Performativity 2.0

By Phoebe Zeller

A thesis submitted to Victoria University of Wellington in fulfilment of the requirements for the degree of Masters of Design Innovation

University of Wellington

2018

Abstract

The practice of quantifying users' performativity online by collecting personal data has become commonplace. This thesis explores performativity in a Web 2.0 climate relating to social media and online data with the intention to better understand this relationship. This exploration of internet culture is conducted through the perspective of a designer who is concerned with how people interact with social media, how this affects us, and the data footprint we leave behind. The theory of performativity is used to understand the ways in which users' curate their online image. Social media can capture a user's performance through the collection of data. This representational data must be considered in the context in which it was created. This thesis discusses how assumptions are made on users by analysis of their data and the ways in which this is misrepresentational. These themes are expanded upon by the creation and design of artifacts.

Acknowledgements

I'd like to give thanks to everyone who helped me along the way, but special mention to:

The staff at the Faculty of Architecture and Design, VUW including Catherine Caudwell and Tom White. Most of all Walter Langelaar, without your unwavering encouragement and advice this thesis wouldn't have made it to completion.

To my Family, Steve, Donna and Meghan (Smeagol) for the support and compassion.

To the fellow Masters' students but especially Bryan, Chris, Ivy, Sophie, Stef and alternate Stef. Your willingness to provide help, feedback and friendship has been invaluable.

To my friends' whose synergistic support, motivation and cute animal pictures made this thesis less of a struggle, particularly Tobias, Rose, Jordan, Tom and Joseph.

Preface

This thesis is written by a child of the internet. I was hooked up to the internet at a young age, and from then I have been fascinated. One of the first things I remember being impressed by was the online version of the Encyclopaedia Britannica. This encyclopaedia is not particularly interesting because it was just a publication of a book, but one published online meaning it was more accessible than ever. From here on I learnt about things that people were making on the internet, for people on the internet. A wealth of information and snapshots of different cultures, all accessible for free. By the time I was 13 or so I was soaking up so much of what the internet had to offer, from online shopping to the Dark Web archives. And I have watched as others around me grew up with the internet as well, and how our habits and behaviours have changed in order to make room for the internet as a presence in our lives.

I consider the internet to be a presence in my life, a character that I interact with daily. I share my thoughts and feelings with, I let them into my most intimate spaces, I cry to them and I celebrate with them. While the internet knows about the details of my life, it only knows about the things I choose to share with it. I can curate myself in exactly the way I want to portray myself.

For the internet, I perform my best self. Through all of this I create data, data that is then used to represent me.

Contents

Abstract	III	
Preface	V	
Introduction	1	
Literature review		
Web 2.0	4	
Big Data and the Privacy Paradox	6	
Expression of self and performance	8	
The effect on our data	10	
Methods	15	
The in breath	15	
The out breath	16	
Responding to Data Selfie	19	
My Data Selfie	20	
Projects	35	
Intro	35	
Embroidery:		
Imagining others complexly	37	
Popping Hashtags	43	
Haemorrhaging data	47	
Discussion	55	
Conclusion	59	
Works Cited		

Introduction

The stage for the self is now digital. Currently, most interpersonal communication on the internet is conducted through social media. It is on social media platforms that we express ourselves, connect with others and build communities. Self-expression online can be publicised through a wide variety of digital media, for example, photos, text posts and video clips. These forms of self-expression are captured and stored by the internet in the form of data. Through our posts online we create a data footprint that can be quantified and analysed by machine learning algorithms. This thesis takes some of the issues around data creation and its analysis and discusses them in relation to the theory of performativity. Through the application of this theory, we can better understand the self-curation behaviour exhibited on social media. The ways in which data is created are dictated by how users perform their sense of self. This research also takes into discusses if data alone can be an accurate representation of a person's self-image. Through applying the theory of performativity to the expression of the self online, this thesis answers the question: How can design illustrate the ways in which users' personal representational data is affected by their performative activity on social media? This research first considers the foundational literature around the topics of performativity, Big Data and privacy. Creative responses that highlight issues around these themes are then also examined. In this section, the work of Erving Goffman is considered when discussing the behaviours people express through their interaction with social media. The perspective on self-performance from this literature is then applied to observations on the large amounts of data that users upload daily and privacy concerns around this. From this background research, a response is designed in the form of abstracted creative outputs that respond critically to the current internet climate.

Literature review

You can tell a lot about a person from how they use the internet. In Lanier's Manifesto (2010), the internet is talked about as an extension of people and culture. The most integral aspect to consider when discussing this technology is how it changes people (Lanier, 2010). This literature review will explore how the internet and the behaviours that facilitate connectivity have evolved during the era of Web 2.0. It goes on to explore how data on the internet is being collected ubiquitously and stored in large amounts. Data collection raises privacy concerns which lead to further inquiries into users' behaviour online. This review will cover how the Internet culture we surround ourselves with dictates the amount of personal information we share as well as some of the contributing factors to how we select the media we upload. Online behaviour surrounding this is explored through the considerations of the theory of performativity. It is human nature to perform in accordance with the culture a person surrounds themselves in and through these considerations, we can achieve a greater understanding of the internet climate we see today. By understanding the present internet climate, data and users' performance we can make more informed designed decisions about the internet we want to interact with.

Web 2.0

Integral to understanding the current internet climate is understanding the paradigm shift that came with Web 2.0. The phrase Web 2.0 was first mentioned by Darcy DiNucci in 1999 (DiNucci, 1999). The popularity of this phrase took off after a brainstorming session with Tim O'Reilly went viral, to use 2.0 phrasing. Describing the web as having moved into a second generation meant considering participatory and communicative web interactions as an integral part of how the internet was to be used going forward (O'Reilly, 2007). The uptake of this phrase was an acknowledgment that the internet had moved into a new era. The era of Web 2.0 shifted the focus away from mass instant publishing, and to a web that is hyperconnected, collaborative and interactive. Parts of the internet that we pay no mind to today where being pioneered during this time: hashtags, open source software, dynamic web design, social networking, tracking and prediction markets (O'Reilly, 2007).

"Web 1.0 --> Web 2.0 DoubleClick --> Google AdSense Ofoto --> Flickr Akamai --> BitTorrent mp3.com --> Napster Britannica Online --> Wikipedia personal websites --> blogging evite --> upcoming.org and EVDB domain name speculation --> search engine optimization page views --> cost per click screen scraping --> web services publishing --> participation content management systems --> wikis directories (taxonomy) --> tagging ("folksonomy") stickiness --> syndication " (O'Reilly, 2007, p.2)

It is within this understanding of Web 2.0 that we find social media and personal blogs. With this formative period of the internet, we started seeing platforms that facilitated self-publishing to the internet (Lovink, 2011). Anyone with access to the internet can now selfpublish in whatever content they want and however they see fit. Lovink (2011) states three distinguishing features of Web 2.0 as ease of use, the facilitation and encouragement of social behaviour and the ability to publish content of any type for free (p. 5). The idyllic nature of this collaborative and free internet based on free speech is a hopeful and liberating notion. The hyper-connected Web 2.0 lends itself well to building rich online communities, often found around niche topics. The communities of the internet are not without flaw and do not exist entirely harmoniously. While the hopes of early Web 2.0 enthusiasts are not completely dashed in the current web climate, the "Techno-Libertarian utopia was a strong meme" (Lovink, 2011, p. 39). While the utopia has not been crushed under bureaucracy, it has certainly been reined in somewhat. People are still able to express themselves online, but the internet currently is far from idyllic.

The freedom allowed by Web 2.0 led to a lot of wonderful creations that made their way to the internet, be this communities, movements or manifestations of other sorts. A large factor in this was that as communities appeared in forums and on social media, people were encouraged to share personal experiences (Boyd, 2016). The sharing of personal information creates a strong foundation for communities to be born from, even if this information is associated with a screen name as opposed to someone's personal identity. We can find this evident in gaming communities who use almost exclusively screen names but still have a strong sense of community (Lovink, 2011). Personal and meaningful interactions can still occur even without the disclosure of real names. This considered, no online community is created freely. In order to have something meaningful, like a dimensional community, effort and real work must be put in. For example, the rise of Web 2.0 saw more and more images being uploaded to the web. We can take these images as units of information that we share for the exchange of attention and the views of others. This exchange is real work, the work of watching and being watched (Rushton, 2012). It is this exchange that communities are built on.

The internet itself can be considered a character within the online communities, and this is due to the design of the internet. When we consider the internet and the computers we use to access the internet, we talk about them in terms of personhood. We discuss its memory (disk space), its eves and ears (webcams) and the forms that we fill out are said to 'ask' us questions (Lanier, 2010). By considering the internet in terms of a person, we feel a closer connection to it, allowing us to share more, connect more and in turn, we are further shaped by our own internet habits. A behaviour observed on the internet that has become more prevalent in recent years is the use of self-branding with the intention of becoming a micro-celebrity within your internet community (Page, 2012). These practices focus on the construction of an identity with the intention of it being consumed by the masses. This behaviour is like existing in a community, but that community is centralised around whoever has chosen to connect with you. It is this focus on the self which is present in social media behaviour that will be continued to be explored in the coming sections.

