

This version is a preprint. Please cite following version:

Smith, C. H.. (2020). 'Corporatised Identities ≠ Digital Identities: Algorithmic Filtering on Social Media and the Commercialisation of Presentations of Self'. In C. Burr and L. Floridi (Eds.) *Ethics of Digital Well-Being*. Springer.

Corporatised Identities ≠ Digital Identities: Algorithmic Filtering on Social Media and the Commercialisation of Presentations of Self

Author:

Charlie Harry Smith

Oxford Internet Institute
University of Oxford
1St Giles, Oxford,
OX1 3JS, United Kingdom

charlie.smith@oi.ox.ac.uk

ORCID 0000-0001-9832-1723

Abstract

Goffman's (1959) dramaturgical identity theory requires modification when theorising about presentations of self on social media. This chapter contributes to these efforts, refining a conception of digital identities by differentiating them from 'corporatised identities'. Armed with this new distinction, I ultimately argue that social media platforms' production of corporatised identities undermines their users' autonomy and digital well-being. This follows from the disentanglement of several commonly conflated concepts. Firstly, I distinguish two kinds of presentation of self that I collectively refer to as 'expressions of digital identity'. These *digital performances* (boyd 2007) and *digital artefacts* (Hogan 2010) are distinct, but often confused. Secondly, I contend this confusion results in the subsequent conflation of corporatised identities – poor approximations of actual digital identities, inferred and extrapolated by algorithms from individuals' expressions of digital identity – with digital identities proper. Finally, and to demonstrate the normative implications of these clarifications, I utilise MacKenzie's (2014, 2019) interpretation of relational autonomy to propose that designing social media sites around the production of corporatised identities, at the expense of encouraging genuine performances of digital identities, has undermined multiple dimensions of this vital liberal value. In particular, the pluralistic range of authentic preferences that should structure flourishing human lives are being flattened and replaced by commercial, consumerist preferences. For these reasons, amongst others, I contend that digital identities should once again come to drive individuals' actions on social media sites. Only upon doing so can individuals' autonomy, and control over their digital identities, be rendered compatible with social media.

Keywords: Goffman, Identity, Relational Autonomy, Performance, Digital Artifact, Social Media

Introduction¹

As the online/offline distinction has blurred, digital identities have become part of daily life (Floridi 2011a, p. 477; Hongladarom 2011, p. 534), drawing critical attention to the “construction of personal identities in the infosphere” (Floridi 2011b, p. 550). Indeed, most liberal democratic citizens now remain constantly connected to the internet through their smartphones, continually contributing to and maintaining various online personas. Although not the only sites where such digital identities play an important role, they are perhaps most closely associated with social media platforms (Ellison and boyd 2007, p. 210) – sites like Facebook, Instagram, YouTube and Twitter.

Accordingly, this chapter’s primary aim is one of precision: to refine a conception of digital identities and clarify their role(s) on modern social media sites. This objective is complicated by various usages of the term ‘digital identity’ in the literature so², to be clear, I will utilise a *performative* account of digital identity, following in Goffman’s (1959) ‘dramaturgical’ footsteps, to consider how social media can affect individuals’ presentations of their identities online. This requires considering digital identities in a highly subjective and personal manner: as tapestries of intersubjective experiences, woven from ongoing presentations of self. Such identities are therefore understood to be inherently unique and always fluctuating, changing and reacting as they are continually performed, re-negotiated and re-presented to numerous audiences.

This approach allows me to distinguish what I will argue are digital identities ‘proper’ from the substandard *approximations* of digital identities – which I term ‘corporatised identities’ – that social media companies covertly and algorithmically produce once they have identified individuals. Although often confused with one another, these two kinds of identity are distinct, and I will suggest that their conflation obscures fundamental, important questions surrounding the self and personal identities as they are continuously performed and constructed online.

However, understanding how and why corporatised identities are being conflated with digital identities begins with recognising the differences between *digital performances* (boyd 2007) and *digital artefacts* (Hogan 2010) which, together, I term ‘expressions of digital identity’. Whilst these interrelated elements both express individuals’ identities and can be understood to be “presentations of self” (boyd 2007, p. 128), they do not contribute equally to individuals’ digital identities. Consequently, in the first two sections, I elucidate the differences between all these

¹ My thanks to Kai Spiekermann, Christopher Burr, and Laura Valentini for feedback on earlier drafts of this chapter. I am also grateful for Jaimie-Lee Freeman’s advice regarding conceptions of digital well-being.

² For instance, ‘digital identities’ can refer to merely technical tools for *identification* (Whitley et al. 2014, p. 18).

various related and often confused presentations of self (§1) and types of identities (§2). This chapter's key theoretical contribution is thus the disentanglement of these concepts.

This conceptual work offers real value to the political theorist. To illustrate, in the third and final section, I first draw on MacKenzie's (2014, 2019) relational account of autonomy to argue that designing social media sites around the algorithmic production of corporatised identities, rather than encouraging the performance of digital identities, has undermined multiple dimensions of this vital liberal value, including self-governance, self-determination, and self-authorisation. Then, I utilise the eudaimonic account of well-being to frame this as a harm to individual functioning (Devine et al. 2008)³. In particular, I argue that the pluralistic range of authentic values and preferences that should structure flourishing human lives are being flattened and replaced by commercially-motivated, consumerist preferences. It is for these reasons, amongst others, that I contend digital identities should come once more to drive individuals' actions on social media sites.

In summary, this chapter has two main aims. Firstly, to reframe the debate surrounding the social media industry in terms of its negative effects on digital identities, not mere expressions of digital identity or the corporatised identities inferred from those expressions. And, secondly, from this new vantage point, to argue that (relational) autonomy, particularly over digital identity formation, is currently being impaired by social media companies and the algorithms that drive their systems. The conclusion then readily follows that the continued production of corporatised identities will be deeply damaging for individuals on a eudaimonic account of digital well-being.

³ The *eudaimonic* approach to well-being emphasises the individual's flourishing and functioning, standing in contrast to *hedonic* approaches that focus on their subjective happiness alone (Ryan and Deci 2001).

§1 – Expressions of Digital Identity

§1.1 – Digital Performances

Preliminary examples of digital performances include posts, status updates, photos, likes, livestreams, tweets and retweets, as well as purchasing choices, ad clicks and many more interactions. Unsurprisingly, these kinds of expressions of digital identity bear strong resemblance to Goffman's (1959) notion of a *performance* in the analogue world, whereby social interactions are metaphorically understood to be theatrical performances delivered to an audience⁴. For Goffman, performances amount to “all the activity of an individual which occurs during a period marked by his continuous presence before a particular set of observers and which has some influence on the observers” (Goffman 1959, p. 13). Such performances essentially aim to impress a particular identity upon an audience, and individuals then work to maintain that impression over time and ensure that their performances will be acknowledged as genuine (Goffman 1959, p. 10). Identity is therefore inherently social both because it is relational by nature (between actor and audience) and because performances are continually being “negotiated” with the audience (as those involved figure out how they relate to one another and try to shape how they are treated) (Phillips 2009, p. 304).

Importantly, however, individuals regularly perform many *different* identities depending on the situation, with various identities (or combinations of identities) being more or less salient in a given context (Davis 2016, p. 139). Individuals can thus be understood to wear different “masks” in varying situations that emphasise different aspects of their multiple identities (Bullingham and Vasconcelos 2013, p. 101). And, in much the same way, when performing online, individuals attempt to do the same (boyd and Heer 2006, p. 4).

Asserting a *digital* identity also involves the performance of an idealised identity to an audience (Marwick and boyd 2011, p. 114). When it comes to self-presentation, social media users personalise their profile pages, choose photos and write pithy ‘bios’ that describe themselves, and generally attempt to present a coherent identity (Marwick and boyd 2011, p. 115). Furthermore, the content they actually perform takes specific forms (e.g. livestreaming, memes, etc.) that will be appropriate for the identity they are adopting. Online, this is particularly noticeable when considering the numerous personae that individuals create and maintain on different platforms –

⁴ On social media audiences are often “imagined” as it is frequently impossible to know who will witness an individual's performances (Marwick and boyd 2011, p. 115).

the sanitised performances most people give on LinkedIn, for example, are likely to be completely different to those in a WhatsApp chat with close friends.

All of this performative work results in a unique impression of a digital identity being formed with and for the audience in question. Although mediated by technological means, digital identities thus have much in common with their analogue counterparts: they are the result of ongoing social processes of negotiation that stem from a struggle to present and maintain a particular impression of oneself in the eyes of another. Each identity is thus highly individualised, distinct and contextual, made up of the entirety of the relevant online interactions that have occurred between a given set of individuals up until that point⁵. They are, in other words, far more than the sum of their digital parts, a social impression that results from concerted efforts to present a particular version of the self.

