Government Office for Science

Foresight

Future Identities: Changing identities in the UK – the next 10 years

DR3: Social Media and Identity

Nicole Ellison

Michigan State University

January 2013

This review has been commissioned as part of the UK Government's Foresight project, *Future Identities: Changing identities in the UK – the next 10 years* The views expressed do not represent policy of any government or organisation.

Contents

1. Introduction

Identity is a construct that encompasses the way we think about ourselves and our role in larger social environments; identity is enacted through social interactions with others and our relationships with them. Along one continuum, identity can be understood as containing both individual, personal differentiators (such as one's personal history) as well as components related to one's role in social groups, such as the way we interact with others in online communities. A second spectrum for understanding identity concerns the character of the attributes we associate with an individual. Identity characteristics can be written on the body, such as gender or ethnicity, or are elective, such as our chosen political affiliations. How we see ourselves and our role in the larger social environment can have consequences for how we behave, what we believe, and who we affiliate with. Because online information and communication technologies have the potential to shape identity processes in meaningful ways, it is important to consider the identity implications of social media practices as well as the role of computer-mediated communication (CMC) more generally.

In online settings such as social network sites (SNSs), chat rooms, or discussion groups, identity processes are complicated because many identity cues (such as gender or age) are masked and can be purposefully shared, withheld, or misrepresented. In these and other online contexts, identity is essentially typed into being (Sunden, 2003). Individuals can adopt multiple online personae, and online activities often leave visible traces which can be captured, tracked, packaged, and shared.

Some of the questions that emerge around this issue include: Can we be anything we want online? How are online identity performances different from offline? To what extent do individuals share multiple or conflicting identities online? What are the consequences of conducting so many of our social interactions online?

Social media such as social network sites, blogs, wikis, and online discussion forums contain a set of social and technical affordances that have the potential to affect identity – to reshape how individuals view themselves and others. These affordances include the ability to engage in selective self-presentation when *presenting identity,* the ability to enact *multiple identities* in online settings simultaneously, and issues of *data, computation and identity.* This last factor stems from the fact that social media applications typically capture vast amounts of behavioral data about their users which can be used to track individuals across sites and identities, especially when paired with newer technologies such as face recognition software and geo-locational tagging.

1.1 Social Media and Social Network Sites

Use of social media has surged in recent years, initially spurred by young people but now used by all demographic groups of the global population. In the U.K., 60% of Internet users are members of a social network site, a 43% increase from 2007 (Dutton & Blank, 2011). Social network sites such as Facebook are perhaps the most commonly referenced social media applications, but other popular social media sites in the UK include Foursquare, Twitter, Tumblr, and LinkedIn. There is no universally agreed upon definition of social media although most agree that a key differentiator between social media and traditional, broadcast technologies (such as television or print newspapers) is found in the fact that social media allow users to create, share, consume, and collaborate around content in ways not previously supported by earlier technologies. Other activities enabled by social media may include rating, recommending, remixing, and sharing text, video, or image content.

Much of the research on social media focuses on social network sites (SNSs). A commonly accepted definition of social network sites was proposed by danah Boyd and Nicole Ellison in their 2007 essay. They define SNSs as “web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system.” In a recent update to their definition, they amend this to consider the changing influence of the media stream within SNS practices and to highlight the role of communication in SNS use. A more recent version, updated to reflect technical and social changes since the first definition, reads: “A social network site is a networked communication platform in which participants 1) have uniquely identifiable profiles that consist of user-supplied content, content provided by other users, and/or system-level data; 2) can publicly articulate connections that can be viewed and traversed by others; and 3) can consume, produce, and/ or interact with streams of user-generated content provided by their connections on the site” (Ellison & Boyd, in press, p. 158).

2. Presenting Identity

Erving Goffman's (1959) work on self-presentation is helpful for understanding the identity cues we share with others, both intentionally and unintentionally, and how this process is affected by how we see ourselves and want to be perceived by others. Goffman uses the metaphor of the stage to illustrate the differences between situations in which self-presentation concerns are salient and those in which they are less pronounced. When individuals are in public settings where they are trying to inculcate a specific impression among an audience, we can think of this as "front stage," similar to an actor on stage presenting a performance. This is in contrast to the "back stage" – a place where performers can relax and step out of character. In everyday situations, we are enacting performances for specific audiences: trying to be perceived as competent at work or entertaining at the pub, but not necessarily the other way around. It should be noted that engaging in impression management is not manipulative or deceptive, but rather a natural aspect of human relationships that in many ways can make interactions flow more smoothly and enable individuals to meet their personal and professional goals.

Social media offer new opportunities for sharing self-presentational content, or "branding" oneself online. One such form is through micro-celebrity, defined as "the commitment to deploying and maintaining one's online identity as if it were a branded good, with the expectation that others do the same" (Senft, 2012) or "an emerging online practice that involves creating a persona, sharing personal information about oneself with others, performing intimate connections to create the illusion of friendship or closeness, acknowledging an audience and viewing them as fans, and using strategic reveal of information to increase or maintain this audience" (Marwick, 2010). Marwick (2010) examines these practices among members of the Silicon Valley "tech scene," explicating the labor involved in performing seemingly authentic brand management of one's identity via social media each day, every day. Marwick and Boyd (2011) describe conflicting views on micro-celebrity and its practices among their informants, some of whom saw the ability to command attention as a signal of status and success. Others viewed micro-celebrity practices as inauthentic, 'phony,' and explicitly self-promotional. Their discussion highlights one of the tensions associated with identity and social media – the desire to be authentic contrasted with the need to speak to multiple audiences (associated with multiple personal goals), including unknown audiences, while engaging in self-presentational (and promotional) practices.

