The Digital Subject. Data and Persons in Calculative Infrastructures

Journal:	*Theory Culture & Society*
Manuscript ID	16-214-POST.R1
Manuscript Type:	SI: Posthumanities
Key Words:	Data, posthumanities, neural networks, social networks, individual, algorithm, subject

Abstract:	This essay explores the return of the subject in the computational context, which I address as a digital subject. This digital subject encompasses a digital identifier, correlations in data or a data profile, moving between biological characteristics and symbolic expression. I focus on the processes through which digital subjects are constructed by matching, correlating, modelling, as well as how they become enactive. The ways of pulling data together into a digital subject is often presented as a logic of fact, where data is equated with documentary evidence. Instead, I propose the notion of the distance in which digital subjects are produced. Indexicality comes from outside of data, whereas the regard for the thick distance becomes a mark of the form of knowledge. I conclude by arguing for a posthumanities approach that establishes the distance while allowing for different subjects to be called upon.

The Digital Subject. Data and Persons in Calculative Infrastructures

A few years ago, Chiara Bernardi, at the time a doctoral student of mine, was working as a consultant in digital marketing. She attended a lecture delivered by a representative of MaxPoint, an American advertising company entering British and European markets. MaxPoint combined Census data, Telmar data (a large database of advertising media information and consumer data), bank transaction data, private browsing, online activity data beyond cookies and a mapping algorithm that tracks peoples' behaviour based on location, to name just a few sources. MaxPoint's main solution is "HyperLocal Advertising." For as little as £5000 the service can group people with similar online behaviours living in the same neighbourhood into "Digital Zips" and deliver personalised ads. Though the scale is calibrated into neighbourhoods exhibiting commonalities at the household level, the lecturer proceeded to describe an individual dot on the map in terms of available data. That individual was most likely a woman, who lived at a particular South London street, ordered organic vegetables and many books that she kept, used London Overground, but was not on Twitter. As more details followed, Chiara realised that it was I.

What is the digital entity presented that she identified as me? What relation does it have to me? How do I relate to it? How is it able to stand in for me and enact me for advertisements delivered to the living I and thus alter me while being reliant on my activity? How is it produced outside of my awareness, mobilised, and recruited? It is clearly not I, and yet it's no one else than I. What other Is are out there, labouring in the legal, medical, industrial, aesthetic spheres?

To engage with this encounter, I propose the term *digital subject*. The digital subject is a concept that includes a subject of a data profile alongside a Facebook stream, a history of browsing or search engine queries, a record of mobile phone positioning, bank transactions, sensor data, structural codes of face or movement recognition, as well as mail inboxes, amongst other things. The digital subject thus moves between captured unique and persistent biological

characteristics and premeditated forms of symbolic expression, judicially inferred subjects of actions and performed identities. It is this very entanglement of physical, legal, sensual and cultural that warrants the use of the term.

Various digital subjects connect and feed off experiences of the living persons. The digital subject is therefore distinct from the living self, addressed in the continuity of its experience in critical psychology and psychoanalysis, though one does not necessarily need psychological tools to talk about it. Digital subject is an abstracted position, a performance, constructed persona, data, profiles and other records and aggregates. While it would be a fruitless task to disentangle the living self, for instance, from the digital persona, in defining the digital subject I follow a tradition of looking for forms of connection between subject (subject position) and subjectivity (seen as "more than the sum total of discourses from birth," a complexity of experiences possessing a non-derivative ontology) (Henriques et al., 1984; Blackman et al., 2008). A digital subject comes after the subject and announces a new search for the ways in which it connects to subjectivities of living persons.

It could be argued that subjectivity itself, partly an effect of cultural and disciplinary individuating techniques (Foucault, 2005) and a product of work on the self through technology, such as the diary (Kittler, 1990) is formed in structural coupling with its computational environment, making all subjects digital. Hybridity is a *de facto* mode of contemporary existence. The omnipresence of hybrid technical forms might seem to warrant suspension of the redundant "digital," whereas the modern "subject" as a site of thought, moral judgment and action, has been understood as conceptually responsible for the atrocities of patriarchy, colonialism, and many other practices of subjugation. However, there are new forms of constructing the subject, the digital subjects, which arise out of computational procedures and are employed by various forms of power to singularise, map and capture not only subjectivities, but also non-humans and physical things that inhabit the world. Digital subjects become a new arena for struggle. In this paper I aim to demonstrate the distinctness and usefulness of the term by

discussing digital subjects in terms of some of their modes of production and critical perception. I aim to look at how they perform by focusing on what the subjects of me that are constructed digitally are and how they connect to the living I that is always a multiplicity (Braidotti, 2008). The digital subject therefore builds on the established critique of the subject to engage with the processes of its construction, recruitment and enactment that occur in the computational regime. Even if digital subjects are solely a response to neocapitalist desire for construction of fixed personhood, we firstly need to study them in detail to be able to imagine the alternatives.