Nonetheless, a corollary of digital identities constantly evolving is that attending to each distinct performance in isolation – looking at a single tweet or status update – only provides part of the fuller picture. Each is only ever one expression of a salient identity (or several identities) in that specific context; one quarantined scene that constitutes only a slice of a larger part that is being performed over time, and which responds to audience reactions (Goffman 1959, p. 5). “Performances”, as Schieffelin argues (1998, p. 198), “are a living social activity [...] While they refer to the past and plunge towards the future, they exist only in the present”. Identities are thus always being performed more or less effectively, with a fuller picture of each digital identity only emerging through repeated successful performances. This means that whilst a particular performance might therefore have contributed effectively towards an identity when first performed, it can easily fail to continue doing so when removed from its original situation, altering the meaning of that performance and its relevance to the individual. In the context of modern social media, this is not unproblematic, but to see why we must consider the other half of expressions of identity.

§1.2 – Digital Artefacts

Goffman chiefly investigated social interactions that occurred when “in one another’s immediate physical presence” (Goffman 1959, p. 8), so his work needs modification to remain relevant for internet communications. Largely, this requires acknowledging that the digital traces of performances which are recorded by social media platforms actually fail to meet Goffman’s

⁵ This thumbnail sketch of a digital identity should suffice until the concept is more thoroughly fleshed out in §2.3.

criteria for a performance (Hogan 2010, p. 377). Whilst these digital remnants of a performance are still undeniably “a form of presentation of self” (Hogan 2010, p. 377), once later accessed and processed it would be inaccurate to call them performances. Instead, they are better understood (metaphorically-speaking) as online “exhibitions” made up of stored performances that are recalled and re-presented to audiences as required – a new kind of expression of digital identity that Hogan terms “artifacts”⁶ (Hogan 2010, p. 381). Thus, although similar, digital performances and digital artefacts are not in fact identical.

This terminological change is primarily required for two reasons. Firstly, stored artefactual performances are endlessly re-presented by social media companies to different audiences, not just the audience to which the performance was originally delivered, and, secondly, this re-presentation is done by algorithmic “curators” that make complex and opaque decisions when selecting what to exhibit (Hogan 2010, p. 380). For instance, the audience that watches a videogaming live-stream in the moment are no doubt exposed to a full-on performance, but if a recording is made available to view a year later, it is likely that a completely different audience is being shown that artefact, selected by an algorithm trying to maximise engagement⁷ (Hogan 2010, p. 381).

Exhibitions are therefore not made up of true digital performances but assembled from the digital artefacts of what *used to be* performances: the digital traces of an individual’s identity as it was at a particular moment. Consequently, whilst earlier researchers recognised that without “live” updates social media profiles became “frozen performances” and “outdated representations of the self” (boyd and Heer 2006, p. 9), Hogan successfully formulates the need to move beyond performances altogether and extricates the similar, but distinct, concept of a digital artefact.

Reconsidering the preliminary examples of performances that began our discussion, we can now see that many decay in just this way. Still, in my taxonomy, both digital artefacts and digital performances are merely *expressions* of digital identities which, when taken together, are not equivalent to digital identities themselves (See Fig. 1). This is because digital identities emerge *from* digital performances but are not reducible to them. Nonetheless, when considering social media and its normative issues, analysis often focuses on these expressions of digital identity, taken in part or as a whole, as these are easy to access and link back to an identified individual⁸. After all, it is now trivial for social media companies to harvest and store vast stores of digital data. But,

⁶ Hogan uses the American spelling, but I have opted for the British ‘artefacts’.

⁷ E.g. Facebook’s ‘memories’: curated content from an individual’s past which generated the most engagement at the time is selectively re-presented to them today to share to new audiences. For more, see: (Hod 2018).

⁸ Consider, for example, studies charting racism on Twitter (Chaudhry 2015) or fake news on Facebook (Bakir and McStay 2018) that focus on expressions of digital identity rather than individuals’ actual digital identities.

unfortunately, much of the analysis also ends here, conflating digital performances and digital artefacts despite them not being equally relevant to digital identities.

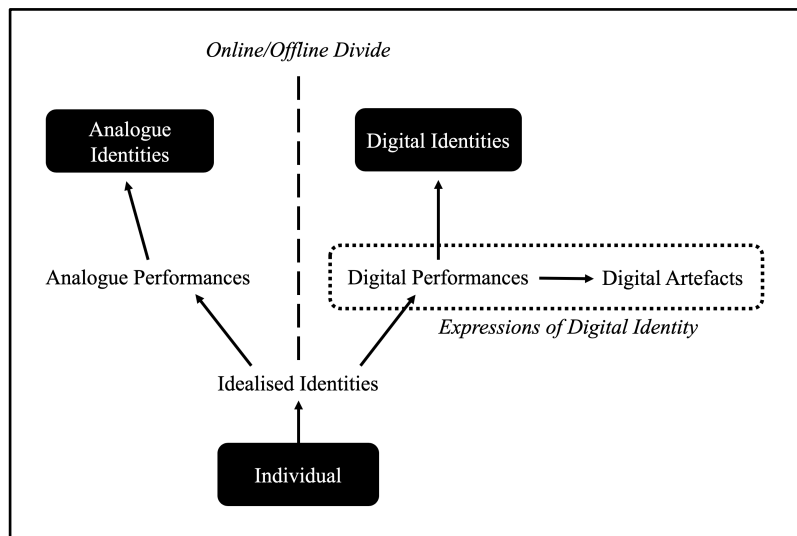


Figure 1. A schematic delineating an individual's various kinds of identities, and their relationships to different presentations of self, both online and offline

§2 – From Expressions to Identities

To maximise engagement, and therefore advertising revenues, social media sites strive to show individuals content that ‘people like them’ have interacted with (Wu 2017) – i.e. identify their digital identity markers (Cheney-Lippold 2011, p. 165). To do so, machine learning algorithms are used to infer anything from an individual’s “sensitive attributes (e.g., race, gender, sexual orientation)” to their interests and “opinions (e.g., political stances)” from the data they post and consume⁹ (Wachter and Mittelstadt 2019, p. 4). These psychometric categorisations allow advertisements (and content) to be targeted at specific ‘audiences’ (Bay 2018, p. 1723) – the same process by which Cambridge Analytica categorised possible voters into “universes” and social marketers sort potential customers into “buckets” (Bartlett 2018, p. 83). Does this categorisation, however, amount to individuals being assigned a digital identity based on their digital performances?

I think not. Indeed, I believe that conflating digital artefacts and digital performances has led to an equivocation between digital identities and *corporatised identities* – a term for the audiences, buckets and universes used to wrangle economically-exploitable categories out of the masses of data generated by social media usage. These corporate amalgamations of digital identities are not digital identities proper (Fig. 2) and, as I will demonstrate, prioritising their production at the expense of digital identities themselves results in commercial values being forcefully impressed upon individuals. The term ‘corporatised’ is thus fitting, as it means to make something “corporate by introducing or imposing the structures, practices, or values associated with a large business corporation; to commercialize” and, hence, “to deprive” that thing “of independence or individual character” (OED Online 2006). Bearing this definition in mind, I will argue below that the unique, fluctuating digital identities that individuals seek to create and maintain over social media are being corrupted by the production of corporatised identities that identify individuals for advertising and tracking purposes. Indeed, as it will transpire, it is damaging to confuse this forensic, corporate process of individuation and identity *assignment* with the fluid and social *construction* of digital identities, as production of the former is harmful to individuals’ digital well-being in its current form.

⁹ As these conclusions often cannot be understood by humans such algorithms are often described as being ‘opaque’ (Villaronga et al. 2018, p. 308).