Big Data and the Privacy Paradox

With the hyper-connectivity of the Web 2.0 generation came social media networks. These platforms facilitated aspects of Web 2.0 such as the sharing of images and interpersonal connection. Presently, social media is more prevalent than ever, with the largest user base being Facebook with approximately 1871 million active users as of 2017 (Chaffey). Should the current trends in social media uptake continue, platforms such as Facebook will be unlikely to show a loss of popularity. Social media sites encourage social behaviour, which leads to a large amount of self-published information. The information shared facilitates community and connectivity, but this self-published information and personal data raise issues around privacy. The challenges around publicised information are complex and intertwined with the current social-technical norms (Boyd, 2016). This section discusses the issues around data and privacy but keeps these challenges in consideration. A result of the uptake of sharing information and the publication of personal data is the creation of Big Data. Big data can be defined as "the storage and analysis of large and or complex data sets using a series of techniques including, but not limited to: NoSQL, MapReduce and machine learning." (Ward & Barker, 2013, p. 2). The scale of Big Data is such that it is impossible for a person to comprehend how large it is. The data being collected comes in the forms of text, sensor data, audio, video and click streams to name a few. This data can be used to make highly specific assumptions and predictions (Kosinski, Stillwell, & Graepel, 2013). Our actions can only be predicted because the data sets our personal data is being compared to are so large.

Despite knowing the risks of publishing personal data, and understanding the importance of internet security, people continue to post large amounts of highly personal data daily (Taddicken, 2014). It is not a requirement of the social internet to publish personal data, many people have rewarding online experiences by creating conversation around information that is not self-centric (Naaman, Boase, & Lai, 2010). Our current web climate encourages the sharing of social data, but the user experience of the internet is mostly unaffected by the user's choice to share personal posts as opposed to more information-based publications.

Despite the vast scope of social media, we choose to connect with friends and people we find interesting. Because of this, the content presented to us aligns with our interests and opinions, with few variations (Boyd, 2016). It is part of human nature to seek out 'tribes', and this reflects in our social media use. Taddicken's study (2014) suggests that our behaviour online is greatly dictated by the actions of our peers. For example, if a user's friends and family share many photos online, they are likely to share similar content. While this means that we are presented with content that we want to see, we also perpetuate the habits of our peers that lead to the masses of data being uploaded every second.

The existence of Big Data has resulted in data-based privacy considerations which are encouraged to be considered when posting on social media. The act of disclosing private information despite understanding the security implications is called the Privacy Paradox (Taddicken, 2014). This paradox can be broken down to when privacy -"selective control of access to the self" (Altman, 1975) - is not considered when posting online despite any/all privacy concerns an individual may have. Andrejevic (2014) argues that people have the impression that they are not able to control what data is being collected on them. There are many possible factors which play into this sense of not being in control of your data. The most prominent factor is the scale difference between the data that which is collected from a single user and that which makes up Big Data. It is worth noting, to have meaningful human interaction a degree of self-disclosure is required. On the internet, this could mean posting an original thought or unique moment in your life through whatever means the user sees fit. The majority of users will disclose these thoughts and moments even if they are concerned about their privacy on the internet (Taddicken, 2014). When we consider a user's inability to imagine the scope of Big Data, the data produced by individual users seams less significant. This feeling of insignificance is arguably a contributing factor to some elements of the Privacy Paradox. These inconsistencies are even more common in groups of people that see others disclosing similar information (Boyd, 2016). We follow the social conventions of our peers as we see them on social media, we know that the communities that we are part of online dislike change and outsiders, so we produce an image of ourselves that is like the content that we see. If it is the case that the content that we see is highly populated with posts containing personal data, then we too will share personal information ourselves (Boyd, 2016). It is that personal information that presents the most privacy concerns, but it is the most rewarding to share.

The personification of Big Data by describing it as 'the internet's memory' allows for its vast scale to be considered in relation to the human condition. This makes the topic more approachable as well as allowing for more abstract interpretations of data. Van Bree (2017) states that by utilising the technological advances that the internet has provided, human memory has been externalised for the purpose of enhancing our collective intelligence. Van Bree goes on to suggest that people are becoming more aware that through the internet we are living with an external memory which is incapable of forgetting. Van Bree refers to this as Digital Hyperthymesia. Van Bree goes on to state that by having such an extensive record of "digital memories" (p.31) people are going to begin to consider their own identity as something quantifiable. This notion conversely has not resulted in a major drop in personal data being made publicly available (Taddicken, 2014) by internet users. In

her online essay, Carmen Hermosillo AKA humdog (1994) describes personal expression within 'electronic communities' (para. 10) to have a permanent record which becomes open to scrutiny despite "maintaining an illusion of transience"(para. 19). Recently this transience illusion has begun to break down and more people are becoming aware of the importance of data privacy.

The understanding of people's behaviour on the internet in relation to their data is important for understanding the data that people produce. Another key element to understanding the data that is being collected into Big Data is understanding how people present themselves. Any data that is being collected on people has been self-expressed, the next section will discuss factors of that expression of self.

Expression of self and performance

Data which is uploaded to the internet creates a data profile that poses privacy considerations. This knowledge raises the question of how truly personal the data is that we are uploading. Social media has its origins in the desire to share with others and be social. One method of understanding sociality is through the theory of performativity. Performativity describes the way in which we give select access to aspects of ourselves. Through this theory, we can better understand how people curate their online persona and by extension their data profile.

Theories of the performed self and performativity have been around longer than modern computing and the internet. Erving Goffman states in his book *The presentation of self in everyday life* (1959) that humans in a social setting are always performing. This line of thinking is derived from the perspective of dramaturgy, a school of thought heavily based on stage performance ideals. It can be said our actions and mannerisms can fluidly change and adapt depending on who is watching, as well as the social situations we place ourselves in. Goffman's book suggests that based on the role we are expected to uphold within particular social situations, we adapt to be the most suited versions of ourselves for that time and place. Goffman goes on to liken aspects of the human condition to stage plays, stating that our lives have a backstage as well as the front stage that is on display for people to see. This clear difference between the front and back stages of our lives would be distinguished by the presence of "impression management" (Goffman, 1959). We are not doing ourselves an injustice when we mute the undesirable or inappropriate aspects of our personality in certain social settings; we are still presenting a version of our true self. The front stage performance is merely a curated version of the self, not a falsity. The social performance gives us tools to fit in with our peers and adapt to different social situations. We embarrass ourselves and misplay the role we desire to play all the time. We often say the wrong things or act in a way that is not totally appropriate. This is also part of being human, even the most

sociable person can misstep, but it is not the end of the performance.

Despite the age of Goffman's work, many of his findings are still applicable today. Goffman's ideas on the performance are, though, somewhat rigid compared to modern standards. The front stage and backstage are now considered to be more of a spectrum than a duality. Our performance is heavily based on what is expected of us (Butler, 1990). In a typical social situation, if someone is trying to carry out a performance that is not their own, it will be noticed by the observers. Butler considers the performance of self to be heavily influenced by the culture that we surround ourselves in. Our performance is socially constructed, and from these constructs, we can approximate our true self (Butler, 1993). While Butler's primary interest in performativity is when considering gender, performativity can be considered in a Human-Computer Interaction and social media use context. Online the ability to self-censor is magnified. In a classic social situation, if someone is trying to carry out a "performance" that is not their own, people will notice. In an online environment, this is not the case as we have then absolute control to censor parts of our life that we deem unappealing. Selfcensorship is mentioned by Miller in his paper (1995), illustrating how self-performance can be applied to internet behaviour. However, due to the era of the internet in which Miller's writing is situated, the points made fall flat when compared to the current hyper-social Web 2.0. His arguments are valid but are mainly discussing performance through email, whereas presently we have far more variety in the ways in which we present ourselves.

There have been misunderstandings of how performativity interacts with the internet, notably since the beginning of Web 2.0. Mark Zuckerberg was quoted in 2010 saying 'Having two identities for yourself is an example of a lack of integrity' when discussing impression management with different audiences (Zimmer, 2010). This is either a misunderstanding of how self-performance works in a social setting, or a belief that social media can break down our social front entirely. In contrast, we could consider the different personalities portrayed through social media as a relief from our everyday life through forms of escapism (Turkle, 2011). Though we can be considered to escape our current situation by engaging with the internet. Turkle argues that our time spent online is not a meaningless diversion. By curating our online persona, we are able to use our virtual experience for personal growth. Taking the example of gamer communities and their use of aliases, they can create meaningful interpersonal engagement without giving out their true identity. Through this anonymity, they can be said to become distanced from their true self and exist online only behind an endless series of masks which they could change as they see fit (Lovink, 2011). By living behind masks, members of such a community cannot foil the privacy flaws of the modern internet, but further disregard them.

