times-about-a-pos-1624707346>
- True, E. (2014). The gaming journalist who tells on her internet trolls – to their mothers. Retrieved November 30, 2014, from <http://www.theguardian.com/culture/australia-culture-blog/2014/nov/28/alanah-pearce-tells-on-her-internet-trolls-to-their-mothers>
- Tsotsis, A. (2011). Sean Parker On Why Myspace Lost To Facebook | TechCrunch. Retrieved December 12, 2013, from <http://techcrunch.com/2011/06/28/sean-parker-on-why-myspace-lost-to-facebook/>

- Tufekci, Z. (2017). *Twitter and Tear Gas*. London: Yale University Press.
- Tumblr. (n.d.). Adult Content | Tumblr. Retrieved June 15, 2015, from <https://www.tumblr.com/docs/en/nsfw>
- Tumblr Staff. (2013). Tumblr Staff — All, we've heard from a bunch of you who are... Retrieved July 23, 2013, from <https://staff.tumblr.com/post/55906556378/all-weve-heard-from-a-bunch-of-you-who-are>
- Twitter. (2006). Version 1 | Previous Terms of Service. Retrieved September 12, 2015, from https://twitter.com/tos/previous/version_1
- Twitter. (2015). Muting users on Twitter. Retrieved November 17, 2015, from <https://support.twitter.com/articles/20171399>
- U.S. Department of Health and Human Services. (2011). The Surgeon General's Call to Action to Support Breastfeeding. Retrieved May 27, 2015, from <http://www.ncbi.nlm.nih.gov/books/NBK52688/>
- Ullman, E. (1995). Getting close to the machine | Harper's Magazine. Retrieved May 3, 2016, from <http://harpers.org/archive/1995/06/getting-close-to-the-machine/>
- Valverde, M. (1999). The Harms of Sex and the Risks of Breasts: Obscenity and Indecency in Canadian Law. *Social & Legal Studies*, 8(2), 181–197. <http://doi.org/10.1177/096466399900800202>
- Vaughn, K. I. (1978). John Locke and the Labor Theory of Value. *Journal of Libertarian Studies*, 2(4), 311–326.
- Vimeo. (n.d.). Reporting Abuse and Violations on Vimeo. Retrieved August 31, 2015, from <https://vimeo.com/help/faq/watching-videos/reporting-abuse-and-violations#how-does-vimeo-determine-the-difference-between-pornography-and-artistic-or-non-sexual-nudity>
- Wachtler, S., Kaye, J. S., Hancock, S. F., Bellacosa, J. W., Titone, V. J., & Simons, R. D. THE PEOPLE & C., RESPONDENT, v. RAMONA SANTORELLI AND MARY LOU SCHLOSS, APPELLANTS, ET AL., DEFENDANTS. (S. Wachtler, J. S. Kaye, S. F. Hancock, J. W. Bellacosa, V. J. Titone, & R. D. Simons, Eds.), Legal Information Institute 875–232– 601 (July 7, 1992). New York.
- Wagner, Kyle. (2014). The Future Of The Culture Wars Is Here, And It's Gamergate. Retrieved October 14, 2014, from <http://deadspin.com/the-future-of-the-culture-wars-is-here-and-its-gamerga-1646145844>
- Waldman, A. (2008a). Twitter refuses to uphold Terms of Service. Retrieved September 12, 2015, from <http://arielwaldman.com/2008/05/22/twitter-refuses-to-uphold-terms-of-service/>
- Waldman, A. (2008b). Twitter responds. Retrieved September 12, 2015, from <http://arielwaldman.com/2008/05/23/twitter-responds/>
- Walsh, J. (2007). Men who hate women on the Web. Retrieved September 12, 2015, from <http://www.salon.com/2007/03/31/sierra/>