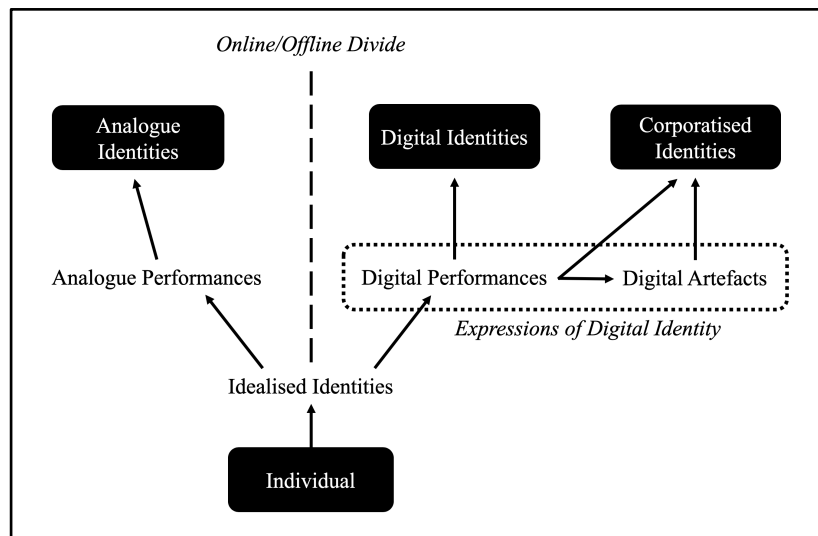


Figure 2. A revised schematic accounting for an individual's 'corporatised identities'

§2.1 – Corporatised Identities

Corporatised identities are no more than commercially useful extrapolations inferred from the deluge of expressions of digital identity recorded by social media companies. Whilst such “social sorting” (Lyon 2014, p. 10) no doubt captures, perhaps quite accurately, elements of a performed identity, it therefore cannot ever hope to emulate or equal an individual’s ever-changing digital identities. There are at least four reasons for this.

Firstly, algorithmically generating corporatised identities relies on analysing many *different* individuals’ expressions of digital identity together (de Vries 2010, p. 77; Manders-Huits 2010, p. 45), comparing them to draw useful boundaries around similar groups (Wachter and Mittelstadt 2019, p. 13). Doing so allows social media companies to “draw non-intuitive and unverifiable inferences and predictions about the behaviors, preferences, and private lives of individuals” (Wachter and Mittelstadt 2019, p. 4). Numerous individuals’ expressions of digital identity are thus constantly being processed into discrete, machine-readable categories. At most, this means that what Goffman would term ‘group’ or ‘role’ identities might be understood by social media companies, inferred from common features of a population, but not *an* individual’s digital identities. Indeed, this is inevitable, as individuals are only ever members of a category for their algorithms – points in a dataset at a relative level of abstraction rather than persons (Cheney-Lippold 2011, p. 176).

Secondly, because new data are always being gathered, the algorithms continually redefine and refine the boundaries of corporatised identities (O’Neill 2016), tweaking them to improve their

effectiveness¹⁰ (de Vries 2010, p. 81). Who I ‘am’ (which categories describe me) is therefore in a constant state of flux, not only due to my actions, but also how the algorithm understands those actions to relate to others’. Whilst I may be statistically identified as ‘male’ today, tomorrow I might be recategorised as ‘female’ based on new performances (Cheney-Lippold 2011, p. 173). When this happens, however, the entire advertising infrastructure shifts to treat me completely differently, targeting me with different adverts and content, destabilising the contexts in which I act online. This means that algorithms now partially co-create and “supplement existing discursive understandings of a category” (Cheney-Lippold 2011, p. 173). Gender classifications are just one example, but they illustrate that fickle, group-level distinctions define corporatised identities, not anything approaching the granularity of an individual’s actual performed identity. In contrast, as we have already seen from Goffman, individuals generally try to consistently perform coherent identities across situations, doing intersubjective work to ensure their identities are being correctly understood. This clearly conflicts with the continuous modulation of categories and the pursuit of economically effective, but not necessarily accurate, categorisations.

The third difference stems from the commercial motivations behind this process. Individuals are categorised to show them content that similar users have found engaging (de Vries 2010, p. 77), but also to make decisions “about how best to predict, persuade, and ultimately control the behaviour of the user” (Burr and Cristianini 2019, p. 463). The end goal of this process, after all, is converting attention into profit through various monetisation techniques¹¹ (Wu 2017; Zuboff 2015). So, despite rhetorical protestations to the contrary, social media companies are not audiences seeking to truly interact with the performing individual, despite the “power of performativity” turning “crucially on its *interactive edge*” and the “relationship” between performer and audience (Schieffelin 1998, p. 200). Rather, leveraging their ubiquity as platforms, these companies are inserting themselves *between* genuinely interacting individuals as parasitic accessories to their performances in order to monetise their interactions.

Accordingly, social media companies are agnostic about digital identities tout court, because only *comparable*, machine-readable identity characteristics need to be inferred from individuals’ expressions of digital identity. Indeed, only those inferences that might be economically-exploitable are relevant for rendering their vast stores of raw data useful (Cheney-Lippold 2011, p. 170). Elements of identities that do not serve this purpose or cannot be inferred from expressions

¹⁰ This distinguishes corporatised identities from idealised identities that shape individuals’ performances. The former are transient, with their boundaries constantly being re-drawn, whilst the latter are socially-determined and so somewhat fixed by the possibilities of a particular society.

¹¹ International markets, not concern for “data justice” or ethics, therefore drives this process (Taylor 2017, p. 3).

of digital identity are therefore discarded or computationally unobservable, leaving companies with limited, but commercially useful, shadows of a potential consumer's identities.

Fourth, and finally, this means that digital performances and artefacts are wilfully conflated and mined in search of any exploitable insights. My past is, after all, as exploitable as who I now am and wish to be. Beyond the conflation of self and other we have discussed, past and present, performance and artefact, and all an individual's multiple identities, too, are bundled together and mined for commercial predictions. As such, whilst corporatised identities reflect salient, generalisable features of the *expressions of digital identity* that linger in the network *and can be computationally modelled* (itself a large caveat) (Hildebrandt 2019, p. 92), due to their indiscriminate and inferential construction they will never be equivalent in terms of detail, scope or contemporaneity to any true digital identity. No genuine consideration of self – no relational, negotiated social interaction – occurs, in favour of inferences derived from collated and often outdated expressions of self.

To appreciate this, recall the imagery of artefacts and exhibitions. Evidently, an archaeologist cannot 'know' the identity of an individual they find buried, beyond perhaps in the simple sense of identification, because they never had a social relationship with them. At best they might be able to infer some educated guesses from artefacts left behind (e.g. diaries). But, similarly, the algorithms utilised by social media companies can do little better, if at all, despite such guesses being how these companies claim to 'know' their users. In reality, both archaeologist and algorithm can at best only speculate, as they are removed from their objects of interest – there are no negotiated performances occurring between them.

Furthermore, whilst algorithms do not only consider digital artefacts – they can also access ongoing performances – the former are almost certainly given more weight in their categorisations than a human might. Whilst human memories are relatively undependable, and analogue performances can fade over time, data storage is cheap and reliable, so digital performances are rarely, if ever, 'forgotten'¹² (Manders-Huits 2010, p. 52). But this means that artefacts can easily outweigh those most fleeting performances an individual is currently giving. A recently reformed smoker might thus continue to be pigeon-holed by adverts that seek to sell a product aimed at who they were but no longer wish to be.

§2.2 – Initial Issues with Corporatised Identities

¹² Unless legally obligated to 'forget' by, e.g., Article 17 of the European General Data Protection Regulation. For more, see: (Politou et al. 2018; Villaronga et al. 2018).

Corporatised identities, then, are clearly not digital identities. Nonetheless, like many similar “flawed models” (Bridle 2018, p. 34), corporatised identities can exert a potentially problematic influence on the reality they are only meant to be modelling (Hildebrandt 2019, pp. 91–92). In practice this means, once assigned a corporatised identity, individuals are repeatedly shown content matched to their ‘type’. A cycle of reinforcement therefore proceeds (Elmer 2003), as individuals are shown digital artefacts nudging them to act in ways which re-confirm the original categorisation, in turn deepening the algorithm’s confidence in a correct classification¹³ (Lessig 1999, p. 154). Unsurprisingly, this then feeds back into the identity formulation process and individuals’ future performances (Burr et al. 2018; de Vries 2010). As Lanier (1995, p. 67) colourfully puts it, the result is that individuals are reduced to a “cartoon model” of their purported interests, and one that is self-reinforcing¹⁴. This cannot authentically be described as mapping an identity, though, but at least partially defining it – shaping fluid human identities to match rigid categories that algorithms can compute, rather than reflecting the complexities of identities that actually exist ‘out there’ in the social world.

However, at this point a sceptical reader might respond that, in everyday life, we are also always sorting people into various groups, getting those classifications wrong and recategorizing people – and that a central part of identity construction involves negotiating with our peers where we stand in their respective social categories. Are social media companies’ categorisations not just further examples this?