As we go through the motions of expressing ourselves online, curating

our image and showing others what we choose, we are proving that we are anything but passive consumers of a product. We no longer live in the "society of the spectacle" (Rushton, 2012, p. 87) but excite interest in others in the subject of ourselves. In this way, our performance is the principal object of production, an object that we impart immense value upon (Rushton, 2012). Self-branding and the construction of an identity place great emphasis on creating a cohesive self-image. The intention of this image is to be consumed by the masses (Page, 2012). The conscious act of curating and managing a brand that is based on one's personality is generally for the purpose of achieving micro-celebrity status and obtaining social or economic benefit (Page, 2012). It is becoming more important for creatives hoping to make themselves known through the internet to laboriously make their online presence curated and in line with their work (Erica Scourti, 2017). We put effort and work into our online performances, and because of this, we can create great communities of people. This work though is encouraged by a desire to be observed, which stems from a desire to be valued and appreciated by others in our community. Following the social conventions that we see in our communities, we produce an image of ourselves that is like the content that we see. If it is the case that the content that we see is full of personal data, then we too will show that personal side of ourselves (Boyd, 2016). It is that personal side that is the most valuable, but that is the most satisfying to share.

The effect on our data

The present-day internet with its hyper-connectivity, perfect memory and large amounts of stored data has changed how we present ourselves online. We consider our peers by how they perform on social media, and social media considers us in relation to our peers. This perspective and the assumptions data makes of us creates the climate of the internet presently. This climate is explored by creatives to get a better understanding of how these elements connect. Boyd (2016) states that despite any online behaviour we observe, we still try to make attempts at expressing ourselves. We have the desire to share as a means to have better bonds with others. We observe others online and through that act of watching we project our assumptions (Boyd, 2016). These assumptions can be critiqued and drawn attention through creative projects. It is these critical projects that will be explored in this section. Using Big Data, we have gotten very good at approximating details about a person from the data they produce. An approximation of a person's personality and demographics is only possible in any detail by utilising Big Data. A relatively small amount of data is needed to profile a single person because it is checked against similar data to make personalised generalisations (Kosinski et al., 2013). Using a tool created by Cambridge University called Apply Magic Sauce you can see the assumptions that your data is making about you (Kielczewski, 2018). Apply Magic Sauce, an open-source application programming interface

(from here onward referred to as an API) that can look at any data that you give it and extrapolate what qualities that data indicates. This API allows users to glimpse the ways in which their data represents them, and the assumptions being made of them. Using this API, users also can determine inconsistencies and flaws in how data is represented in relation to personhood. Apply Magic Sauce brings issues around data accurately representing us online to consideration. Without awareness around these issues, we are likely to repeat habits that compromise our data privacy. Designer Tanne van Bree's response (2017) to the internet's perfect memory is a program that can approximate images, to imply what an imperfect memory is. It uses computer vision to look at an image, and then makes a composite image as a response. This composite image is made up of the algorithmically determined subject matter, compiled of images scraped from the internet. This design raises an important consideration when designing for the internet, how something will be recalled. The idea of imaging technology diminishing over time follows a similar reasoning Mark Fisher uses in his book (2014). The internet's perfect memory is like a .mp3 file in that it will always play back the same way, giving the impression of timelessness. Vinyl records in contrast crackle which "makes us aware that we are listening to a time that is out of joint; it won't allow us to fall into the illusion of presence" (Fisher, 2014, p. 21). Vinyl is more honest, as we know we are observing the past in the moment of the crackle. On social media, the content we observe is always from the past, even by just a few seconds. Posts are always being preserved and reflected on. While interacting with social media, people look at the past and deem they are missing out on what is happening in the now (Macmillan, 2017). Through this, creative and abstracted viewpoints are afforded new perspectives on social media phenomena.

The way we use social media does not drastically vary from day-today. The artwork 'You like my like of your like of my status' (Grosser, 2016) highlights the cyclical nature of unconstructive data production. This artwork consists of a generated text that is being read aloud by a synthesised voice. The text describes the recurring act of liking status updates and then liking these updates again. The artwork becomes repetitive after a short amount of time, but despite this, there are enough subtle differences that each moment of the artworks performance is slightly different. This work prompts the viewer to consider how we create data that is like our peers' (Boyd, 2016). Grosser's work shows how shallow a lot of the data in Big Data is. This is an analogy to how current social media contains a lot of vapid data that is unable to capture the human experience but rather drown out mundane life (Lovink, 2011). These considerations are in contrast to the idealised views of the hyper-connected internet present in the early Web 2.0 era.

Automated work like that of Grosser is a medium that is well suited to providing critique towards how people perform online. One such automated medium is the Twitter bot. The phrase Twitter bot comes from Twitter robot, meaning a Twitter account that will post tweets according to its programming. Erica Scourti (2017) states that because her Twitter bot is fed her diary as input text and then posts automatically, then perhaps it is more truthful than her. Her bot is run through a filter to make sure it does not say anything rude or offensive but otherwise is created from a preserved form of Scourti. It produces images and text with each tweet that can contain personal or sensitive messages, messages that cannot be edited or polished by the creator before publication (Erica Scourti, 2017). Because of this, the Bot's performance is personal but unwavering, allowing for a critical demonstration of the personal performance and self-branding behaviours (Page, 2012) prevalent on Twitter.

Though the aforementioned abstract and artistic responses to the digital performance we are afforded a critical eye on the subject. Digital performance and the data surrounding it is a subject which can be further investigated as some areas are outside of the scope of this literature review. It is hopeful that through artistic and qualitative investigations of online phenomena further innovative theoretical approaches will arise. Through this innovation, we will be better equipped to make informed design decisions when creating content for the internet.

Methods

This thesis makes use of methodologies that involve a deep reflective process as well as valuing the expressive artistic responses. The methodologies used in this thesis can be considered in terms of the description of the creative process used in the comic "Creativity is like breathing" (Inman, 2016). For effective and inventive creative practice, one must "breathe in" before an attempt can be made at creating. This breath in is the act of observing and taking in the information around you. Breathing in is just as important as breathing out, and vice versa. I chose Inman to be the influence to this school of thought because I found it natively on the web during my browsing. Inman's notion has had an impact in the way I view my work and my personal creative process more than the other texts cited.

The in breath

The importance of "tacit knowledge" is highlighted by (Friedman, 2008) by stating its integral role in "embodied individual and social knowledge provides the existential foundation of all activities, including intellectual inquiry." (p.157). This notion is key to the personal approach this thesis takes. Utilising tacit knowledge of the internet allows for design considerations that are relevant to the context in which the design is intended to be situated. This notion comes from a constructivist background which recognises that one's own background shapes interpretation (Creswell, 2009). As someone who has been integrated with internet culture since childhood, the implicit knowledge gained during this time has shaped this research greatly. Aspects of culture already understood were backed up by literature. Archer (1995) indicates that in order to conduct good research it is important to have a consolidated understanding of primary sources and the provenance of key ideas within the relevant field. The intent and standpoint of these primary sources which must be considered to fully detain the value of the work. Archer goes on to state how Action Research, the practical action intended to systematically generate communicable knowledge, and a transparent and honest depiction of actions can be a highly useful tool when exploring a proposition. The practice of design as an action method while actively acknowledging the tacit knowledge enabling practice (Friedman, 2008) is the basis of this thesis's use of research through design as a method (Frankel & Racine, 2010)

The out breath

The expressions present in this thesis are made in consideration to Steve Brown's writing on abstract experimentalism (2012). Experimentalism in the context of this literature is a method "of inventing or creating new forms in which the world is deemed able to 'speak'" (Brown, 2012, p. 4). This ideology draws from reductionist ideas but emphasises the "desire to uncover the new" (p. 10) by artistic means.

Responding to Data Selfie

In this section, I respond to the interpretation of my data through the extension Data Selfie. Data Selfie is a free open source Firefox and Chrome extension designed by DATA X. It utilises natural language processing and machine learning algorithms from IBM Watson and the University of Cambridge to explore how data is represented online. As this is a personal reflection, my response is written using a personal tone and in the first person.

The extension can track a variety of actions. Looking, typing, clicking, and liking. This extension does not interact with the Facebook API but instead mimics the methods that Facebook uses to gather data. Data Selfie does so in such a way that you can download and observe your data in a raw format. This extension does not store any of your data on a server, so if you choose to delete the data that Data Selfie has collected on you it is deleted permanently. This is a very humanitarian way of hacking your own data, it gives control back to the user. It is empowering to be able to track yourself much like a corporation such as Facebook does, and even more so because the designers allow you to interact with your data and take ownership of it. All of the design decisions for this extension enforce this ethos. Their aims are to "provide a personal perspective on data mining, predictive analytics and our online data identity- including inferred information" (DATA X, 2018, para. 3) as well as illustrate how your data can make assumptions about you. Coming to terms with how much power and influence your data profile has over you is a personal experience, so Data Selfie strives to be as transparent as possible. This is communicated through the visual language of the extension as well as the explanation features of the extension. The aesthetics of the extension are used to further reinforce the idea that we are being given an inside look at data that is not normally made available to us. This hacker aesthetic is clean and informative, allowing the users to come to their own conclusions about the information being displayed.