2.1 Characteristics of CMC That Affect Identity Processes

The affordances of computer-mediated communication (CMC), which encompasses social media channels, complicate identity processes due to the fact that individuals have more control over the self-presentational messages they exchange with others and the fact that many identity cues (such as gender and age), which are difficult to hide or alter face-to-face, are able to be masked or misrepresented in online contexts. It is much easier to engage to pretend to be someone you aren't - "gender-bending" for instance – online than in person. As a rule, people want to create positive impressions and will act in ways consistent with this goal, meaning they will take advantage of the opportunities presented by mediated communication to do so.

CMC refers to communication between individuals that takes place through or is facilitated by the use of computers; a commonly adopted definition of CMC is "communication that

takes place between human beings via the instrumentality of computers" (Herring, 1996). In this day and age, of course, "computers" may include a smart phone or other handheld device. CMC differs from face-to-face communication in several ways that have implications for interpersonal communication and identity processes; one key difference involves the ability to engage in selective self-presentation in CMC.

2.1.1 Selective self-presentation

Selective self-presentation refers to the fact that individuals can choose which identity cues they claim in online environments. The ability to share only the cues that one wants others to have, thus curating a desired impression, is enabled by two characteristics of CMC: asynchronicity and reduced cues. As Joseph Walther (2011) writes, "Online, one may transmit only cues that an individual desires others to have …. CMC senders may construct messages that portray themselves in preferential ways, emphasizing desirable characteristics and communicating in a manner that invites preferential reactions" (p. 461). Individuals may choose different strategies for enacting their online self-presentational practices. For instance, Livingstone (2008) finds that the way in which teenagers construct their online identities evolves over time, with younger teens emphasizing more aesthetic elements of their profiles and older teens aiming for more austere profiles that emphasize their connections with others.

2.1.2 Asynchronicity

Asynchronicity refers to the time delays inherent in many forms of mediated communication that give individuals more control over the messages they construct. In asynchronous media such as email and many forms of social media, individuals are able to carefully compose and edit messages -- even showing them to others for feedback before sending. The fact that messages are not exchanged simultaneously means that individuals have more control over their self-presentational messages – they can carefully select a profile photograph or even use photo manipulation software to embellish the image. Many social media applications include asynchronous communication channels and the act of creating a profile is almost always asynchronous. Ellison, Hancock and Toma (2012) found that online dating participants were aware of the temporal gap between when they created their profile and when it might be read, and used this to justify the creation of a profile that reflected characteristics that they might be able to achieve in the future.

Greater control over self-presentation is also enabled by the fact that mediated communication – where we don't see one another face-to-face – often masks or hides the physical cues to identity that are ascribed on the body, such as wrinkles that indicate age or skin tone which can indicate ethnicity. The *reduced cues* available in CMC means that identity information cannot be inferred and thus has to be explicitly communicated – sometimes accurately, sometimes not as accurately. Additionally, non-verbal cues that aid in interpreting communication, such as eye contact, tone of voice, and hand gestures, are not visible in text-based online communication. Because ascribed aspects of identity, such as gender, ethnicity and age, can be withheld or fabricated in online contexts, individuals can choose whether to disclose aspects of the self, both ascribed (such as race/ethnicity) or elective (such as an interest in motorcycles). This can have positive and negative implications for identity, as discussed below.

Early in the history of the Internet, there was some hope that the reduced cues context meant that visually triggered biases such as racism would be eliminated, but unfortunately this was not the case. People made assumptions about one another and found other biases to use. Nakamura (1999) found that in some online settings, when race and gender were not

mentioned, people assumed the writer was white and male, and calling attention to the fact that one wasn't was seen as making trouble by introducing what others saw as a divisive topic.

This process of selective self-presentation can have implications for relationship development. Walther's (1996) research on the hyperpersonal model argues that the selective self-presentational capabilities of asynchronous media enable interaction partners to form idealized or overly positive impressions of one another. This suggests that two individuals who encounter one another online and have an extended period of online interaction before meeting face-to-face are more likely to be disappointed than those that meet offline immediately, because they will have had more time to build up idealized impressions of one another. As more sites that bring together strangers incorporate technologies that minimize asynchronous communication (such as instant messaging or chat) and introduce more identity cues (such as video-chatting), we might expect to see different dynamics emerge. For instance, users who communicate over video might be more likely to have more realistic expectations about what their online communication partners look like and how they will behave in an offline setting.