This, I believe, is a misunderstanding – social media companies’ classifications are markedly different. Ordinarily, the categorisation (or stereotyping, pigeon-holing, and so on) that individuals do to one another is *multifarious*; varied beliefs and biases mean everyone treats each other in slightly different ways. This results in exposure to a valuable plurality of viewpoints, against which individuals can examine their own lives and identity trajectories (See, e.g. Muldoon 2015, pp. 182–184) – empowering them to figure out who they want to be, which identities to perform and how they want to be seen by others. It supports, in other words, Mill’s “experiments of living” (Mill 1984, p. 115), contributing directly towards their well-being. And, far from being undesirable, this helps individuals to grow and explore new identities, contributing to a flourishing pool of potential lifestyles in liberal society¹⁵.

¹³ Success is achieved when a “categorization fits the behavior of a user”, without regard for whether a user actually “embodies that category” (Cheney-Lippold 2011, p. 179).

¹⁴ My thanks to Christopher Burr for drawing my attention to this formulation of the issue.

¹⁵ I gesture here to the rich liberal literature on *value pluralism*. For more, see e.g. (Crowder 1998; Galston 1999).

Online, however, things are different. As we have seen, the construction methods of corporatised identities ensure that the *only* categorisations made by social media companies are those which ultimately serve a narrowly single-minded and economic impetus. This is radically unlike the diverse stereotyping and categorising that friends, colleagues and strangers do. Indeed, appreciating this “strips away the illusion that the networked form has some kind of indigenous moral content” to reveal the socially-parasitic, commercial impetus behind social media’s supposedly-social design, and shows that individuals trying to authentically socially interact are actually being used as the “brazen means to others’ commercial ends” (Zuboff 2019a, p. 19). I will return to the normative implications of this shortly.

§2.3 – Digital Identities

Nonetheless, having distinguished corporatised identities, we can now better ascertain the nature of digital identities themselves. An immediate and major difference is that digital identities, much like their analogue equivalents, are intensely personal and so largely defined by their uniqueness. Whilst corporatised identities attend to similarity and sameness at the level of categories, digital identities are characterised by their distinctive *individuality* at the level of persons; no two digital identities will likely ever be the same in either aspiration or performance.

Ontologically, this follows from the “informational nature” of identities, according to which digital identities are comprised of the rich perceptions (or narratives) generated by an individuals’ actions (Floridi 2011b, p. 556), far beyond those data required for identification and monetary exploitation. Indeed, it is through individuals experiencing and remembering digital social interactions that they perceive others’ complex identities. Since two individuals’ minds will never possess all the same information, they will never perform identical identities, as the manifold diversity of previous experiences (stored as memories) will frame and condition their future interactions. As such, two distinct individuals will always interact in unique ways and perform subtly different identities, whether online or offline.

Another vital difference is that, both online and offline, the motivation for identity-relevant action stems from the desire to perform a particular idealised identity successfully. Evidently digital identities, not corporatised identities, drive this process online, not least because individuals usually lack direct epistemic access to the corporatised identities that companies extrapolate. Although moving performances online therefore changes the ways in which successful negotiation might be achieved, and introduces new mechanisms and techniques for identity management, the fundamental aim – of successful performance – remains consistent. This aim, crucially, is not one

shared by corporatised identities and the social media companies that create them for an identification and governance agenda.

As we have seen, the creation of corporatised identities is principally commercially motivated. As such, whilst digital identities primarily obtain *between* persons, existing as shared constructions regardless of the medium by which they are performed, corporatised identities exist on mainframes, belonging to the companies that generate them. The former are irreducibly social; negotiated, ongoing, and reciprocal. By contrast, the latter are one-sided, privileged and readily exploited by one party. Even if companies partially construct the performance environment, conditioning individuals' performances, corporatised identities are not a shared endeavour but a powerful tool for behavioural manipulation¹⁶.

Similarly, the importance of digital *performances* to digital identities, rather than digital artefacts, cannot be stressed enough. Recalling how misguided it would be to suggest that an archaeologist 'knows' a corpse, the centrality of *intersubjective* online experience to digital identities should now be clear. A performance's ephemeral nature stems directly from this intersubjectivity, as without the possibility of reacting and adapting to an audience what is actually being created is a digital artefact, because performances "exist only in the present"¹⁷ (Schieffelin 1998, p. 198). Identities are fluid and, although somewhat fixed idealisations shape how individuals act, they are not fixed in the manner which corporatised identities require to enable artificial comprehension (Hildebrandt 2019, p. 91). This conflicts with the desire to categorise that provokes social media companies' constructions of corporatised identities in the first place. And, even though the boundaries of corporatised identities are constantly changing, this does not emulate the fluidity that runs through an individual's intentional, shifting performances of an identity. Algorithmic redefinition is a disjointed exercise in reactive speculation, not one of coherent but evolving presentation.

Credence is lent to this claim of fluidity, or aversion to identity-fixing, by appeal to psychological studies. For instance, being confronted with artefacts that individuals do not currently identify with, even if those prior performances were integral to their identity at the time, can generate significant discomfort (Tian 2017, p. 204). Such "mismatched expectations" surrounding who individuals believe themselves to be can lead, in particular, to embarrassment and anxiety (Tian 2017, p. 205). As I have shown, this is because identities are ongoing and

¹⁶ Facebook's emotional contagion experiments, for instance, explored their ability to influence individuals' mental states without their knowledge (Kramer et al. 2014). For further discussion of behavioural manipulation in the context of digital well-being, see (Klenk, this collection). In particular, Klenk's assertion that such manipulation can deny individuals both autonomy and control over valuable aspects of their lives chimes well with the argument I later advance in §3.

¹⁷ That said, prior performances "inform present ones" (Davis 2014, p. 514), so there is some *continuity* of identity.

multifaceted constructions, constantly changing and adapting. Fixing snapshots of identities in social media networks therefore creates the potential for conflicts between currently lived identities and the fossilised remains of identities that have been left behind.

This leads to a final key insight regarding the nature of digital identities. As a “continuous” process of reflection, identity work consists in “evaluating and *identifying* with one’s attributed identifications” (Manders-Huits 2010, p. 50, emphasis mine). Individuals, in other words, must feel that the identities they are performing are truly ‘theirs’. In Paul Ricoeur’s terminology, this feeling links to an individual’s *ipse* identities, those which they first-personally experience and recognise as their “unique selfhood” (de Vries 2010, p. 74). People clearly identify with their digital identities, given their continued attempts to present them, but would not necessarily feel the same about a corporatised one – indeed, like with any other group categorisation, people often find it alienating, arresting or uncomfortable to see how they have been (mis)characterised (de Vries 2010, p. 81; Newman and Newman 2001, p. 526). This arises from indignance; a feeling that ‘that’s not who I am’.

Bearing all of this in mind, I therefore tentatively submit that digital identities should once again be allowed to underpin individuals’ online performances on social media if we are to avoid these issues. Indeed, before corporatised identities were created, digital identities surely did fulfil this role: individuals intersubjectively performed their identities, utilising the internet to explore them in new ways. With the increasing production of corporatised identities, though, individuals’ digital identities – those they lived and felt to be their own – have been systematically displaced. Instead, as platforms have realised their potential monetary value, they have begun building and exploiting corporatised identities at the expense of digital identities. This development, however, has damaging corollary effects: namely, reducing individuals’ autonomy over their own identities and how they are presented to others. And, having already touched on these issues, we are now better equipped to explore them in detail.

§3 – Relational Autonomy and the Harms of Corporatised Identities

Clearly, as corporatised identities are algorithmically generated behind closed doors, individuals cannot straightforwardly exercise autonomy over these identities. However, at the same time, corporatised identities readily condition individuals' social contexts, enabling social media companies to exert influence over their digital well-being. This primarily occurs through the filtering of information, as these companies control the internet's advertising and social infrastructures – a fact that is potentially worrying because, as individuals, “we interact, flourish, or suffer depending on the flows of information in which we partake” (Floridi 2019, p. 379). But, again, because all this filtering is oriented towards maximising engagement, ad clicks and revenue, social media sites have been redesigned to be increasingly addictive (Andreassen 2015). After all, for these companies we only “appear as statistical objects of study, abstracted from our personal preferences and life plans, and from our individual capacities and freedom to choose” (Manders-Huits 2010, p. 45).