My Data Selfie

My interaction with this extension began on the 2nd of May 2017. This extension has been installed on my main personal desktop computer which resides in my bedroom at home.

The data was retrieved for the purpose of analysis on the 13th of June 2018.

Since installation, I have spent 89.63 hours on Facebook. This is not a representation of the total hours I have spent on Facebook over the past year as I mainly access Facebook from my phone. Not all the hours logged were hours spent actively engaging with Facebook. A large portion of the hours logged is from having Facebook open on a secondary monitor. With Facebook open, the extension would read me as looking at whatever was displayed on the page, even if I was not paying active attention. I do not consider this a skewing of the results as I was still witnessing the contents of the page, if somewhat subconsciously.

From the data collected over the 89 hours, Data Selfie has created this summary of me by analysing my data. This summary is a broad statement created using predictive analytics and machine learning.

Figure 1: Data Selfie: Summary

My summary: You're a laid-back, liberal female who eats out frequently and doesn't prefer style when buying clothes and is more satisfied in life than most.

This is a poorly constructed sentence as it is not a core focus of this extension to produce natural sounding text. Anything on this extension that does not have a level of polish to it I assume is created algorithmically because any man-made elements are highly considered. This, therefore, is created using natural language processing as a response to my data. Natural language processing techniques are used in an attempt to algorithmically create text that sounds 'natural'. This is done by stringing together parts of speech in a way that best mimics how a person would construct a sentence. As this summary results in a poorly constructed statement, I already have the impression that my data does not have a full understanding of me. This attempt of a sentence demonstrates how the assumptions of data can be interesting and show insight but are imperfect at representing complexities. It also makes generalisations that are difficult to disagree with, except I do eat out frequently.

Top friends (10 of 361)	Top pages (10 of 224)	
Time spent (in sec) on friends' posts	Time spent (in sec) on pages' posts	
15328 Fisher Alfredo Rivero Connell	21752 Bad Memes for Suffering Victoria University Teens	
4311 Micheal Baron	11770 i-D	
3146 Rose Lastovicka	3368 I fucking love science	
2764 Stefan Peacock	2574 Victoria University of Wellington - Architecture and Design	
2732 Sacha Judd	2116 Corgi Overload	
2583 Edward Taylor	1575 Game of Thrones Memes	
1639 Alex MacKay	1386 Futurism	
1470 Aimee Rose Bennett	1350 Ninjaflower	
1269 Fintan Nelson	1345 Skyrim Guard	
1162 Daniel Peres Dos Santos	1242 Victoria University of Wellington	
۲ ۲	۲ (۲ (۲ (۲ (۲ (۲ (۲ (۲ (۲ (۲ (
updated: now	updated: now	
Figure 2: Data Selfie: Top Friends	Figure 3: Data Selfie: Top Pages	
Top Likes (10 of 19)	<u>ـ</u>	
Likes for posts, photos or videos		
2 Rose Lastovicka		
2 Ninjaflower		
1 Gerrit Jacobus van Rooyen		
1 Shiv Andrews		
1 Overwatch		
1 NZ swing voters against dogmatic par	ty affiliated memes	
1 Phoebe Zeller		
1 LUSH New Zealand		
1 Sessils Foote		
1 DisneyLovesOuntes		

Figure 4: Data Selfie: Top Likes

Top friends & Top Pages & Top Likes

These first sections (Seen in Figures 2, 3 and 4)are just data points, laid out cleanly. These points are not processed in any way, they are raw and unassuming. They are interesting to reflect on because it shows my online performance without any algorithmic processing. What is seen is the clinical results of my actions. I cannot disagree with what is seen here, but I can see how my actions on the internet when quantified paints a different version of myself than what I would normally consider. If I made custom "top friends" and "top liked" lists they would differ greatly to what is seen here.

This shows the power of the work of watching (Rushton, 2012); if we invest time in things that we do not stand by completely we are doing a disservice to ourselves and our data profile. It is this raw data that is the foundation for all the assumptions machine learning and natural

language processing make on us. The data that has been collected is a reductionist view on a social medium which is being presented in a way that it was never intended by the user to be viewed as. This presentation is not a short falling of the extension, but of how social interaction is being captured via data.



Figure 5: Data Selfie: Object Detection

Object Detection

This section (see Figure 5) uses machine learning and computer vision to detect the subject matter of images I have seen on Facebook. This section shows that machine learning can deduct some interesting statistics from various kinds of data. This section is also somewhat unnerving, as it gives the impression that Facebook is 'watching' or viewing not only the text I am exposed to but also images. This section shows me how social my Facebook is, with 249 people present in the 207 images evaluated. People are by far the most common identifiable object, with 'cup' being the second most at 14. This is to be expected as connection with others is the focus of my Facebook use.

- Entitie Relevan	s ce (0 to 1	1) and sentime	NLP ent (-1 to 1 / negative to positive) of people, organizations, places, a.o. mentioned in the recent
content	you nave	TOOKEG at	
0.74	0.78	positive	Tony Stark (Person)
0.44	-0.86	negative	Facebook (Company)
0.36	-0.68	negative	Tribeca (Location)
0.35	0	neutral	Lambton Quay (Facility)
0.35	0.44	positive	Michael Fowler Centre (Facility)
0.35	0.77	positive	Kanto (Location)
0.33	0.36	positive	Yas (Person)
0.32	0	neutral	Te Auaha (Facility)
0.31	0	neutral	Liv (Person)
0.3	0	neutral	Naith (Location)
4			
			updated: Jun-13-18

Figure 6: Data Selfie: Entities

Entities

This section (see Figure 6) uses natural language processing to make assumptions on material directly referenced in my data. I do not remember viewing or interacting with a lot of these entities (some of which are fictional or misconstrued). The 'relevance' of these findings does not seem to correlate with my memory of viewing these entities. This illustrates data's perfect memory (van Bree, 2016) and it's ability to recall facts that would have been lost if not recorded. The fact that I have a reasonably negative sentiment towards Facebook as a company is now recorded within Big Data along with the other shallow data points present here.



Figure 7: Data Selfie: Catagories

Categories

This section (Figure 7) gives a score on how closely a category that is not directly referenced in my data relates to my data self. This section is not very meaningful but is a good example of how irrelevant things can still be connected to you despite having never interacted with them. It is unsettling how my data knows that I like birds despite apparently not having directly referenced them.

Personality Prediction					
conservative + traditional	h nanese	liberal + artistic			
impulsive + spontaneous	conscientiousness	organized + hard working			
contemplative		engaged w/ outside world			
competitive	+ Agreeableness	team working + trusting			
laid back + relaxed	+ + Exotional range	easily stressed + emotional			
Ø percentile average 100th percentile					
		updated: Jun-13-18			

Figure 8: Data Selfie: Personality Prediction

Personality Prediction

This section (see Figure 8) is of great interest because it attempts to make predictions and assumptions on my personality itself. It clearly demonstrates the difference between what I have been recorded as looking at and with what I have typed. A lot of my active use of Facebook is to type in chat with friends, so the data that comes from my typing is the most interesting and I would assume well-informed. This personality prediction work with the Big Five Personality traits, also known as the Five Factor model (Digman, 1990). The five personality traits represented are Openness, Conscientiousness, Extraversion, Agreeableness and Emotional range or Neuroticism. These are sometimes referred to as OCEAN for short.

I consider myself more liberal than conservative and this is reflected by the data collected from my typing. Considering what I look at, it seems that my Facebook has presented me with content that is mostly conservative. I can only make assumptions on what makes the content I look at "conservative", perhaps that is the content I am being advertised, or perhaps my friends are posting conservative material that I am viewing. This data point makes me consider using a more critical eye when browsing Facebook.

From what I can gather from my extraversion data point is my data assumes I am highly contemplative. This is likely a representation of how I use Facebook, as opposed to how I present. My Facebook use is introverted because I mainly type in chats with my friends, as opposed to connecting with others through group or event pages. I consider this a misjudgement on my character, but an accurate analysis of my social media use. This is much the same as my interpretation of my Agreeableness and Emotional range values.



Figure 9: Data Selfie: Religious Orientation

Religious Orientation

Spirituality being assumed by data points. This, in my opinion, is one of the most personal estimations made by this extension. The advantage of looking at religion in terms of a bar graph shows that there is some play between different schools of thought. It is interesting to think that to the machine learning engine my religious views can be a blend of different ideologies when typically religions have been averse to choosing which parts one would like to believe. In this respect the insight of how someone could potentially communicate their religion is nuanced and progressive.

On the other hand, the assumptions made about me are quite false in many ways and I feel uncomfortable with being referred to as 3% Mormon (see figure 9).



Political Orientation

Politics lend themselves well to being on a spectrum, so seeing a bar graph showing a blend of positions is a familiar sight. I personally try to stay out of politics on Facebook due to the toxic comment wars that I have so often bore witness to. The correlation between what I view on Facebook and my political leanings hold little weight in my opinion. This is in part because my results do not skew drastically one way or another. This is probably because it is difficult to assume someone's political leanings based on the content that appears on my Facebook as I mentioned; I try to stay out of Facebook politics.