2.2 Honesty and Deception in Online Self-Presentation

Selective self-presentation also has implications for how we feel about and see ourselves. In many situations, both online and offline, individuals curate positive impressions by withholding disclosures which might reflect poorly on them and sharing those that are more positive. In contrast to earlier forms of CMC that emphasized fantasy identities, such as MUDs, self-presentational messages in many forms of social media tend to be quite accurate. Social media sites that allow individuals to connect to offline acquaintances and friends (such as social network sites) tend to be more honest than ones in which individual profiles exist in a social vacuum (such as online dating sites), in part because users are broadcasting their disclosures to hundreds of people they know in offline context and these "Friends" might call attention to any fibs. In fact, research shows that individuals form similar initial impressions of others from Facebook profiles and from face-to-face interactions, suggesting that those with online profiles are providing relatively accurate portrayals of their offline selves (Weisbuch *et al*, 2009).

One strategy by which individuals can assess the credibility of online is through *warrants*. The concept of warranting, introduced by Sandy Stone (1995) and refined by Walther and Parks (2002) describes the way in which some kinds of online cues signal a link between one's online and offline self. The credibility of this cue, or the warranting value, is based on the extent to which a cue is believed to be susceptible to manipulation by the profile owner. For instance, a photograph of someone climbing a steep boulder would have more warranting value than just a textual statement about one's fondness for rock-climbing. Recent research has found that providing warrants such as one's real name, a real photograph, and shared offline social connections is associated with lower rates of deception in CMC channels (Warkentin *et al*, 2010).

Similarly, experimental work has found that when an individuals' self-reported level of attractiveness was contradicted by comments posted by their SNS contacts, participants tended to privilege the opinion held by others over self-reported information (Walther, Van Der Heide, Hamel, & Shulman, 2009)

One trend identified by Ellison and Boyd (in press) is the fact that profiles on SNSs are increasingly *co-created*. Although early incarnations of SNS profiles were typically static portraits occasionally updated by users, today's profiles consist of user-supplied content as well as activity reports, content supplied by other users, and system-provided content such as a list of their activities on third-party sites. They write, "Over time, the profile has shifted from a self-presentational message created by the individual to a portrait of an individual as an expression of action, a node in a series of groups, and a repository of self and other-provided data" (p. 154). This trend can be expected to continue in the future, and may have implications for identity processes in that social media profiles will increasingly contain identity markers contributed by others and behavioural data traces, in addition to information explicitly constructed by the profile owner.

2.3 Psychological Implications of Online Identity Work

The ability to emphasize positive presentation can have implications for self-concept (how we see ourselves) and self-esteem (how we feel about ourselves). In an experiment assessing the impact of viewing one's own Facebook profile vs. looking at oneself in a mirror on self-esteem, Gonzales and Hancock (2011) found that viewing and editing one's profile resulted in increases in self-esteem among college undergraduates. They speculate about the increased control over presentation, presumably enabling the production of more positive self-presentational messages, and write, "By allowing people to present preferred or positive information about the self, Facebook is a unique source of self-awareness stimuli in that it enhances awareness of the optimal self" (p. 82). In a related study, Toma (2010) argues that SNS profiles can serve to remind users of important aspects of their lives and thus looking at one's own profile can boost one's moral and feelings of self worth. The above studies suggest that curating a profile in social media may be a pathway towards increased self-esteem and self-awareness.

On the other hand, what are the effects of viewing profiles of other people? Users who view others' social media content may encounter feelings of inadequacy or lowered self-esteem when comparing themselves to the uniformly positive and self-promotional boasting of their peers, consistent with social comparison theory's premise that the social attractiveness of others impacts perceptions of one's own social attractiveness (Festinger, 1954). Given that SNS users may be motivated to selectively self-present online, those on these sites will be more likely to encounter positive, socially attractive information about their online peers than negative information. This over-representation of positively valenced information about others might lead to upward comparisons, encouraging users to feel poorly about his or her own accomplishments and characteristics.

2.3.1 Context collapse

Context collapse describes the possible complications associated with online self-presentation and identity management (Boyd, 2010; Marwick & Boyd, 2011). Context collapse is the process through which various connections representing different aspects of one's identity are flattened into a uniform, one-dimensional group such as "Friends" or "contacts." In offline contexts, we typically interact with a specific set of people in one particular place or occasion, which enables self-presentation specific to that context. But in the "networked publics" (Boyd, 2010) found in social media, it is difficult to segment distinct contexts and often one's network consists of individuals representing different aspects of one's identity. According to Vitak *et al*. (2012), professional adult users of Facebook engage in various strategies to manage context collapse, such as keeping professional contacts out of their Facebook network, creating multiple Facebook accounts, or avoiding sharing any information that could potentially damage their professional relationships. Farnham and Churchill (2011) approached the concept of identity online by characterizing humans' real and online identities as "faceted" and complex, arguing that each individual has different aspects of his or herself that they present to different individuals or groups of individuals, contingent on the social situation. Arguing against the notion of "singular" identities, Farnham and Churchill (2011) propose a theoretical framework that emphasized the shortcomings of social technology systems that operate as though users are one-faceted or as though all facets of one's life are integrated.

Users who feel more confident in their ability to use online tools and who have the knowledge to use these tools in ways that minimize the negative outcomes of context collapse and emphasize the advantages of a diverse social network are more likely to benefit from their use. Individuals' SNS use has been linked to levels of social capital – their perceptions about their ability to access information and social support from their social networks (Ellison *et al*, 2007; 2011). Of course, expressing a need (for information related to a job search, for instance, or support after the death of a relative) is a prerequisite for receiving these informational or social support benefits. If individuals respond to context collapse pressures by refusing to disclose anything on social media, they may be denying themselves these benefits. Knowing how to create different groupings of one's network of online Friends, through use of tools like Google's Circles, in order to target content to one group but not another would enable users to harness the benefits of their SNS use. There is an increasing focus on Internet skills as a determinant of the benefits one can access through their Internet use (e.g., Hargaittai, 2010).