In other words, a substantial problem with producing corporatised identities is that individuals are not treated as autonomous agents – they are not afforded adequate moral concern for their capacity to choose. This has significant implications for eudaimonic approaches to well-being, because possessing sufficient autonomy is vital (Deci and Ryan 2008, p. 6). Indeed, autonomy is of central importance to all conceptions of the “fully functioning person” that defines eudaimonic well-being (Ryan and Deci 2001, p. 161)¹⁸. Harm to an individual's digital well-being is therefore an inevitable outcome of the production of corporatised identities if this process is damaging to autonomy. Thus, to support these conclusions, the remainder of this section fleshes out a particular conception of *relational* autonomy and explores how it is adversely affected by corporatised identities.

§3.1 – A General Account of Relational Autonomy

Respect for individual autonomy is woven into the very fabric of liberal democracies as a “cardinal moral value” defended morally and in law (MacKenzie 2008, p. 512). Traditional understandings of autonomy, however, have been criticised as excessively masculine, individualistic and atomistic, driving the development of relational approaches to autonomy that

¹⁸ See (Calvo, Peters, Vold, Ryan, this collection) for more on this relationship, as seen through the lens of self-determination theory. Their work make great strides towards the development of a more systematic approach to diagnosing autonomy-compromising digital design.

seek to recognise individuals as “emotional, embodied, desiring, creative, and feeling” (MacKenzie and Stoljar 2000, p. 21). At a fundamental level, all such theorists agree that individuals are irreducibly socially rooted and that their “identities are formed within the context of social relationships” (MacKenzie and Stoljar 2000, p. 4). The parallels with the intersubjective account of identity above should therefore be self-evident, and this also sits comfortably with eudaimonic well-being’s recognition of the need for an “appropriate and situated notion of autonomy” (Devine et al. 2008, p. 132). Nonetheless, navigating the numerous strands of relational autonomy would be beyond our requirements¹⁹, as MacKenzie (2014, 2019) has developed a multidimensional model that integrates the various approaches along three interrelated axes. These are:

- *Self-governance*, which “involves having the skills and capacities necessary to make choices and enact decisions that express, or cohere with, one’s reflectively constituted diachronic practical identity” (MacKenzie 2014, p. 17);
- *Self-determination*, which “involves having the freedom and opportunities to make and enact choices of practical import to one’s life, that is, choices about what to value, who to be, and what to do” (MacKenzie 2014, p. 17); and,
- *Self-authorization*, which “involves regarding oneself as having the normative authority to be self-determining and self-governing... [i.e.] authorized to exercise practical control over one’s life, to determine one’s own reasons for action, and to define one’s values and identity-shaping practical commitments” (MacKenzie 2014, p. 18).

Along each dimension, various circumstances can therefore either support or restrict an individual’s autonomy. This includes internal conditions, such as the individual’s own psychology, as well as external conditions, like “social norms, institutions, practices, and relationships” that can “effectively limit the range of significant options available” (MacKenzie and Stoljar 2000, p. 22). Brainwashing children, for instance, can not only limit their beliefs and desires to those their parents find acceptable, but can also impede the development of their critical faculties, leading to dependencies. Relational approaches to autonomy recognise that this would rob them of authority over their own lives, replacing their freedom to live as they wish with a narrow conception of the good life. And, in line with this, the eudaimonic understanding of well-being, too, explicitly recognises the detrimental effects for well-being entailed by such losses to autonomy (Devine et al. 2008, p. 113).

¹⁹ See, e.g.: (Barclay 2000; Baumann 2008; Christman 2004; MacKenzie 2014; Westlund 2009).

As I will now demonstrate, all three dimensions of MacKenzie's model can ground criticisms of corporatised identities. Consequently, I remain broadly neutral towards particular interpretations of relational autonomy because, if deploying corporatised identities potentially impinges upon elements of all three axes, it would suggest they are generally incompatible with relational autonomy and so undermine digital well-being, too. If liberal democracies seek to promote respect for relational autonomy, then I propose the practice of creating corporatised identities should thus be altered – or even rejected altogether. Regardless, MacKenzie's model illuminates the theoretical utility of distinguishing corporatised identities from the digital.

§3.2.1 – *Self-Governance*

The dimension of self-governance fundamentally considers (i) the individual's *competency* or internal capabilities to make and act upon free decisions in line with their wishes, and (ii) whether their choices and preferences are their own – i.e. the *authenticity* of their intentions (MacKenzie 2019, p. 149). It particularly focuses on the individual's moral psychology; whether they possess the capacity to be autonomous, or whether disability or dependency has made them unable to form or execute their intentions (MacKenzie 2019, p. 147). As a relational theory, however, attention is also paid to the social and institutional environment required for the development and sustenance of effective self-governance, and issues such as stereotyping and adaptive preferences pertinently illustrate how inauthenticity can be forced upon individuals (MacKenzie 2014, p. 32), alienating them from themselves and generating internal conflict over their identity and value commitments (see e.g. Khader 2011).

Where corporatised identities are utilised, both competency and authenticity conditions are potentially undermined, both online and offline. Firstly, when individuals rely on algorithms to recommend purchasing choices, they can often “surrender to technology” and settle for inferior products (Banker and Khetani 2019, p. 2)²⁰. Importantly, corporatised identities underpin these recommender systems, as without companies ‘understanding’ individuals they cannot guide their actions. This, however, can undermine an individual's competency and breed dependence, with algorithms simplifying complex situations and nudging users into commercially-motivated actions (Hildebrandt 2019, p. 105) despite these products affecting how individuals can perform identities, and despite purchasing them in itself being a kind of identity performance. Consider, for instance,

²⁰ Conversely, individuals' best interests can sometimes be *best* served by autonomously surrendering to technology, as is the case with fitness trackers. However, concerns surrounding dependence clearly remain in these cases.

Facebook and Google recommending ‘promoted’ restaurants paying for more exposure. But, in relying on recommendations, individuals plainly cede competency and opportunity for self-governance to a company using the influence granted by corporatised identities for profit-related purposes.

Secondly, algorithmic reliance can also remove opportunities for the development and performance of *authentic* digital identities. Algorithms can nudge us, for instance, to buy more than we need, not only reducing self-control but also promoting consumerism – identity traits that individuals may not reflexively endorse. Furthermore, as these systems expand beyond simple purchases it will only become more difficult to avoid their recommendations. The game ‘Pokémon Go’, for instance, was quietly monetised by Google/Alphabet through the sale of virtual land in real-world locations. So, companies like Starbucks paid for in-game monsters to reside near their cafés, herding players to stores and boosting sales, whilst players were not informed this is how monsters were distributed (Zuboff 2019b, ch. 10). Following the blockbuster success of this gambit²¹, a business model built on covertly manipulating individuals has thus been realised, with social media companies keen to release the latent value of vast stores of digital identity data (Zuboff 2019a, p. 19). Nonetheless, if the values that individuals are pushed to endorse through these systems are not their own, then authenticity is lost, with commercial values supplanted for individuals’ own. In this case, for example, a desire for videogaming (a kind of digital identity performance) in nature is being contorted into an opportunity for coffee sales. Consequently, the rich variety of preferences that underpin a digital identity are being collapsed or flattened in favour of those preferences which can be economically exploited, stymieing the authentic development of individuals’ varied preferences and harming their well-being. Indeed, these systems all display significant biases for increasing “sales, ad exposure, user engagement, and [...] other strategic goals” that are likely to conflict with an individual’s own values (Banker and Khetani 2019, p. 4).

Corporatised identities are thus leveraged to target individuals with sophisticated behavioural manipulation systems designed to service these companies bottom lines at the expense of an individual’s capabilities for self-governance, undermining both the authenticity of their preferences and potentially their competency to make agential decisions. Whilst these systems are only just emerging, the potentially damaging effects for digital identities are serious: with diminished self-governance, individuals will be less able decide what identities to perform, depending on algorithms whilst being herded into performing identities that suit a commercialised agenda. Indeed, because these autonomy-damaging systems all rely on corporatised identities to operate,

²¹ Pokémon Go broke five world records, including fastest mobile game to gross \$100m (Swatman 2016).

this gives prima facie reasons, on grounds of preserving self-governance, to want digital identities to underpin individuals' actions on social media instead. Individuals, after all, retain far more control over digital identities than the corporatised identities companies generate from harvesting their online interactions.

§3.2.2 – *Self-Determination*

The dimension of self-determination is preoccupied with external, structural threats to autonomy and its development in individuals. MacKenzie defines these in terms of (i) *freedom* conditions – the personal and political liberties protecting individuals from coercion, domination and exploitation (MacKenzie 2014, p. 25) – and (ii) *opportunity* conditions, which canvas the significant options individuals can choose from in society (MacKenzie 2014, p. 26). The two are interlinked, as in situations where liberties have been curtailed, individuals often also possess inadequate or insufficient meaningful life-options to choose from, and so their autonomous status is undermined (MacKenzie 2019, p. 147). In particular, MacKenzie is clear that having significant options cannot be reduced to an unlimited “array of consumer choices” (MacKenzie 2019, p. 148); there must be a legitimate variety of life-choices available for individuals to pursue, free from dominating forms of power and interference, for self-determination to be achieved.