Figure 11: Data Selfie: Other Predictions

Other Predictions

These other predictions (see Figure 11) are very here nor there. It is already somewhat absurd to be able to assume anyone's intelligence regardless of the tests you put them through. To say you can assume intelligence from data scraped from the internet is fantasy. Likewise with life satisfaction and leadership scores. Gender is already a social construct and performed in accordance to social biases anyway (Butler, 1990) so that statistic can also be disregarded. Conversely, I am sure advertisers pay great attention to the gender statistic, being the reason I am constantly bombarded by advertisements for Clear Blue the "Only test that tells you how many weeks" ("Clearblue Digital Pregnancy Test with Weeks Indicator," 2015).

It is always refreshing to see gender being represented as a spectrum and it is fitting to say that I am near the more androgynous grey zone of being female.

 Shopping Preference 	ces	ML
Based on what you	type your personality (Big 5), values and needs are predicted and these preferences can be produced	
Likely to	prefer quality when buying clothes	
Likely to	prefer style when buying clothes	
Likely to	prefer comfort when buying clothes	
Not likely to	be influenced by brand name when making product purchases	
Likely to	be influenced by product utility when making product purchases	
Not likely to	be influenced by online ads when making product purchases	
Not likely to	be influenced by social media when making product purchases	
Likely to	be influenced by family when making product purchases	
Not likely to	indulge in spur of the moment purchases	
Likely to	prefer using credit cards for shopping	
		-
4		►.
		updated: Jun-13-18

Figure 12: Data Selfie: Shopping Preferences

Shopping Preferences

This section (see Figure 12) is likely a section that is very highly sought by advertisers. It is empowering to be able to see how I am viewed by potential advertisers. I cannot escape from seeing advertisements, but it is nice to know why I am being shown some of them. I do not consider my shopping habits an integral part of who I am, so I do not find this section invasive. My behaviour on Facebook has a lot of traits rooted in consumerism. I see and engage with advertisements on this platform, so I find it reasonable that predictions be made on my shopping habits. This contrasts with how I feel about my spiritual views or intellect being predicted by my data.



Figure 13: Data Selfie: Concepts

Concepts

This category has since been removed from the most updated version of DataSelfie. This screenshot (See Figure 13) was captured early on in my interaction with the extension, so is based on less data than the other figures. This category attempted to make assumptions on a subject that had shown up in your data, but you had not directly referenced recently. I can assume that this section was removed because the claims being made were too random. I have limited recording of content I saw in this section, but anything I witnessed being inferred in this section was generally very confusing. This brings cause for concern if my data is making these obscure conclusion about me, why is it being collected at all. My data is not very good at guessing my opinion on things I have never experienced, therefore, are these guesses being taken into consideration when I am being profiled?

I am frustrated that my data is being used to make assumptions about me on things that I have never interacted with. I personally, unlike many others (Andrejevic, 2014), don't mind personalised advertisements. The issue with them is their subject matter often consists of things I don't normally interact with. Because of this, they are being targeted to me not by my profile data directly, but by data adjacent. From an advertisers perspective this makes sense, make me want what I don't know I want yet, but so often these predictions are too far off for me to care about the product. I have zero interest in 'Laser medicine' (Figure 13).



Figure 14: Data Selfie: Health + Activitys + Other Preferences

Health + Activity + Other Preferences

These two sections (see figure 13 and 14) are an amalgamation of all of the above data to create expectations of online activity. This is the product that Facebook makes of you. The reason Facebook and other social media are free is a result of these vague predictions being sold to advertisers. These predictions that in my case, after a year of selftracking, are largely false.

This data is claiming itself redundant saying I am "not likely to be influenced by social media when making product purchases". My data is tracked largely because advertisers would like me to purchase their product. Why track my data if you know that social media is not going to affect my chance of purchasing your product? Advertising works, if I see something I am more likely to buy it. In this case, I disagree with this assumption, I am likely to be affected by my social media. How dare my data tell me otherwise.

It is disappointing to see that after all the data points previously stated, my Data selfie comes to the conclusion that social media does not affect me. A conclusion I disagree with.

Findings

I found being able to hack your own data empowering. Currently, social media users feel powerless to control how much data they are uploading to the internet (Andrejevic, 2014). By allowing users of Data Selfie to see how much and what kind of information they are uploading they are given access to some of the utility of Big Data. They are still creating data, but at least now they have been given something in return in the form of personalised graphs and figures. Self-data mining allows us to gain a better understanding of how interacting with social media has a real impact on your data footprint (Andrejevic, 2014).

The predictive ability in this extension is impressive and seemingly gives a lot of detail, although the detail is almost always slightly off. This could very well be because it was only a relevantly small amount of data collected. This extension does not track all of the data points that are possible to track, such as location and what your friend's activities are. If more data were available to the extension, then perhaps it would be more effective at making personal predictions. It was insightful to see how over the course of using Data Selfie, my data profile developed as more information was collected. Through watching this relatively small amount of self-tracked data become more intelligible, I considered how I was producing data on other sites. As is, I feel like my Data Selfie is not a very good representation of me, furthermore, I find the data collected on me a poor proxy of my identity. Despite this, I know that any information about my Facebook usage the internet has on me is captured and catalogued, not to be deleted (van Bree, 2016). My internet proxy will be remembered.

Projects

Intro

The next section focuses on the creative experiments that were manifested after considering the content of the literature review and my experiences with Data Selfie. This section covers the documentation and elaboration on the design experiments created in response to the themes this thesis has explored so far. These artistic expressions are the culmination of my research.

First is a physical illustration of an attempted re-connection with a person; attempted at distance.

Second is the response to internet self-branding and Web 2.0 culture.

The third is the physical manifestation of my data.

These experiments are the Action Research conducted as part of my implementation of the research through design method (Frankel & Racine, 2010). The creation of these artifacts allows for a greater exploration of potential ways to increase understanding and critical thinking around the topics of online performance, data collection and social media's effect on people. The outputs in this section are physical, tangible objects. The intent behind this is to highlight how Human-Computer interactions have meaningful, lasting impacts that are as important as interactions that manifest physically. Each project could be summarised simply by its statistics, how long it took to make, what its materiality is, the snapshot of how it exists now. The meaning comes from the process of creation as well as the intent behind each piece.



Figure 15: Embroidery: Connection

Embroidery: Imagining others complexly

Project data<> Time took to complete: Approximately 1 hour 45 minutes Unfinished Medium: Embroidery Size: Approx. 17cm by 12cm Subject: Close friend, flatmate. Facebook Profile picture.

The profile picture was current from 10th February to 24 March 2018 A key aspect of a person's online presence is how they represent themselves. In the Web 2.0 era of the internet, special attention is spent on the images and visual representations of things and people. Therefore, an important part of your internet portrayal is your visual representation, the profile picture. The image next to your name on any social media. Care and consideration are taken when selecting this image as it is the image that will most often be associated with you. Profile pictures appear next to anything a person posts. This frequency of appearance on a social media means that this image above all others is the one that is most strongly associated with a personal profile. This project is about connecting. I used it as a form of a meditative process with the intention to honour my friendship with the subject. From the consideration of performativity and my Data Selfie interaction, I consider a person's social media profile to be an inaccurate portrayal of them. This is because of the impression management techniques that a person uses while creating the content of their profile. Furthermore, the data that makes up a person's profile is not complete enough to approximate personhood. This project attempts to explore a friendship by considering the space between the profile, and the person. It does this by combining the intent of the social media profile of a person, with my thoughts and feelings about that same person.

Interacting on social media is something that is typically done alone, with the strive for connection. I wanted to simulate this but through a different medium. A medium that did not have a data footprint but did have a tangible output. This attention-seeking effort through the act of observing was heavily influenced by the way in which Steve Rushton (2012) describes the act of watching as a form of creating. I chose embroidery as my medium because it takes a long time to create, meaning I could spend time with the image I was recreating and put thought into my creation. It is also an activity that I did while alone, striving for a better connection with my friend, in the same way, I do through social media. This is also a medium that my friend uses to



Figure 16: Embroidery: Friendship

express herself, but I have never attempted before. As I was working I could think about her, and how she would have experienced her first attempt at embroidery. I now have an experience that is both shared with her and adjacent to her.

I chose a profile picture to embroider that I saw every day but glanced over. I know what my friend looks like, I see her often, this image is not special to me. But it is to her because this is how she (at the time of embroidery) chose to represent herself. During the embroidery, I am looking at my friend's data, and then thinking about her. This is in contrast to how algorithms such as those that power Data Selfie looks at her data. The algorithms cannot 'think about her', they can only make assumptions about her. They can look at her profile picture using computer vision and identify 'person', but this is a shallow interpretation compared to knowing a person and their background. Through the creation of this embroidery try to strengthen my connection with this person by not only putting my intention into an embroidery of her but also considering her as she has curated herself online to be seen. I can imagine her more complexly than an algorithm, but there is no substitute for sharing an interaction where both parties are present.