When thinking about future implications of social media use on identity in terms of concepts like self-esteem, it is difficult to predict outcomes. Some research has found that those with higher self-esteem were more likely to benefit from their Facebook use (Ellison *et al*, 2007). On the other hand, other work has suggested that when it comes to positive outcomes of Internet use, the "rich get richer," as found in very early studies on Internet use and depression (Kraut *et al*., 1998). Those who are already prone to feelings of inadequacy may have these feelings amplified by viewing the overwhelmingly positive posts of their friends. This could be further problematized by web content, sites or services that feed upon feelings of inadequacy, such as "pro-ana" and "pro-mia" sites. These websites support anorexia and bulimia as "lifestyle choices" and often include "thinspiration" -- images of dangerously thin women -- and tips for hiding one's disease and ignoring signs of hunger or starvation. In some cases, engaging with these communities may serve to prolong harmful behaviours, because individuals are able to connect with others who celebrate these values and delegitimize those who try to stigmatize them. Other examples of this dynamic can be found in pro-white, right-wing extremist online communities, such as the "Stormfront" website,

where individuals who feel like they are stigmatized in their offline interactions value the site more than those who do not and see it as a refuge (De Koster & Houtman, 2008). Although there are examples of stigmatized populations coming together for positive outcomes, such as information-sharing, social support, and political mobilization, it is important to remember that these opportunities are available to all groups, even ones that have goals that are generally accepted as dangerous for society at large.

3. Multiple Identities Online

M*ultiple identities* are enabled in online spaces and these identities may or may not correspond to one's offline identity. For instance, one person can create multiple accounts on different sites, or even the same site, each of which reflect a different aspect or facet of their identity. As Nancy Baym (2010) writes, "In lean media, people have more ability to expand, manipulate, multiply, and distort the identities they present to others."

Early forms of social media that enabled identity play, such as MUDs (Multi-User Dimensions or Dungeons), were embraced by participants who created fantastical descriptions of the online characters they inhabited. Sherry Turkle (1996) framed this as a freeing, liberating experience when she wrote, "The anonymity of MUDs gives people the chance to express multiple and often unexplored aspects of the self, to play with their identity and to try out new ones. MUDs make possible the creation of an identity so fluid and multiple that it strains the limits of the notion" (Turkle, 1996).

Newer forms of social media may discourage identity play. On some SNSs, such as Facebook, individuals are encouraged to have one profile that closely corresponds to their offline identity. Facebook's terms of service, in fact, specify that users create only one account and that users won't provide any "false personal information" to the site (Facebook, 2012). Profiles on SNSs like Facebook are connected to other accounts that the user affiliates with, creating a "social graph" that captures these webs of relationships. Some social media sites allow users to participate without associating their activities with a profile, as when a user watches a video on Youtube without logging in; their activities may still be tracked by advertisers or the site, but are not visible to other users and typically activities such as rating or commenting cannot be done without logging in. In other cases, users may use a pseudonymous account for participating, as when a user logs into her administrator account on Wikipedia and edits an entry. Of course, users can strategically choose to identify themselves or not, depending on their particular goals and activities. For instance, a user may browse stories on the discussion site Reddit while at work without logging in; log into a long-term account to post a story likely to result in reputation "karma"; or create a fake "throwaway" account to share a personal experience she does not want associated with her "real" long-term account, which may or may not be linked to her offline identity.

In some online contexts, taking advantage of the ability to create new online personae that may not reflect one's offline self is seen as deceptive and problematic. This is particularly true in contexts where this linkage is critical. For instance, in a discussion forum where medical professionals exchange information about health treatments and diagnoses, falsely claiming to have medical training would be considered a breach of trust. Similarly, on online dating sites, where individuals are looking for romantic partners and want to assess the extent to which they are attracted to a particular online partner, lying about one's appearance could be alienating and counter-productive.

The ability to hold multiple profiles that are not associated with one's "real name" is an evolving discussion among those in the tech industry and users. Google+, for instance, began to enforce a policy whereby users were only allowed to use their actual name on their profile by deleting accounts that didn't adhere to their name restrictions. This was concerning to many users who had reasons for not wanting to be identified, such as victims of stalking or assault, activists, gay and lesbian people, and others (Boyd, 2011). There are many legitimate reasons why individuals would want to engage in interactions online that are not linked to their offline identity. For instance, anonymity can enable more honest self-disclosure, which is an important precursor to accessing social capital benefits, as those who do not share events like a health diagnosis are less likely to receive social or informational support (Ellison *et al*, 2011). Similarly, anonymity can potentially engender more open information dynamics, which is why protecting the anonymity of its sources is an important goal of Wikileaks (Wikileaks, 2012).