- Walters, C. (2009). Facebook's New Terms Of Service: 'We Can Do Anything We Want With Your Content. Forever.'. Retrieved January 7, 2014, from <http://consumerist.com/2009/02/15/facebooks-new-terms-of-service-we-can-do-anything-we-want-with-your-content-forever/>
- Warren, C. (2013). Tumblr Starts Hiding Porn From Search and Tag Pages. Retrieved July 24, 2013, from <http://mashable.com/2013/07/19/tumblr-nsfw-policy-changes/>
- Warren, E., Black, H., Douglas, W. O., Clark, T. C., Harlan, J. M., II, Brennan, W. J., Jr, et al. *Jacobellis v. Ohio* | US Law | LII / Legal Information Institute. (E. Warren, H. Black, W. O. Douglas, T. C. Clark, J. M. Harlan II, W. J. Brennan Jr, et al., Eds.), Legal Information Institute (June 22, 1964). Washington D.C.
- Wauters, R. (2011). Illogical? William Shatner Gets Booted From Google+ (Update: He's Back). Retrieved November 24, 2014, from <http://social.techcrunch.com/2011/07/18/illogical-william-shatner-gets-booted-from-google/>
- West, L. (2014). A Man Is Facing Jail for Harassing Women Online, and It's a Big Deal. Retrieved January 9, 2014, from <http://jezebel.com/a-man-is-facing-jail-for-harassing-women-online-and-it-1497322836>
- West, L. (2015). How Reddit's Ellen Pao survived one of "the largest trolling attacks in history." Retrieved June 6, 2017, from <http://www.theguardian.com/lifeandstyle/2015/dec/22/reddit-ellen-pao-trolling-revenge-porn-ceo-internet-misogyny>
- Williams, E. (2008a). Ev Williams on Twitter: "Note: Before joining a mob, you might want to check if everything they're saying/assuming is true.." Retrieved September 12, 2015, from <https://twitter.com/ev/statuses/817807577>
- Williams, E. (2008b). Ev Williams on Twitter: "Our response to the Ariel situation: <http://tinyurl.com/5b2jul>." Retrieved September 12, 2015, from <https://twitter.com/ev/status/817907116>
- Wilson, J. (2011). Playing with politics: Political fans and Twitter faking in post-broadcast democracy. *Convergence: the International Journal of Research Into New Media Technologies*, 17(4), 445–461. <http://doi.org/10.1177/1354856511414348>
- Wilson, M. (2017). Google still hasn't given up on Google+ and Topics is the latest attempt to keep people on the service. Retrieved June 2, 2017, from <https://betanews.com/2017/04/28/google-plus-topics/>
- Wingfield, N. (2014). Feminist Critics of Video Games Facing Threats in "GamerGate" Campaign. Retrieved October 16, 2014, from <http://www.nytimes.com/2014/10/16/technology/gamergate-women-video-game-threats-anita-sarkeesian.html>
- Wired Staff. (1994). Meme, Counter-meme. Retrieved November 10, 2016, from <https://www.wired.com/1994/10/godwin-if-2/>

- Wofford, T. (2014). Is GamerGate About Media Ethics or Harassing Women? Harassment, the Data Shows. Retrieved October 25, 2014, from <http://www.newsweek.com/gamergate-about-media-ethics-or-harassing-women-harassment-data-show-279736>
- Women Action & the Media. (2013). Open Letter to Facebook | WAM! Retrieved May 21, 2013, from <http://www.womenactionmedia.org/facebookaction/open-letter-to-facebook/>
- Wong, J. C. (2016). Mark Zuckerberg accused of abusing power after Facebook deletes “napalm girl” post. Retrieved September 9, 2016, from <http://www.theguardian.com/technology/2016/sep/08/facebook-mark-zuckerberg-napalm-girl-photo-vietnam-war>
- Wong, J. C. And O. Solon (2018). "Google to shut down Google+ after failing to disclose user data leak". Retrieved December 29, 2018 from <https://www.theguardian.com/technology/2018/oct/08/google-plus-security-breach-wall-street-journal>
- Working to Halt Online Abuse. (2012). Comparison Statistics 2000-2012. Retrieved January 10, 2014, from <http://www.haltabuse.org/resources/stats/index.shtml>
- Wortham, J. (2009). Facebook Won't Budge on Breastfeeding Photos. Retrieved October 13, 2013, from <http://bits.blogs.nytimes.com/2009/01/02/breastfeeding-facebook-photos/>
- Yin, R. K. (1994). Case Study Research: Design and methods (2nd ed.). Thousand Oaks, CA: Sage Publishing.
- York, J. C. (2010). Facebook: “No Palestinian Pages.” Retrieved January 10, 2014, from <http://jilliancyork.com/2010/07/25/facebook-no-palestinian-pages/>
- York, J. C. (2014). @jilliancyork: “@thelindywest Private companies that aim to be the new public sphere should not be in the business of censorship..” Retrieved January 11, 2014, from <https://twitter.com/jilliancyork/status/421898743894384640>
- Youtube. (n.d.). YouTube Copyright Center - YouTube. Retrieved December 6, 2013, from <https://www.youtube.com/yt/copyright/>
- Zavadski, K. (2015). “Terrorist” Troll Pretended to Be ISIS, White Supremacist, and Jewish Lawyer. Retrieved September 11, 2015, from <http://www.thedailybeast.com/articles/2015/09/11/terrorist-troll-pretended-to-be-isis-white-supremacist-and-jewish-lawyer.html>
- Zeilinger, J. (2015). Here's What the Free the Nipple Movement Has Really Accomplished. Retrieved August 21, 2015, from <http://mic.com/articles/124146/here-s-what-the-free-the-nipple-movement-has-really-accomplished>
- Zeldman, J. (2008). A Tweet Too Far - Zeldman on Web & Interaction Design. Retrieved September 14, 2015, from <http://www.zeldman.com/2008/05/22/a-tweet-too-far/>

- Zhang, X. (2015). Sharing block lists to help make Twitter safer | Twitter Blogs. Retrieved June 11, 2015, from <https://blog.twitter.com/2015/sharing-block-lists-to-help-make-twitter-safer>
- Zimmer, M. (2010). Zuckerberg's Remarks Aren't Surprising, Nor New, Nor True. Retrieved October 15, 2015, from <http://www.michaelzimmer.org/2010/01/12/zuckerbergs-remarks-arent-surprising-nor-new-nor-true/>
- Zuckerberg, M. (2010). D8: Facebook CEO Mark Zuckerberg Full-Length Video | Video. (K. Swisher & W. Mossberg, Eds.) *Allthingsd.com*. All Things D.
- Zuckerberg, M. (2017). Building Global Community. Retrieved February 17, 2017, from <https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634/>