The stitching itself was a tedious process, but it was satisfying to have a tangible output as opposed to the digital expression of friendship I am more accustomed to. I missed the ability to hit an undo button, and any stitches I was not happy with had to be painstakingly unstitched and redone. This is a fitting analogy as to how information once made publicly available is difficult to reclaim. While I stitched I considered how much easier this task would be if I had gotten an embroidery machine to create the project for me. The final product could have been cleaner and more precise. Clean and precise are two words I would use to describe raw data, but not as descriptors of a friendship. This is the issue with describing relationships that occur online, while they are captured through data they are more than the information they are comprised of. While I know the subject of this embroidery in my offline life, I have had friends and connections that exist purely through the internet. While the moments I have shared with these people are stored in databases, they are unquantifiable as interactions because human interaction is very difficult to capture in its entirety. Getting to know someone online is more than observing their profile and making assumptions, it is spending the time through real interaction and intent.

This artifact's intention is to reconnect with the person themselves and to see if this is possible through embroidering with the intention to reconnect while viewing a profile picture. This re-connection is important because too often people get hung up on numbers, statistics and data on social media (Macmillan, 2017). It is one thing to leave someone a 'like' on Facebook, but it is more personal to pay someone a compliment. Unfortunately, we too often get fixated on the statistics of social media, instead of welcoming the more natural social aspects. While it is a



Figure 17: Embroidery: Detail

relatively grand gesture to embroider my friends face, it allowed me to put our social media interactions into perspective. It is not necessary to make art of your friends to connect with them, but more effort is needed than simply clicking 'like'. Through this, I believe this embroidery has been better for my friendship with the subject than if I had just observed her Facebook but is not as beneficial as taking her out for coffee would have been.

With all the good intentions of this piece, I still consider it unfinished. The hair could have been filled in. More detail could have been included. When I consider why I did not finish this piece, I think it comes down to how I express my modern relationship with my friend. After an hour and forty-five minutes of solid intention-based meditative work, I had had enough. When I show my friend attention and respect normally, it is through a quick message or sharing of a photo I know she would enjoy. This dedication of time spent ruminating was unusual for me and showed me that I had fallen into the habit of wanting instant gratification. I wanted my actions to have an effect as soon as I was through with them. The final product, while unfinished at least looks like the profile picture.

This project has not resulted in a perfect recreation of an image, but an approximation. While this embroidery can't assess my friend on her OCEAN values like a dataset could (Digman, 1990), it is intended to be more meaningful to her because I made it while thinking of her.



Figure 18: Popping Hashtags: Flatlay



Figure 19: Popping Hashtags: Detail

Popping Hashtags

Project data<> Gathering and assembly time: 20 minutes Item count: 34 Pill weight: 40mg Subject: '#' and Beta Blockers

This project is best understood when you take into consideration my personal relationship with pills. When I was younger I could not take them, but now they are part of my daily existence. As someone with chronic pain, I require pills to function most days. The materiality of this artifact has the desired intent to represent my personal journey of normalisation. The small, green pill is a "Propranolol", prescribed to ease the physical symptoms of anxiety. Just as I grew up with the internet and learnt to navigate it with great dexterity, I have learnt to interact with pills.

The issue with entering Web 2.0 with such idealised intentions (Lovink, 2011) is once our internet experience does not live up to them we feel like we have missed out. On social media sites such as Twitter and Instagram, we are presented with others that have large followings and the community's behind them. To gain these communities, microcelebrities have had to spend a lot of time and effort in curating their image, and their techniques of posting in a way that best facilitates community (Page, 2012). Often this involves the considered use of hashtags. The issue here comes from how to use the labelling and compartmentalisation of the internet to better market oneself goes against wider internet community ideals. Hierarchy on the internet is unavoidable, but it is not helped by giving those who hyper curate their image to an unrealistic and impossible ideal.

The bombardment of new hashtags, memes and online trends can be overwhelming. To keep up many feel the need to observe their peers and keep up the work of observing, which can become exhausting (Erica Scourti, 2017). My intent with this project is to draw attention to the excess found in the posting of shallow materialistic content but in a playful manner. Hashtags are a creation of Web 2.0 and they provide a useful tool for users to categorise the content they produce. The fact they are used to systematically categorise cat pictures is less of an issue, and more of a distraction from the more important issues surrounding privacy and interpersonal communication.

From my embroidery project, I discovered my tendency to desire instant gratification and feedback. I applied this thinking to my internet habits and found that this was a theme there also. This spawned the



Figure 20: Popping Hashtags: Lights



Figure 21: Popping Hashtags: Soft Focus

idea of "popping a hashtag" and feeling better. This lead to the first iteration of this concept. When I need a distraction or require a step back from the context I find myself in I often turn to scrolling social media. Using platforms which categorise content such as Reddit and Instagram, it requires little effort to be gratified with pleasing images that follow a theme. For example, on Instagram, there is #asmr, #corgibutt, #satisfying and numerous other hashtags that are associated with relaxing or calming images and videos. The content found on these hashtags is extensive and, I would argue, excessive. While it is a great asset for the internet to have extensive lists of similar images, there must come a saturation point. To interact with these images or posts, I do not have to contribute to a community or show any original thought, I simply observe. Though this act of observing is important to note in its own right (Rushton, 2012), it is not a constructive activity. This phenomenon further enforces the escapist aspect of social media mentioned by Turkle (2011). Taking in my daily hashtag dose in the long term does not better me as a person, it simply distracts me from the here and now.

The expression of self-online is exaggerated as well as curated, which facilitates escapism further. People are not pleased with their achievements, they are thrilled and overly expressive. Sarcasm online is difficult to communicate, so users are very accustomed to hyperbole. This also is reflected in the way in which people post images online. Through filters and face tuners, users can look the way they would most like to look. Through overexposure to these highly produced images, they have slowly become the norm. I chose to mimic one of the most common techniques I have observed on Instagram to create more visually appealing images, the use of shallow focus and fairy lights (seen in figure 20 and 21). This is a semi-subversive approach to making the message behind this image set more palatable to those who are already engulfed in hashtag culture. The visual aesthetics of images are so often considered over all else when curating a profile. This enforces a notion that when going about daily activities that bring about joy, social media uses should take a moment out of whatever they are doing to document it in a visually pleasing way. While the sharing of experiences is an important part of building online communities, it is beginning to influence the way in which we enjoy our lives. From the perspective of social media viewers and our data footprint, our posts online encapsulate our existence. If you don't take a picture of something and share it, it did not happen from the perspective of the internet. This produces the pressure to share, to be more 'genuine' and have a 'complete' profile. Conversely, we only share in accordance with the image we wish to display. This is just an example of how our curated self performance affects our social media habits, but our social media habits affect our performance of self.



Figure 22: Data Haemorrhaging: The Ingredients



Figure 23: Data Haemorrhaging: Mix Together

Haemorrhaging data

How to make data:

¹/₄ Cup of resistors. Can be of any variety, but best to be of the keyboard warrior type.

Add a microcontroller for the ability to look back at old posts and cringe. Also useful for micromanaging your self-branding.

Add diodes to preference. Some to keep the data moving and prevent clumping. Some bright shiny ones to keep whoever is watching entertained.

To keep present with the current state of the web, add some connection. Be this spider webs, insulated wiring or Internet Protocol addresses.

1 cell: like the cell of a form you fill with all your intimate details. This is the powerhouse of the data.

Optional: capacitor, for the capacity to empathise with the situations of others. Also increases the likelihood of online friendships, built on sending cat memes, forming.

3 cup of lube, would recommend a partially zesty controversial opinion on a public platform.

And as much of your true self as you can muster. In a pinch, store bought is fine.

Apply a blue filter, because that's comforting and friendly.

Mix together, share with everyone, have your life be summarised

by it, treat without caution.



Figure 24: Data Haemorrhaging: Uncomfortable



Figure 25: Data Haemorrhaging: Specimen

Project data<> Gathering and assembly time: 40 minutes Duration: Ongoing Subject: Personal data reimagined Components: Poem, Electronic waste and blue matrix.

The initial concept for this idea was to show a physical representation of data. The themes of ownership of data, how data is taken from users and how that data is representational of the self, are communicated through this recipe style poem and its physical manifestation. This project intends to encourage engagement and awareness around the topic of personal data.

The use of electronic waste components was a way of recycling objects that I had around me as well as having their original uses be considered in a metaphorical sense. The recipe style poem intends to be reminiscent of DIY culture and internet instructional posts. It serves as a commentary on behaviours that lead to personal data being shared, internet slang, the lack of privacy concerns and internet culture's predisposition towards curating their online performance. This abstraction of the original purpose of the electronic waste highlights the abstract ways we talk about our data and Human-Computer interactions. We use metaphor when talking about how we interact with a computer, from the 'desktop' to the 'web' itself. I, therefore, chose to represent data in a highly metaphorical sense both into increase its tactility and as an in jest nod to current HCI norms. The abstraction of the original purpose of these electrical components also serves to mimic the way in which our data is misconstrued by being analysed algorithmically. From my Data Selfie interaction, I found assumptions where being made about me on topics I had not interacted with. As a response to my data drawing parallels between loosely related aspects of my character, I chose to represent data in a similar way. Just as a capacitor is not an accurate portrayal of empathy, data is also a poor proxy for a person.