On the other hand, real name policies are attractive to social media companies, who recognize that being able to identify a specific user enables more targeted and relevant marketing messages. Knowing that a particular user is interested in estate planning is useful to a financial institution, but knowing that this user also shops at high-end online boutiques, frequently logs in from an ip address in Japan, or spends a lot of time on mortgage default and couponing sites is even more useful. Being able to link a particular user with their behaviour across the web is extremely valuable to online sites and those that advertise on them.

Of course, for society, enabling anonymous communication in social media platforms can have negative implications. When people are anonymous, they tend to act differently and adhere less to social norms. Research shows that anonymity can prompt more honest self-disclosures (Joinson & Paine, 2007) but it can also result in deindividuation and enable destructive behaviours (Zimbardo, 1970). One example of this in the social media realm is trolls. Trolls are individuals who act in ways intended to upset others. For instance, some trolls look for memorial pages on social media sites – pages dedicated to people who have passed away and typically created as a place for people to share positive memories of the deceased and to exchange social support. Trolls will post deliberately inflammatory messages to these sites, such as gory images they say are pictures of the cause of death (e.g., car wrecks), upsetting information allegedly about the deceased, or insults directed at mourners. Social media memorial pages offer an interesting artefact to consider when thinking about the costs and benefits of anonymous communication online. Although it is difficult to understand the motivation behind this behaviour, interviews with trolls suggest that some trolls are reacting to "grief tourists" and the public nature of what they believe should be a more private experience (Phillips, 2011). In the UK, proposed legislation aimed at curtailing defamatory messages in online forums offers greater protection to websites who help identify trolls and other anonymous online participants who post problematic messages (Guardian, 2012).

Social media present challenges in the realm of identity questions, and many of these issues are likely to be exacerbated in the upcoming decade. Processes of impression management are complicated by the kinds of data that are being captured and shared about our online behaviour 24/7, sometimes without users being fully aware of what is being captured and who it is being shared with. Advertisers are one obvious audience, but as more applications take actions on our behalf – such as posting updates for us – the potential for individuals to inadvertently "leak" identity information is heightened, and context collapse issues make it more likely that we may be embarrassed or hurt by these disclosures.

3.1 Identity and Networked Publics

Above and beyond the characteristics of CMC described above, social media has specific affordances that are unique to newer communication technologies. Danah Boyd (2010) discusses four affordances of digital content in networked publics:

- "Persistence: online expressions are automatically recorded and archived.
- Replicability: content made out of bits can be duplicated.
- Scalability: the potential visibility of content in networked publics is great.
- Searchability: content in networked publics can be accessed through search."

These characteristics complicate the audiences for our utterances via social media, and in many ways make it impossible to identify the actual audience.

Sharing via social media involves speaking to a networked audience consisting of known and unknown readers – many of whom may be connected to one another. As Marwick (2010) writes, "The networked audience is the real or imagined viewers of digital content who are connected to the content creator and each other. Many Web 2.0 sites digitize formerly ephemeral social information, causing all manner of complicated social problems as this information moves across boundaries and contexts" (p. 8). Marwick explicates how this audience is a critical component of identity construction in online spaces, quoting Markham's (2005) assertion that "The common phrase 'I think, therefore I am' is woefully inadequate in cyberspace. Even 'I speak, therefore I am' is not enough. In cyberspace, the more appropriate phrase is 'I am perceived, therefore I am.'" Without an audience, identity as enacted via social media does not, by many accounts, exist.

The persistence of social media utterances, as explicated by Boyd, is not always apparent to creators of social media content. Many Twitter users were surprised to learn that their postings to the site would be archived at the U. S. Library of Congress, although Twitter later clarified that it wouldn't include private tweets and that these archives would only be available to researchers (Parry, 2010). Some users, concerned about the privacy implications of their social media use, employ technical and social workarounds.

For instance, one application, Snapchat, sets an “expiration date” on photos so that they will delete themselves after a specific amount of time (Wortham, 2012). But not all users are savvy enough to know about these workarounds or to effectively use them.

The persistence of personal information can have identity implications for younger users. Although children younger than 13 are not technically supposed to have accounts on SNSs, a 2010 report found that 46 of children at age 12 in America use social network sites (Lenhart, Purcell, Smith, & Zickuhr, 2010) and that in many of these cases, parents have knowingly helped create accounts for their children (Boyd, Hagittai, Schulz, & Palfrey, 2011). These children are sharing photographs and textual information that may follow them into adolescence and beyond. In other cases, parents are sharing information about their children before they are of age to consent. Implications for identity include the fact that future audiences will have access to these materials, potentially limiting users’ ability to enter into adulthood with the freedom to reinvent themselves as younger generations did.

4. Data, Computation, and Identity

Social media often capture behavioural data from users, and advanced computational techniques can mine these data for valuable insights into consumer behaviour, sometimes leading to complications regarding issues of privacy and the ability to be identified. Throughout history, individuals have been linked with a persistent identity, associated with a name and ID number. In online contexts, individuals can create multiple accounts on a variety of sites each of which reflect a different facet of one's identity.

Companies sometimes anonymize user data and then sell or share them, but even anonymized datasets, such as those released by Netflix, have been re-identified by researchers using fairly simple techniques, such as comparing the anonymous movie ratings with those submitted to a public website (Porter, 2008).