Social media companies cannot currently forcibly curtail their users' freedoms. However, they are nevertheless narrowing the range of available significant opportunities. Filter bubbles and echo chambers, for example, are polarising individuals' identities, making their beliefs more extreme (Burr et al. 2018; Pariser 2012). Specifically, the most engaging emotions are anger, jealousy and outrage, so content generating these reactions is shown to individuals more often (Fan et al. 2014). Whilst (negatively speaking) individuals can post what they like, as only illegal and explicit materials are actively censored, the addictive mechanisms that elicit individual contributions are thus tuned to amplify and encourage content that generates these extreme emotions, polarising discourse. Consciously or not, if they wish to maximise engagement with friends, individuals are therefore conditioned to post such content by an external, coercive algorithmic force that only understands the world through corporatised identities. To expound on this point, consider the rise of ‘influencers’ on social media.

Influencers are the celebrities of social media culture (Cotter 2019, p. 896). Essentially, they sell a lifestyle through the products that go with it, and therefore rent themselves as vehicles for advertisements on social media with the hopes of nudging their followers into purchasing sponsored products (Brown and Hayes 2008). Accordingly, many potentially authentic social relationships between followers and influencers are reduced to little more than a friendly-faced

exchange of possible consumer choices. This is only compounded by the planned nature of many influencer endorsements – posts must often be signed off by advertising executives months in advance, meaning that such posts are artefacts, not intersubjective performances, long before they have even been posted. But, accordingly, influencers’ identities are definingly shaped by the algorithms aimed at promoting those who best maximise engagement and revenues²² (Cotter 2019, p. 901). Success on Instagram, after all, requires moulding your identity into one that the algorithm will favour.

This has clear determining effects on influencers’ identities but also affects the performances and identities of their followers – normal people interacting in the disguised marketplaces of social media. This is because companies that competitively sponsor influencers usually gain the most publicity. Importantly for self-determination, however, this comes at the expense of significant alternative options – even those that might be free – regardless of whether the alternatives might actually be in individuals’ better interests, or better support their social interactions. After all, simply having more (purchasing) options is not enough; alternatives must also be *significant* as, “if one has an inadequate range of significant options to choose from, one’s autonomy is diminished” (Brison 2000, p. 285).

Even putting influencers aside, however, social media companies generally control which of individuals’ digital identity performances are revealed to others; algorithms, after all, are constantly using corporatised identities to decide which posts to highlight to friends or not. Whilst self-determination over digital identities is thus not entirely undermined by these companies, the structural barriers to resisting their inferences are clear to see. As Shoemaker puts it (via Manders-Huits 2010), corporatised identities undermine individuals abilities to “present [their] self-identity to others in the manner [they] see fit”, meaning their “autonomy is undermined” as they are “unable to be the manager of [their] own reputation” (Shoemaker 2010, p. 13). Indeed, an algorithm’s mediation of digital identity performances removes “a key element of self-determination” even if they would have shared the information themselves (Shoemaker 2010, p. 13). In other words, individuals are reduced to their interests and behaviours as can be understood by these algorithms and which serve their ends (Williams 2005, p. 108). But, in these cases, an external imposition is clearly determining who individuals are seen to be – defining their identities for others by filtering posts, and hence limiting opportunities for self-determination, through the algorithm’s role in structuring how and what information is revealed on social media sites.

²² Influencers might therefore fail to fulfil the authenticity conditions of self-governance.

§3.2.3 – *Self-Authorisation*

Self-authorisation is concerned with an individual seeing themselves as deserving of self-respect, self-trust and self-worth/esteem (MacKenzie 2019, p. 149). This reflective element is irreducibly social, as it is through our relationships that these evaluative attitudes are built up. Individuals are only ever empowered to “speak and answer” for themselves if they are treated as people in their own right and so regard themselves as autonomous (MacKenzie 2019, p. 149). Accordingly, socially stigmatising practices that undermine self-authorisation can in turn undermine an individuals’ self-governance and self-determination (MacKenzie 2019, p. 150).

Immediately, and even discounting Facebook’s emotional manipulation experiments, the notion of self-worth can be easily linked to the expansive literature on social media’s negative effects on self-image. Humans naturally compare themselves with others but social media disproportionately exposes individuals to the highlights of others’ lives, damaging their self-esteem (Vogel et al. 2014, p. 206). Heavy users of social media are therefore generally more depressed (Feinstein et al. 2013), and report lower levels of well-being (Kalpidou et al. 2010; Shakya and Christakis 2017)²³. Crucially, social media usage also increases the perceived gap between who individuals want to be and who they think they are (Haferkamp and Krämer 2010). Or, in other words, the gap between an individual’s idealised identities and actually performed digital identities, likely due in part to the re-presentation of embarrassing or outdated expressions of digital identity. In other words, it can damage the “properties or beliefs about ourselves we value and respond to emotionally in relation to our self-esteem” (Manders-Huits 2010, p. 46).

Taken to its extreme, this alienation of the individual from themselves, and the accompanying erosion of self-worth and self-respect, is terminal for autonomy along this dimension. Indeed, due to the increased prevalence of negative self-evaluative feelings, social media users often report feelings of listlessness and isolation, and that they could have spent their time more fruitfully (Primack et al. 2017). Additionally, excessive social media usage often precludes individuals from pursuing activities that could actually increase genuine face-to-face interactions and a sense of fulfilment (Newport 2019, pp. 168–169). For some individuals, then, social media takes time away from more fulfilling pastimes that increase feelings of self-respect and self-esteem, directly undermining their sense of self-authorisation.

²³ It is worth noting some of these conclusions have been recently challenged by work testing the link between digital technology and adolescent well-being (Orben et al. 2019; Orben and Przybylski 2019), although this bears no impact on the identity effects that Haferkamp and Krämer researched.

What is more, algorithms are inherently socially stigmatising, given that they operate through classification at a generalisable level: they are, quite literally, stereotyping processes²⁴. It is this that allows companies to use individuals as a means to their economic ends. However, stigmatisation that treats people as less than human can damage self-trust and their understanding of their own digital identities. Not only can individuals surrender to algorithmic choices in the face of complexity, but self-authorisation clearly plays a role here, as regarding oneself as unable to make a competent autonomous choice betrays a lack of self-respect and self-trust. Indeed, one reason for handing more responsibility for decisions over to algorithms is that they are often regarded by the public to be more competent choosers (Zittrain 2019); algorithms are trusted to have ‘gotten it right’, so individuals often feel they should obediently follow recommendations suitable for ‘their’ category (de Vries 2010, p. 82). Algorithms, though, only identify statistical correlations between datapoints, not causal links (Zittrain 2019) – there is no reasoning taking place. Taken in by a veneer of algorithmic competency, reliance on machine intelligences can thereby weaken an individual’s understanding of their own identities, challenging who they believe themselves to be and altering the ways in which they perform their identities at the expense of self-authorisation.

Finally, it is vital to remember that algorithmic curators have no respect or understanding for individuals as people – they are, after all, only members of a set. There is no social relationship between the algorithmic ‘audience’ and performer; the relationship is parasitic (Hildebrandt 2019, p. 107). Their aim, additionally, is not to support identity experimentation but to encourage reliance and addiction in order to generate corporatised identities that can be monetised. Autonomous control over digital identities is therefore undermined through a process explicitly designed to generate reliance and compliance with a corporatised classification. These algorithms are, quite literally, “traps” that measure success in terms of their “captivation” and retention of individuals (Seaver 2018, p. 9). I can think of no better metaphor for an autonomy-decreasing mechanism than this. Consequently, and in concert with my previous points, this underlines why a return to the encouragement of digital identities, which individuals can not only exert control over, but also socially perform and develop, is so urgently required.

²⁴ Categorisation appears in traditional marketing but, when achieved algorithmically, deep granularity results in categorically different predictive abilities (See e.g. Canny et al. 2011).