The intent behind showing the "data" in a clear and yellow specimen container (see Figure 25) is to acknowledge the phenomena I call Data Haemorrhaging. The Oxford Dictionary defines haemorrhage as "To seep, grow, or spread uncontrollably; to be rampant." or "To dissipate or expend (something, esp. money) in large amounts, as if by allowing it to drain away." ("haemorrhage, v.," 2018). Data is being created on the internet by users rapidly, and these users describe themselves as powerless in response to concerns about the collection and use of personal information (Andrejevic, 2014, p. 12). Data is a powerful tool we give the internet. Through our outpouring of data, we give those who receive it insight into our actions, locations and behaviours. Our data is extremely valuable and we give it away often without a fully comprehending the ways in which we are compromising our privacy (Taddicken, 2014). I chose to represent data in a medical setting to



Figure 26: Data Haemorrhaging: Leaking Battery Acid



Figure 27: Data Haemorrhaging: Detail

highlight this clinical and purposeful collection of a users' personal information. I considered the question: if the user could see their personal data sitting before them how would they feel? Uncomfortable? Confused? Surprised? Curious? These are feelings I experienced when presented with my Data Selfie, but I wanted to push these feelings further into the uncomfortable and clinical. It is through seeing our data in an uncomfortable setting I highlight the paradoxical nature of giving up something so important willingly. The scale of the data that is pictured is intended to be at a personal, humanised scale. This small specimen container of material is not a representation of Big Data, but of the data produced over a day by a single person. As we haemorrhage data into the servers containing Big Data the scale of our own contribution is easy to dismiss. The importance of considering data on a human scale allows for it to be considered in a tangible way.

The purpose of including the cell battery was to power the LED's. When the components are suspended in the matrix the cell battery will occasionally make contact with the anode and cathode of the LED. This causes the LED to glow. The effect of this adds playfulness to the artifact, encouraging interaction and curiosity. This interaction also illustrates data's sporadic usefulness to the user. Upon occasion, the information stored on us will be of use, but this is infrequent in my experience. One use for the collection of data is targeted advertisements, but the majority of people dislike this feature of the internet(Andrejevic, 2014). The cell battery is also the first part of the artifact that breaks down over time, leaking battery acid into the matrix and turning it opaque (see figure 26). This effect was unforeseen but is a serendipitous indicator of how when left alone data can have more malice than at first thought. The 'data' laced with battery acid is now somewhat dangerous to open because it has been left so long. It has become undesirable to investigate or even look at. This draws parallels to the discomfort of haemorrhaging data for years, and not knowing what to do with the data footprint created. The feeling that important life moments are collected and stored in an inaccessible and confusing way is commented upon by Andrejevic in his paper (2014). Andrejevic raises the importance of informing social media users about their contribution to Big Data to give them back some conscious control.



Figure 28: Data Haemorrhaging: Uncanny



Figure 29: Data Haemorrhaging: Data Reimagined

The images of my data spread on women's sanitary products (see figure 28, 29 and 30) were created from the thought of "Where would personal physical data come from?" The conclusion came when considering in what case I haemorrhage in everyday life. These images are crafted to be both personal and somewhat familiar to some in the hopes of creating an uncanny feeling. Here, the blue not only represents the colour of social media, but also the censoring of menstrual products (Thorpe, 2017). The purpose behind this being that the censorship and secrecy surrounding data collection is as alarming as the collection itself. The visual metaphor also plays with the idea of having data haemorrhage being "normal" but not talked about; much like menstruation (Thorpe, 2017). When taboo topics are not talked about, they gain power over those negatively affected. If we don't speak about the data that is haemorrhaging from our social media accounts, we have no chance of stemming the flow. The intention of this design is to embody a feeling. The feeling of being out of control and losing something valuable. When our data is taken away from us passively during our internet use, we are unable to conceive how it manifests in the eternal servers of Big Data (Andrejevic, 2014). While our information can be used to make very specific assumptions about our character (Kosinski et al., 2013), it is still just a small aspect of our sense of self that is being collected. As discussed in my Popping Hashtags project, on social media we are presented with the idyllic versions of our friends and family. What is not seen is the sinister effect this has on our data. This design intends to confront the users with how uncomfortable it is to have a disconnected experience with an important aspect of your life that is being collected and stored.



Figure 30: Data Haemorrhaging: Disconnected Experience

Discussion

Attempts on quantifying the performance of self through online practices have become commonplace in our current online climate. The act of interacting with social platforms produces personal data, data that is used to build your data footprint. This data footprint can make assumptions about you by comparing your data against others (Kosinski et al., 2013). The original interaction is performed, it is curated in accordance to what you see your peers do (Boyd, 2016), and the persona you wish to convey. This interaction with social media can be portrayed by the user through an anonymous handle or attached to your personal accounts (such as Facebook) that reflect the 'you' desire to be perceived as. The internet only knows about you through interaction, and this interaction is dictated by your performance.

While in typical social situations, if we play a role that is not our's it is easily spotted by those around us. On the internet, we are only able to capture a relatively small part of a person's social performance, so the data captured on us will never be able to encapsulate all the details needed in order to represent a human (Lanier, 2010). Big data is too large to comprehend, but it is only able to approximate the human psyche. Theoretically through complete surveillance and the active attempt to upload as many details about someone as possible then we would be able to make more accurate approximations of someone's 'backstage' as Goffman (1959) would put it. As is, the assumptions made on people's data footprint only make assumptions on the 'front' we put up. My investigation into my own 'front' in my Data Selfie explorations furthered my understanding that whatever we put out into the world will be objectified and simplified by someone else. These judgements are not only applied to my posts but the content I consume as well. As previously discussed, when people interact with social media they produce the content that they take in. It can then be assumed that if all we saw on our social media was content from brands and advertisements then we would also start emulating that content. Erica Scourti raises the question, why are some accounts on social media considered nonfiction, whereas bot accounts are considered fiction (2017). She has created Twitter a bot that contains elements of herself which show aspects of her. The same can be said about her or anyone else's personal Twitter accounts. If these accounts are to be considered fiction, then Turkle's argument on the benefits of escapism onto the internet (Turkle, 2011) can be further understood. Turkle explains that escapism can be a gift of existing on the internet, it gives us a private outlet where we can be whoever we want to be. Through Scourti's argument that we can

consider any profile online fiction because of their limited display of a person's personality, Turkle's description of escapism is even more fitting. It is not to say that internet communities are fictions, but the characters portrayed through personal profiles are less than whole personalities.

The way in which the internet treats interpersonal interaction is described in Carmen Hermosillo's work when she stays how cyberspace "absorbs energy and personality and then re-presents it as spectacle" (1994, para. 3). That energy and personality being absorbed can come from a person interacting with the internet, or from a database containing data of people's online activity. These are both spectacles as opposed to accurate accounts of personhood. With the amount of content, we are exposed to daily, it is difficult to keep in mind that there are real people on the other side of the screen. The concept of reconnection was explored in my embroidery project. This project highlighted the difficulties of surrounding the topic of connection through the internet, some of which could be attributed to our exposure to mass spectacle online. While meaningful friendships through the internet are greatly rewarding, we are often distracted by consuming the huge amounts of shallow, sensationalised content. The difficulties existing in a world where we are being bombarded with shallow context emphasised by spectacle is also an important attributing factor to my Popping Hashtags project. That project highlighted how we consume content as and when we desire it without consideration of its quality or the effects of our actions. Our shallow actions are being absorbed and represented as shallow distracting spectacle. Both distraction and a shift in how we connect with others can attribute to how sensationalism on the internet has become prevalent.

A similar speculation can be made about the blurring lines between over curated social media accounts and brand accounts. Page discusses how personal self-branding culture has come from the marketing of corporations (2012). Corporations in the Web 2.0 climate have their own social media accounts. These accounts don't represent one person, but a brand. Taking into consideration that people perform in a way that is comparable to the media they take in (Boyd, 2016). It is reasonable to conclude that people who are exposed to the accounts designed to market, will alter their online performance to emulate the marketing spectacle they see. By being exposed to highly curated and designed accounts, we are encouraged to brand and curate our own performance online. This only amplifies the curation of personality when posting on the internet, and again arguably is a cause for escapism found in living a duality online described by Turkle (2011). While the duality between the online curated self and offline more genuine self is appealing, it only perpetuates the perceived need to project perfection onto the internet. Although using social media as a form of escapism is greatly appealing, the effects of Digital Hyperthymesia (van Bree, 2016) must still be considered. This perfect memory means recalling facts about a topic

can be done with ease and with great precision. This memory is hugely useful for cataloguing the wealth of human knowledge but has its drawbacks when personal opinions are online. We can scrutinise our actions online and once we have our say on a particular topic, that viewpoint is now permanently connected with us (Lanier, 2010). Because of this, the information that we put on the internet has an impact on our offline life, more so than information shared by more traditional means. The information we put online can be difficult to delete or censor should we want to reclaim it, and because we have very limited control as to who sees the content that we post about ourselves. It this lack of control I illustrated through my Haemorrhaging Data project. The data that is taken from us, and stored forever grows stale over time, obfuscating its original meaning and origin. It is my theory, if our devices seeped data physically we would be far more cautious about our privacy. It would be more difficult to think of data paradoxically if we were forced to clean up after it, should it otherwise be taken from us. Once we have been labelled in a certain way online it is very difficult to retroactively change this label (Lanier, 2010). This is especially prevalent in the current internet climate as we compartmentalise our actions. We have hashtags that we label things with and this is part of the curation of our online persona. The issue comes when we attempt to make judgements on a person's character from the labels that are associated with them. Labels, be they hashtags associated with certain aspects of a person's profile or extrapolations of their data footprint, tend to be misleading. Through my interaction with Data Selfie, I concluded that data collected from social media is unable to accurately portray social interaction use because that data is not complex enough. Our data is just a small part of our expression of self, captured in a fixed medium. The distillation of our pure performance into data is equivalent to the Brown's (2012) argument that decontextualized phenomena is oversimplified and cannot result in findings that would naturally manifest.