4.1 Big Data

Regardless of the extent to which these technologies are incorporated into social media applications, all social media have the option of saving and analyzing behavioural data about their users. The behavioural traces produced by social media applications -- "big data" – are a treasure trove for researchers. Harnessing social media activity data for purposes of targeted marketing or relationship building is a growing business. In contrast to earlier approaches which viewed social media as another channel for broadcasting messages, companies now realize that the information produced by users can be useful. Being able to identify customers is a key part of this process because it enables companies to better target individuals for marketing purposes and to learn more about their customers in order to deepen the relationship (Boorman, 2011).

To the extent that individuals produce content online, they may be identified. Researchers can create a "writeprint" (as opposed to a fingerprint) for each participant based on his or her emails – this is a unique pattern based on things like common typographical errors, word choice, and punctuation patterns. Researchers can correctly identify the author of an email 80 percent of the time using these cues (Goodin, 2011).

Users may have accounts on different sites across the web, both personal and professional, and may have legitimate reasons to keep some accounts separate from other accounts – or their offline identity. However, being able to link disparate identities (which facilitates better user profiling and thus more relevant advertising) is very attractive to many sites. Some sites even change pricing depending on what they know about a user. For instance, an e-commerce site might show a lower (or higher) price to someone whose digital dossier shows they have been to numerous price comparison sites (Ramasastry, 2005).

New techniques such as sentiment analysis mean that information is not limited to customers explicitly "liking" or "recommending" a particular brand, but rather that any text produced by a human can be analyzed for more subjective emotions, such as the extent to which a given utterance is positive or negative, or its intensity or subjective bias.

This is a growing field of great interest to marketers (Wright, 2009). The collection of online activity data is amplified by the fact that mobile devices accompany individuals everywhere and are increasingly embedded in many of their daily activities.

4.2 Quantified Self

Of course, the behavioural traces we produce are not just of interest to third parties – they can also shape our sense of self and thus have identity implications. Individuals can get a sense of how powerful or influential they are by tracking their Klout score, for instance. This is part of a larger trend towards the "quantified self": the self-tracking movement in which individuals continuously chart aspects of their bodies, their behaviour, or their moods, usually through technology or devices of some kinds – and chart their change over time or movement towards a given goal. There are apps that share this with others as well, for instance, posting one's blood sugar level to Twitter. After the data are collected, users can use various tools to visualize them or analyse them. Posting one's statistics to a larger audience can help individuals stay focused on a particular goal, and many applications have this functionality built in. For instance, running applications like Nike+ and RunKeeper track one's runs and offer users the ability to chart their longest or fastest run. These apps also have a social media component whereby users can share their statistics with their Facebook Friends, who might offer encouragement or healthy competition.

The rapid rate of technological innovation means that individuals are less likely to be able to keep up with changes that affect their ability to manage their online identity and privacy issues. These changes include location aware devices and face recognition software.

4.2.1 Location-aware devices and applications

Many online sites are being accessed via mobile devices, which are increasingly equipped with location-awareness capabilities. Users may be broadcasting their location willingly, but without realizing the ways in which information from different sites can be pieced together to discover sensitive information. For instance, the "Girls Around Me" application uses public Foursquare data (which identifies where users are based on their "check-ins" to various locations) with public Facebook profile data (which might include profile pictures, profile information, and contact information) to let users know who is around them, what they look like, and what topics of conversation are likely to be successful. Although the app uses public information, it is combining them in ways that users may not expect or wish. In other cases, users may be unaware of the fact that they are transmitting location information to others. Geo-locational data from multiple users can be aggregated by sophisticated software to determine identity information about users; for instance, people who check into the same venue at the same time are more likely to know one another.

4.31 Face recognition software

Technological innovations in face recognition software have reached the point where it is possible to identify individuals based on photographs from public sites. In one study, researchers could identify one out of ten members of an online dating site by using face recognition software to link the (public but de-identified) photograph on an online dating site to a Facebook profile and were able to find the Facebook profile of approximately one out of three students on a college campus, using an image taken with an inexpensive webcam (Acquisti *et al*, 2011). Because users often incorporate a photograph in their profiles, these photographs can provide a way to link different profiles across the web. If an individual includes a photograph in profiles created on different sites across the web, these profiles can be linked. If her identity is revealed on one site (such as a professional networking

website), it can then be linked to profiles on sites where she would prefer to be pseudonymous. We may see more face recognition technology incorporated into social media applications in the future; Facebook recently announced it was purchasing Face.com, a facial recognition software company (Reisinger, 2012).

5. Conclusion

In summary, social media offer us new ways to connect with one another, present ourselves, and other activities that are critical to identity formation and expression. Although we can't predict how technological innovation will evolve over the next decade, we do know that individuals will be using these tools in ways that enable them to meet their relational, self-presentational, and identity-based goals.

Although they offer new ways of controlling self-presentational messages, it is clear that online and offline modes of communication for many people constitute aspects of a larger communication ecology. Although social media reflect new possibilities for identity-related activities, the online realm is not a separate sphere of activity that reflects "the real world" as distinct from the "virtual." Rather, how individuals behave online and offline reflect dimensions of the self, and online activities are as "real" as those that take place offline (face-to-face). Online identity performances differ from offline performances due to the affordances of CMC, such as the selective self-presentation that is enabled by the asynchronous, reduced cues context of CMC. However, in contrast to many earlier forms of CMC, individuals on social media are embedded in a social network that consists primarily of people they know offline (Boyd & Ellison, 2007), potentially dampening opportunities for identity fabrication or exploration.