Concluding Remarks

This could all understandably be read in oppressively bleak terms. But, in closing, I want to emphasise that it need not be like this. The relatively unfettered performance of digital identities used to motivate most social media interactions and shows that corporatised identities are not essential for a flourishing social media environment; they are a parasitic addition whose uptake has been driven by a recent drive for monetisation. Nonetheless, social media now permeates every second and sphere of daily life, ensuring that individuals' significant options, values and preferences are being constantly conditioned and constrained by a consumerist agenda. This, however, is a choice, and one that could be overturned in favour of a less harmful mechanism for monetisation. Regardless, on the account I have given here, the production and utilisation of corporatised identities cannot endure without continuing to harm individuals' performances of their authentic digital identities, limiting their relational autonomy in a way that I do not believe is compatible with either liberal democratic respect for autonomy or regard for their (eudaimonic) digital well-being.

So, in summary, having distinguished the constitutive elements of expressions of digital identity towards the beginning of this chapter, I then explained how corporatised identities are being conflated with digital identities proper. This, I believe, constitutes a useful theoretical contribution to modern Goffmanian identity theory and helps expose how the displacement of digital identities on social media has come to potentially undermine individual autonomy when understood in relational terms. Indeed, I have argued that designing social media sites around the algorithmic production of corporatised identities, at the expense of individuals' digital identities, has likely undermined their self-governance, self-determination, and self-authorisation. It was for these reasons that I concluded that digital identities must once again be allowed to motivate individuals' interactions on social media unencumbered. Only upon doing so can those individuals' digital well-being, and control over their own digital identities, be rendered compatible with social media.

Bibliography

- Andreassen, C. S. (2015). Online Social Network Site Addiction: A Comprehensive Review. *Current Addiction Reports*, 2(2), 175–184. <https://doi.org/10.1007/s40429-015-0056-9>
- Bakir, V., & McStay, A. (2018). Fake News and The Economy of Emotions. *Digital Journalism*, 6(2), 154–175. <https://doi.org/10.1080/21670811.2017.1345645>
- Banker, S., & Khetani, S. (2019). Algorithm Overdependence: How the Use of Algorithmic Recommendation Systems Can Increase Risks to Consumer Well-Being. *Journal of Public Policy & Marketing*, 1–16. <https://doi.org/10.1177/0743915619858057>
- Barclay, L. (2000). Autonomy and the Social Self. In C. MacKenzie & N. Stoljar (Eds.), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (pp. 52–71). Oxford University Press.
- Bartlett, J. (2018). *The People Vs Tech: How the internet is killing democracy*. Ebury Press.
- Baumann, H. (2008). Reconsidering Relational Autonomy. Personal Autonomy for Socially Embedded and Temporally Extended Selves. *Analyse & Kritik*, 30(2), 445–468. <https://doi.org/10.1515/aug-2008-0206>
- Bay, M. (2018). The Ethics of Psychometrics in Social Media: A Rawlsian Approach. In *Proceedings of the 51st Annual Hawaii International Conference on System Sciences (HICSS'18)* (pp. 1722–1730). <https://doi.org/10.24251/hicss.2018.217>
- boyd, danah. (2007). Why Youth (Heart) Social Network Sites: The Role of Networked Publics in Teenage Social Life. In D. Buckingham (Ed.), *Youth, Identity, and Digital Media* (pp. 119–142). MIT Press. <https://doi.org/10.31219/osf.io/22hq2>
- boyd, danah, & Heer, J. (2006). Profiles as Conversation: Networked Identity Performance on Friendster. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)* (Vol. 3, pp. 1–10). Presented at the Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06). <https://doi.org/10.1109/HICSS.2006.394>
- Bridle, J. (2018). *New Dark Age: Technology and the end of the future*. Verso Books.
- Brison, S. J. (2000). Relational Autonomy and Freedom of Expression. In C. MacKenzie & N. Stoljar (Eds.), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (pp. 280–299). Oxford University Press.
- Brown, D., & Hayes, N. (2008). *Influencer Marketing*. Routledge. Accessed 18 August 2019
- Bullingham, L., & Vasconcelos, A. C. (2013). 'The presentation of self in the online world': Goffman and the study of online identities. *Journal of Information Science*, 39(1), 101–112. <https://doi.org/10.1177/0165551512470051>

- Burr, C., & Cristianini, N. (2019). Can Machines Read our Minds? *Minds and Machines*, 29(3), 461–494. <https://doi.org/10.1007/s11023-019-09497-4>
- Burr, C., Cristianini, N., & Ladyman, J. (2018). An Analysis of the Interaction Between Intelligent Software Agents and Human Users. *Minds and Machines*, 28(4), 735–774. <https://doi.org/10.1007/s11023-018-9479-0>
- Canny, J., Zhong, S., Gaffney, S., Brower, C., Berkhin, P., & John, G. H. (2011, April 5). Granular data for behavioral targeting using predictive models. <https://patents.google.com/patent/US7921069B2/en>. Accessed 19 August 2019
- Chaudhry, I. (2015). #Hashtagging hate: Using Twitter to track racism online. *First Monday*, 20(2). <https://doi.org/10.5210/fm.v20i2.5450>
- Cheney-Lippold, J. (2011). A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control. *Theory, Culture & Society*, 28(6), 164–181. <https://doi.org/10.1177/0263276411424420>
- Christman, J. (2004). Relational Autonomy, Liberal Individualism, and the Social Constitution of Selves. *Philosophical Studies*, 117(1/2), 143–164. <https://doi.org/10.1023/B:PHIL.0000014532.56866.5c>
- Cotter, K. (2019). Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New Media & Society*, 21(4), 895–913. <https://doi.org/10.1177/1461444818815684>
- Crowder, G. (1998). From value pluralism to liberalism. *Critical Review of International Social and Political Philosophy*, 1(3), 2–17. <https://doi.org/10.1080/13698239808403245>
- Davis, J. L. (2014). Triangulating the Self: Identity Processes in a Connected Era: Triangulating the Self. *Symbolic Interaction*, 37(4), 500–523. <https://doi.org/10.1002/symb.123>
- Davis, J. L. (2016). Identity Theory in a Digital Age. In J. E. Stets & R. T. Serpe (Eds.), *New Directions in Identity Theory and Research* (pp. 137–164). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190457532.003.0006>
- de Vries, K. (2010). Identity, profiling algorithms and a world of ambient intelligence. *Ethics and Information Technology*, 12(1), 71–85. <https://doi.org/10.1007/s10676-009-9215-9>
- Deci, E. L., & Ryan, R. M. (2008). Hedonia, eudaimonia, and well-being: an introduction. *Journal of Happiness Studies*, 9(1), 1–11. <https://doi.org/10.1007/s10902-006-9018-1>
- Devine, J., Camfield, L., & Gough, I. (2008). Autonomy or Dependence – or Both?: Perspectives from Bangladesh. *Journal of Happiness Studies*, 9(1), 105–138. <https://doi.org/10.1007/s10902-006-9022-5>

- Ellison, N. B., & boyd, danah. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210–230. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>
- Elmer, G. (2003). *Profiling Machines: Mapping the personal information economy*. MIT Press.
- Fan, R., Zhao, J., Chen, Y., & Xu, K. (2014). Anger Is More Influential than Joy: Sentiment Correlation in Weibo. *PLOS ONE*, 9(10), 1–8. <https://doi.org/10.1371/journal.pone.0110184>
- Feinstein, B. A., Hershenberg, R., Bhatia, V., Latack, J. A., Meuwly, N., & Davila, J. (2013). Negative social comparison on Facebook and depressive symptoms: Rumination as a mechanism. *Psychology of Popular Media Culture*, 2(3), 161–170. <https://doi.org/10.1037/a0033111>
- Floridi, L. (2011a). The Construction of Personal Identities Online. *Minds and Machines*, 21(4), 477–479. <https://doi.org/10.1007/s11023-011-9254-y>
- Floridi, L. (2011b). The Informational Nature of Personal Identity. *Minds and Machines*, 21(4), 549–566. <https://doi.org/10.1007/s11023-011-9259-6>
- Floridi, L. (2019). Marketing as Control of Human Interfaces and Its Political Exploitation. *Philosophy & Technology*, 32(3), 379–388. <https://doi.org/10.1007/s13347-019-00374-7>
- Galston, W. A. (1999). Value Pluralism and Liberal Political Theory. *The American Political Science Review*, 93(4), 769–778. <https://doi.org/10.2307/2586111>
- Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Anchor Books.
- Haferkamp, N., & Krämer, N. C. (2010). Social Comparison 2.0: Examining the Effects of Online Profiles on Social-Networking Sites. *Cyberpsychology, Behavior, and Social Networking*, 14(5), 309–314. <https://doi.org/10.1089/cyber.2010.0120>
- Hildebrandt, M. (2019). Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning. *Theoretical Inquiries in Law*, 20(1), 83–121.
- Hod, O. (2018, June 11). All of Your Facebook Memories Are Now in One Place. *Facebook Newsroom*. <https://newsroom.fb.com/news/2018/06/all-of-your-facebook-memories-are-now-in-one-place/>. Accessed 17 June 2019
- Hogan, B. (2010). The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, 30(6), 377–386. <https://doi.org/10.1177/0270467610385893>
- Hongladarom, S. (2011). Personal Identity and the Self in the Online and Offline World. *Minds and Machines*, 21(4), 533–548. <https://doi.org/10.1007/s11023-011-9255-x>