Conclusion

This thesis explored how through design, the consideration of a person's behaviour in terms of performativity allows for a greater understanding of the nature of the information that they publish. A person's social performance online is dictated by many factors. The relationship between users' behaviour online and their respective data they upload is greatly dictated by current sociotechnical norms. As with any social performance, people represent themselves online in a variety of ways, and this is reflected in the varying communities found on the internet. It is within these communities that differing self-publishing behaviours occur with varying levels of self-curation, each contributing to Big Data in different ways. This variation online is attempted to be categorised through machine learning and natural language processing techniques, which lead to the assumptions data makes upon users. This representational data must be considered in the context in which it was created, being on social media platforms where people are able to curate their online self-image. Through design, we can highlight these findings in order to bring greater consideration how these factors affect the users of social media.

Further topics that have the potential in being explored could involve the commodification of data and personality from either a corporate or self-branding perspective. Further research that would inform understanding in this field could include in-depth quantitative analysis of a large sample size of social media data that then could be considered in relation to users performativity. Through this further understanding can be made about the specific types of data people publish through use of different social platforms. A qualitative analysis on a large user set would also further unpack the social constructs around online performance. Working with large datasets was outside of the scope of this thesis but in order to reinforce broader claims, large-scale data is necessary.

In conclusion, despite the scope of big data and our understanding of performativity, making assumptions about people from their data alone results in misrepresentation. Data cannot capture the complex social behaviour because it only captures the essence of the performance to the simplest coordinates; it attempts to represent social phenomena from an oversimplified source. It is important to make designs that are considerate of this in order to avoid perpetuating this misrepresentation.

Works Cited

All images created by the author Phoebe Zeller

- Altman, I. (1975). The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding.
- Andrejevic, M. (2014). Big data, big questions the big data divide. International Journal of Communication, 8, 17.
- Archer, B. (1995). The nature of research. *Co-Design Journal*, 2(11), 6–13.
- Boyd, D. (2016). Do You See What I See?: Visibility through Social Media. *In Education and Social Media* (pp. 49–60). MIT Press. Retrieved from http://www.jstor.org.helicon.vuw.ac.nz/stable/j.ctt1c2cqn5.8
- Brown, S. (2012). Abstract Experimentalism. In C. Lury & N. Wakeford (Eds.), *Inventive Methods The Happening of the Social* (pp. 61–75). Routledge.
- Butler, J. (1990). Gender trouble: Feminism and the subversion of identity Routledge: New York.
- Butler, J. (1993). Bodies That Matter: On the Discursive Limits of "Sex." Routledge.
- Chaffey, D. (2017, February 27). Global Social Media Statistics Summary 2017. Retrieved May 7, 2017, from http://www.smartinsights.com/ social-media-marketing/social-media-strategy/new-global-socialmedia-research/
- Clearblue Digital Pregnancy Test with Weeks Indicator. (2015, April 13). Retrieved June 28, 2018, from http://nz.clearblue.com/pregnancytests/digital-with-weeks-indicator
- Creswell, J. W. (2009). Research Design : Qualitative, Quantitative, and Mixed Methods Approaches. (3rd ed.). SAGE Publications, Inc.
- DATA X. (2018). Data Selfie. Retrieved July 7, 2018, from https://dataselfie. it
- Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. Annual Review of Psychology, 41(1), 417–440. https://doi. org/10.1146/annurev.ps.41.020190.002221

DiNucci, D. (1999). Fragmented Future. Retrieved from http://darcyd.com/

fragmented_future.pdf

- Erica Scourti. (2017). ON CARING: a short illustrated reading on self help, "coding a line" & the ethics of AI. Retrieved from https://www.youtube. com/watch?v=3o7Ur1Q02A8
- Fisher, M. (2014). *Ghosts of My Life: Writings on Depression, Hauntology and Lost Futures.* John Hunt Publishing.
- Frankel, L., & Racine, M. (2010). The complex field of research: For design, through design, and about design. In Proceedings of the Design Research Society (DRS) International Conference (No. 043).
- Friedman, K. (2008). Research into, by and for design. Journal of Visual Art Practice, 7(2), 153-160.
- Goffman, E. (1959). The presentation of self in everyday life. New York.
- Grosser, B. (2016). You like my like of your like of my status [Generative text and Synthesized speech]. Retrieved from https://bengrosser.com/ projects/you-like-my-like-of-your-like-of-my-status/
- haemorrhage, v. (2018). *OED Online*. Oxford University Press. Retrieved from http://www.oed.com/view/Entry/243320
- Hermosillo, C. (Humdog). (1994). Pandora's Vox: On Community in Cyberspace. Retrieved from https://gist.github.com/kolber/2131643
- Inman, M. (2016). Creativity is like breathing. Retrieved June 30, 2018, from http://theoatmeal.com/comics/creativity
- Kielczewski, B. (2018). Apply Magic Sauce Prediction API -. Retrieved June 13, 2018, from https://applymagicsauce.com
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings* of the National Academy of Sciences, 110(15), 5802–5805. https://doi. org/10.1073/pnas.1218772110
- Lanier, J. (2010). You are not a gadget: A manifesto. Vintage.
- Lovink, G. (2011). Networks without a cause: a critique of social media. Polity Press. Retrieved from http://dare.uva.nl/search?metis.record. id=368935
- Macmillan, A. (2017). Why Instagram Is the Worst Social Media for Mental Health. Retrieved June 27, 2018, from http://time.com/4793331/ instagram-social-media-mental-health/

- Miller, H. (1995). The presentation of self in electronic life: Goffman on the Internet. In Embodied knowledge and virtual space conference (Vol. 9). Retrieved from https://pdfs.semanticscholar. org/089b/89cbb0af6d6d02122b0dfccd7a124a181b02.pdf
- Naaman, M., Boase, J., & Lai, C.-H. (2010). Is it really about me?: message content in social awareness streams. In Proceedings of the 2010 ACM conference on Computer supported cooperative work (pp. 189–192). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1718953
- O'Reilly, T. (2007, March). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software [MPRA Paper]. Retrieved July 2, 2018, from https://mpra.ub.uni-muenchen.de/4578/
- Page, R. (2012). The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. *Discourse & Communication*, 6(2), 181– 201. https://doi.org/10.1177/1750481312437441
- Rushton, S. (2012). *Masters of Reality*. (V. Ohlraun, Ed.). Piet Zwart Institute & Sternberg Press.
- Taddicken, M. (2014). The 'Privacy Paradox' in the Social Web: The Impact of Privacy Concerns, Individual Characteristics, and the Perceived Social Relevance on Different Forms of Self-Disclosure. *Journal of Computer-Mediated Communication*, 19(2), 248–273. https://doi. org/10.1111/jcc4.12052

Thorpe, J. R. (2017, October 12). This Is The Actual Reason Pad Commercials Use That Weird Blue Liquid. Retrieved July 3, 2018, from https://www.bustle.com/p/why-do-period-product-commercials-useblue-liquid-the-practice-has-a-long-bizarre-history-2957963

- Turkle, S. (2011). Life on the Screen. Simon and Schuster.
- van Bree, T. (2016). Digital Hyperthymesia On the consequences of living with perfect memory. In L. Jenssens (Ed.), The art of ethics in the information society (pp. 28–33). Amsterdam University Press.
- Ward, J. S., & Barker, A. (2013). Undefined By Data: A Survey of Big Data Definitions. ArXiv:1309.5821 [Cs]. Retrieved from http://arxiv.org/ abs/1309.5821
- Zimmer, M. (2010, May 14). Facebook's Zuckerberg: "Having two identities for yourself is an example of a lack of integrity" | MichaelZimmer. org. Retrieved May 28, 2018, from https://www.michaelzimmer. org/2010/05/14/facebooks-zuckerberg-having-two-identities-foryourself-is-an-example-of-a-lack-of-integrity/