Computational advancements and the behavioral traces left on social media are making it easier to link identities across site, which may have problematic implications for users who have legitimate reasons for remaining unlinked to their offline identity. The ability for users to contribute anonymously to sites that are sources of information and social support should be protected.

Although the research presented here provides evidence of both positive and negative potential outcomes of social media use, public perceptions about outcomes of social media use are generally uninformed, mirroring gaps in the literature. Much of the popular narratives circulating about social media focus on privacy concerns, but fewer discuss the relational, informational, and self-esteem related benefits to social media use. By the same token, the fact that these tools are so engaging is also deserving of study. Turkle (2011) expresses concerns about the ways in which our communication tools distance us from one another because we are "alone together" – in the same room but using our devices to communicate with others or engage in other tasks. How the increasing use of social media will affect our face-to-face interactions is a topic for future research.

References

Acquisti, A., Gross, R. and Stutzman, F., (2011) Privacy in the Age of Augmented Reality. In *BlackHat USA Conference.* Las Vegas, Nevada, 3-4 August 2011.

Baym, N., (2010) *Personal Connections in the Digital Age*. Cambridge, UK: Polity.

Boorman, C., (2011) Why data mining is the next frontier for social media marketing. *Mashable,* 25 February 2011, [online] Available at: http://mashable.com/2011/02/25/data-mining-social-marketing/ [Accessed 7 June 2012].

Boyd, D., (2010) Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In: Z. Papacharissi, ed., 2010. *Networked Self: Identity, Community, and Culture on Social Network Sites.* pp. 39-58.

Boyd, d., (2011), August 4. “Real Names” Policies Are an Abuse of Power. [online] Available at: http://www.zephoria.org/thoughts/archives/2011/08/04/real-names.html [Accessed 1 June 2012].

Boyd, d., Hargittai, e., Schultz, J., and Palfrey, J., (2011) Why parents help their children lie to Facebook about age: Unintended consequences of the ‘Children’s Online Privacy Protection Act’. *First Monday*, 16(11), [online] Available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/3850/3075 [Accessed 12 June 2012]

De Koster, W., and Houtman, D., (2008) "Stormfront is like a second home to me": On virtual community formation by right-wing extremists. *Information, Communication, & Society,*11, pp.1155-1176.

Dutton, W. H., and Blank, G., (2011) Next generation users: The Internet in Britain. Oxford Internet Survey 2011 Report. Oxford Internet Institute, University of Oxford.

Ellison, N. B., Hancock, J. T., and Toma, C., (2012) Profile as Promise: A Framework for Conceptualizing Veracity in Online Dating Self-Presentations. *New Media & Society* 14(1), pp.45-62.

Ellison, N. Steinfield, C. and Lampe, C., (2007). The benefits of Facebook 'friends': Exploring the relationship between college students' use of online social networks and social capital. *Journal of Computer-Mediated Communication,* 12(3).

Ellison, N.B., Steinfield, C. and Lampe, C., (2011). Connection Strategies: Social capital implications of Facebook-enabled communication practices. *New Media & Society*, 13(6), pp.873-892.

Ellison, N.B., Vitak, J., Steinfield, C., Gray, R. and Lampe, C., (2011) Negotiating privacy concerns and social capital needs in a social media environment. In: S. Trepte and L. Reinecke, eds., 2011. *Privacy Online: Perspectives on Privacy and Self-Disclosure in the Social Web.* Heidelberg and New York: Springer.

Facebook, (2012) June 8. “Statement of Rights and Responsibilities”, Facebook. [online] Available at: <http://www.facebook.com/legal/terms> [Accessed 10 June 2012].

Farnham, S. D. and Churchill, E. F., (2011) Faceted identity, faceted lives: social and technical issues with being yourself online. In: *Proceedings of the ACM 2011 conference on Computer supported cooperative work (CSCW '11).* New York, NY, USA: ACM. pp 359-368.

Festinger, L., (1954) A theory of social comparison processes. *Human Relations,* 7, pp.117-140.

Goffman, E., (1959) *The presentation of self in everyday life.* New York, NY: Anchor.

Gonzales, A. L. and Hancock, J. T., (2011) "Mirror, Mirror on my Facebook Wall: Effects of Exposure to Facebook on Self-Esteem." *Cyberpsychology, Behavior & Social Networking, 14(1-2),* pp.79-83.

Goodin, D., (2011). Data-mining technique outs authors of anonymous email. *The Register,* 15 March 2011, [online] Available at: http://www.theregister.co.uk/2011/03/15/anonymous_email_algorithm/ [Accessed 12 June 2012].

Guardian, (2012) Internet trolls targeted in new bill to tackle defamation online. *Guardian* [online]12 June. Available at: http://www.guardian.co.uk/law/2012/jun/12/internet-trolls-bill-defamation-online [Accessed 13 June 2012].

Hargittai, E., (2010) Digital Na(t)ives? Variation in Internet Skillsand Uses among Members of the “Net Generation”. *Sociological Inquiry*, *80(1),* pp.92-113.

Herring, S., (1996) *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. Amsterdam: John Benjamins.