- Kalpidou, M., Costin, D., & Morris, J. (2010). The Relationship Between Facebook and the Well-Being of Undergraduate College Students. *Cyberpsychology, Behavior, and Social Networking*, 14(4), 183–189. <https://doi.org/10.1089/cyber.2010.0061>
- Khader, S. J. (2011). Adaptive Preferences and Choice: Are Adaptive Preferences Autonomy Deficits? In S. J. Khader (Ed.), *Adaptive Preferences and Women's Empowerment* (pp. 74–106). Oxford University Press. <https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199777884.001.0001/acprof-9780199777884-chapter-3>. Accessed 30 April 2019
- Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), 8788–8790.
- Lanier, J. (1995). Agents of Alienation. *Interactions*, 2(3), 66–72. <https://doi.org/10.1145/208666.208684>
- Lessig, L. (1999). *Code and Other Laws of Cyberspace*. Basic Books.
- Lyon, D. (2014). Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, 1(2), 1–13. <https://doi.org/10.1177/2053951714541861>
- MacKenzie, C. (2008). Relational Autonomy, Normative Authority and Perfectionism. *Journal of Social Philosophy*, 39(4), 512–533. <https://doi.org/10.1111/j.1467-9833.2008.00440.x>
- MacKenzie, C. (2014). Three Dimensions of Autonomy: A relational analysis. In A. Veltman & M. Piper (Eds.), *Autonomy, Oppression, and Gender* (pp. 15–41). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199969104.003.0002>
- MacKenzie, C. (2019). Feminist innovation in philosophy: Relational autonomy and social justice. *Women's Studies International Forum*, 72, 144–151. <https://doi.org/10.1016/j.wsif.2018.05.003>
- MacKenzie, C., & Stoljar, N. (2000). Introduction: Autonomy Reconfigured. In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (pp. 3–31). Oxford University Press.
- Manders-Huits, N. (2010). Practical Versus Moral Identities in Identity Management. *Ethics and Information Technology*, 12(1), 43–55. <https://doi.org/10.1007/s10676-010-9216-8>
- Marwick, A. E., & boyd, danah. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114–133. <https://doi.org/10.1177/1461444810365313>
- Mill, J. S. (1984). *Utilitarianism, On Liberty and Considerations on Representative Government*. (H. B. Acton, Ed.). Dent.

- Muldoon, R. (2015). Expanding the Justificatory Framework of Mill's Experiments in Living. *Utilitas*, 27(2), 179–194. <https://doi.org/10.1017/S095382081400034X>
- Newman, B. M., & Newman, P. R. (2001). Group Identity and Alienation: Giving the We Its Due. *Journal of Youth and Adolescence*, 30(5), 515–538. <https://doi.org/10.1023/A:1010480003929>
- Newport, C. (2019). *Digital Minimalism: choosing a focused life in a noisy world*. Penguin Business.
- OED Online. (2006). Corporatize. *OED Online*. Oxford University Press. <https://www.oed.com/view/Entry/267634>. Accessed 21 July 2019
- O'Neill, C. (2016). *Weapons of Math Destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.
- Orben, A., Dienlin, T., & Przybylski, A. K. (2019). Social media's enduring effect on adolescent life satisfaction. *Proceedings of the National Academy of Sciences*, 116(21), 10226–10228. <https://doi.org/10.1073/pnas.1902058116>
- Orben, A., & Przybylski, A. K. (2019). Screens, Teens, and Psychological Well-Being: Evidence From Three Time-Use-Diary Studies. *Psychological Science*, 30(8), 1254–1254. <https://doi.org/10.1177/0956797619862548>
- Pariser, E. (2012). *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin Books.
- Phillips, D. J. (2009). Ubiquitous computing, spatiality, and the construction of identity: Directions for Policy Response. In I. Kerr, V. M. Steeves, & C. Lucock (Eds.), *Lessons from the Identity Trail: anonymity, privacy, and identity in a networked society* (pp. 303–318). Oxford University Press.
- Politou, E., Michota, A., Alepis, E., Pocs, M., & Patsakis, C. (2018). Backups and the right to be forgotten in the GDPR: An uneasy relationship. *Computer Law & Security Review*, 34(6), 1247–1257. <https://doi.org/10.1016/j.clsr.2018.08.006>
- Primack, B. A., Shensa, A., Sidani, J. E., Whaite, E. O., Lin, L. yi, Rosen, D., et al. (2017). Social Media Use and Perceived Social Isolation Among Young Adults in the U.S. *American Journal of Preventive Medicine*, 53(1), 1–8. <https://doi.org/10.1016/j.amepre.2017.01.010>
- Ryan, R. M., & Deci, E. L. (2001). On Happiness and Human Potentials: A Review of Research on Hedonic and Eudaimonic Well-Being. *Annual Review of Psychology*, 52(1), 141–166. <https://doi.org/10.1146/annurev.psych.52.1.141>
- Schieffelin, E. L. (1998). Problematizing Performance. In F. Hughes-Freeland (Ed.), *Ritual, Performance, Media* (pp. 194–207). Routledge.
- Seaver, N. (2018). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, 1–16. <https://doi.org/10.1177/1359183518820366>

- Shakya, H. B., & Christakis, N. A. (2017). Association of Facebook Use With Compromised Well-Being: A Longitudinal Study. *American Journal of Epidemiology*, 185(3), 203–211. <https://doi.org/10.1093/aje/kww189>
- Shoemaker, D. W. (2010). Self-exposure and exposure of the self: informational privacy and the presentation of identity. *Ethics and Information Technology*, 12(1), 3–15. <https://doi.org/10.1007/s10676-009-9186-x>
- Swatman, R. (2016). Pokémon Go Catches Five New World Records. *Guinness World Records*. <https://www.guinnessworldrecords.com/news/2016/8/pokemon-go-catches-five-world-records-439327>. Accessed 11 August 2019
- Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, 4(2), 1–14. <https://doi.org/10.1177/2053951717736335>
- Tian, X. (2017). Embodied versus Disembodied Information: How Online Artifacts Influence Offline Interpersonal Interactions. *Symbolic Interaction*, 40(2), 190–211. <https://doi.org/10.1002/symb.278>
- Villaronga, E. F., Kieseberg, P., & Li, T. (2018). Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten. *Computer Law & Security Review*, 34(2), 304–313. <https://doi.org/10.1016/j.clsr.2017.08.007>
- Vogel, E. A., Rose, J. P., Roberts, L. R., & Eckles, K. (2014). Social comparison, social media, and self-esteem. *Psychology of Popular Media Culture*, 3(4), 206–222. <https://doi.org/10.1037/ppm0000047>
- Wachter, S., & Mittelstadt, B. (2019). A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, 2019(2), 1–126.
- Westlund, A. C. (2009). Rethinking Relational Autonomy. *Hypatia*, 24(4), 26–49. <https://doi.org/10.1111/j.1527-2001.2009.01056.x>
- Whitley, E. A., Gal, U., & Kjaergaard, A. (2014). Who do you think you are? A review of the complex interplay between information systems, identification and identity. *European Journal of Information Systems*, 23(1), 17–35. <https://doi.org/10.1057/ejis.2013.34>
- Williams, R. W. (2005). Politics and Self in the Age of Digital Re(producibility). *Fast Capitalism*, 1(1), 104–121.
- Wu, T. (2017). *The Attention Merchants: The epic scramble to get inside our heads*. Vintage.
- Zittrain, J. (2019). The Hidden Costs of Automated Thinking. *The New Yorker*. <https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking>. Accessed 27 July 2019

- Zuboff, S. (2015). Big Other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75–89. <https://doi.org/10.1057/jit.2015.5>
- Zuboff, S. (2019a). Surveillance Capitalism and the Challenge of Collective Action. *New Labor Forum*, 28(1), 10–29. <https://doi.org/10.1177/1095796018819461>
- Zuboff, S. (2019b). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books.