Joinson, A.N. and Paine, C.B., 2007. Self-Disclosure, Privacy and the Internet. In: A.N Joinson, K.Y.A McKenna, T. Postmes and U-D. Reips, eds. 2007. *Oxford Handbook of Internet Psychology*. Oxford University Press, pp. 237-252.

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., and Scherlis, W., (1998) Internet paradox. A social technology that reduces social involvement and psychological well-being? *American Psychologist* , 53(9), pp.1017-1031.

Lenhart, A., Purcell, K., Smith, A., and Zickuhr, K., (2010) Social media and young adults. Report by the Pew Internet and American Life Project, 3 February 2010, [online] Available at: http://pewinternet.org/Reports/2010/Social-Media-and-Young-Adults/Part-3/1-Teens-and-online-social-networks.aspx [Accessed 12 June 2012].

Livingstone, S., (2008) Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. *New media & society, 10(3)* pp. 393-411.

Marwick, A. (2010) Status Update: Celebrity, Publicity and Self-Branding in Web 2.0. Dissertation, New York: New York University.

Marwick, A., and Boyd, D. (2011) I Tweet Honestly, I Tweet Passionately: Twitter Users, Context

Collapse, and the Imagined Audience. *New Media & Society,* 13(1), pp. 114–133.

Nakamura, L., (1999) Race In/For Cyberspace: Identity Tourism and Racial Passing on the Internet. In: V. Vitanza, ed., 1999. *Cyberreader*. Boston, MA: Allyn and Bacon.

Parry, M., (2010) Library of Congress, facing privacy concerns, clarifies Twitter archive plan. *The Chronicle of Higher Education,* 7 May 2010, [online] Available at: http://chronicle.com/blogs/wiredcampus/library-of-congress-facing-privacy-concerns-clarifies-twitter-archive-plan/23818 [Accessed 12 June 2012].

Phillips, W., (2011) LOLing at tragedy: Facebook trolls, memorial pages and resistance to grief online. *First Monday,* 16(12), [online] Available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3168/3115 [Accessed 7 June 2012]

Porter, C. C., (2008) De-identified data and third party data mining: The risk of re-identification of personal information. *Shidler Journal of Law, Communication, and Technology, 5(3)*.

Ramasastry, A., (2005) Web sites change prices based on customers' habits. *CNN (Justice),* [online]
Available at: http://articles.cnn.com/2005-06-24/justice/ramasastry.website.prices_1_price-differentials-price-discrimination-customers?_s=PM:LAW [Accessed 7 June 2012].

Reisinger, D., (2012). Facebook buys facial-recognition startup. *C|NET,* 19 June 2012, [online] Available at: http://asia.cnet.com/facebook-buys-facial-recognition-startup-62216765.htm [Accessed 19 June 2012].

Senft, T., (2012) "Micro-celebrity and the Branded Self." In: J. Hartley, J. Burgess, and A. B. Blackwell, eds., 2012. *A Companion to New Media Dynamics*, Malden, MA: Wiley-Blackwell Aslinger, B. S.

Sundén, J., (2003) *Material Virtualities*. New York: Peter Lang.

Turkle, S., (1996) "Who am we?", *Wired*, pp.148-152.

Turkle, S., 2011. *Alone together: Why we expect more from technology and less from each other.* New York: Basic Books.

Vitak, J., Lampe, C., Gray, R., and Ellison, N. B., (2012) "Why won't you be my Facebook Friend?": Strategies for managing context collapse in the workplace. Paper presented at the iConference, Toronto, Ontario, Canada.

Walther, J. B., (2011) Theories of Computer- Mediated Communication and Interpersonal Relations. *In:* Knapp, M. L. and Daly, J. A. (eds.) *The Handbook of Interpersonal Communication, 4th ed*. Thousand Oaks, CA: Sage.

Walther, J. B., 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research, 2*3**,** pp.3-43.

Warkentin, D., Woodworth, M., Hancock, J. T., and Cormier, N., (2010) Warrants and deception in computer-mediated communication. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, pp. 9–12. New York: ACM.

Weisbuch, M., Ivcevic, Z., and Ambady, N., (2009) On being liked on the web and in the "real world": Consistency in first impressions across personal webpages and spontaneous behavior. *Journal of Experimental Social Psychology,* 45*,* pp.573-576.

Wikileaks, (2012) About: What is Wikileaks? [online] Available at: http://wikileaks.org/About.html [Accessed 10 June 2012].

Wortham, J., (2012) Privacy, please: This is only for the two of us. *The New York Times,* 2 June 2012, [online] Available at: http://www.nytimes.com/2012/06/03/technology/apps-aiming-for-sharing-within-boundaries.html?smid=pl-share [Accessed 7 June 2012].

Wright, A., (2009) Mining the Web for feelings, not facts. *The New York Times,* 23 August 2009, [online] Available at: http://www.nytimes.com/2009/08/24/technology/internet/24emotion.html [Accessed 12 June 2012].

Zimbardo, P. G., (1970) The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In: W.J. Arnold and D. Levine, eds. 1969. *Nebraska Symposium on Motivation.* Lincoln: University of Nebraska, pp. 237–307.

Foresight
1 Victoria Street
London SW1H 0ET
www.foresight.gov.uk
URN 13/505