Scientific notions of identity (such as the unique chemical composition or inner structure of atoms) are employed in the construction of digital subjects to exercise capture. For example, the face becomes the new fingerprint, and movement recognition technology claims it is impossible to move in ways non-identical to one's unique self. Here, digital subjects are mapped onto living persons in attempt to eliminate excess or non-coincidence. Yet, digital subjects also offer sites of reinvention, liberation and play. Fake accounts, performed identities and creative expression are witness to that. It is this very promise of simultaneously identitarian uniqueness, intimate access, legal answerability, wealth of creativity, and a scene for the performance of subjectivity, which characterises digital subjects.

In recent years, notions related to the subject resurfaced with great force and have proliferated. In the exact sciences, it often appears as an unproblematic notion of self or person: the digital self and digital personhood. In social sciences and cultural studies it is discussed as identity (constructed in social media) and as the data citizen (Isin and Ruppert, 2015). In critical theory, it is often a "data double" or a "data shadow" (Raley, 2013). Drastically different in their implications, these terms indicate that new computational subjectification and subject construction processes trigger ontological and epistemological anxiety, while new modes of measurement, evaluation and prediction offer significant monetary incentives and promise new kinds of politics. As a term, the digital subject is called to bridge broadly critical theory-based considerations of selfies, Facebook streams, and search queries in terms of "mediated digital

selves" and the big data analytics-based debates on the construction of data profiles in terms of probability, prediction and control. If the former is often narrated as: "prob(ing) the gap between the perception of our own identity and its distributed representation via social media platforms" (V&A, 2015), the latter is framed in terms of a critique of data analytics as operationalizing and preempting the world largely outside of existing systems of accountability, whether scientific or conceptual (Rouvroy, 2013).

The digital subject thus poses questions that can't be answered within one discipline. While the humanities have a long tradition of thinking subject and subjectivity, for computer science the human is an array. When critical theory engages with data production, it often veers all to soon into the discussions of surveillance. This is massively important, and yet such a well-trodden rut effectively prevents close, conceptually rich examination of the processes comprising data analytics. A political choice here is not to dismiss certain forms of rationality or formalism too easily or too early. As many paradigms and disciplines collide, calling for a radically interdisciplinary posthuman enquiry, the digital subject asks for new ways to understand causality and determination, unity and disunity, interiority and exteriority as well as reason, expression and sense. In this paper, I focus more specifically on the question of the connection between the living person and data, or its digital subject, which I explore by addressing the promise of indexicality of such a connection. Instead of indexicality, I argue, there is distance. I ask again and again: what is the digital entity presented that Chiara identified as me? What relation does it have to me? How do I relate to it? How is this connection made?

Distance

Key to conceptualizing digital subject is the notion of distance. In Valla and Benenson's artwork *Some Sites and Their Artefacts: 123D Catch* (2014), an automated camera glides through 3D rubble, presenting decaying digital objects, from a lion's head to antique busts, created using Autodesk 123D Catch software. An automated voice narrates the decay of 3D objects in the

virtual space, starting with the following statement: “A digital artefact is neither an object nor its representation but a distance between the two." It can be adapted into: "A digital subject is neither a human being nor its representation but a distance between the two". This phrase is key to capturing the anxiety around these notions.

I propose the notion of *distance* to discuss the relation between the digital subjects - and the humans, entities and processes they are connected to. Distance instead of relation poses a very important difference. Distance is not representational; it introduces change. Such distance is Leibniz’s distance: luminous and thick. As non-empty, qualitative distance, it can expand and contract, stretch and collapse; it is full of interference. In his writing on Leibniz, Deleuze specifies that the fold comprises many planes, as it is always a fold of folds. The fold doesn’t consist of points or parts, but is elastic and unites different forms of matter (Deleuze, 2006: 4-6). The distance of the fold, for Deleuze, must contain a metamorphosis: there is always a transformation occurring in the distance. The distance is also temporal; its limit is not quantity, but space-time. Such a distance is not an arrow from 0 to X: it is vast and twisted, found between two constantly changing entities. The person, the starting point of such a distance, is not necessarily a human being. It could be a human body moving through transport networks, but also a smart home, a sensor on a pigeon’s foot, fictional story or figure, a collective of anarchists sharing one mobile phone, car number plates, together with a hundred other alliances, long and short-term. These persons could be things that can enter into relations without necessarily possessing subjectivity. At times it is a compulsively socialising teenager and at other times it is a SIM card shared among a village.

“Its representation” is not a representation. Data are not traces (Rouvroy 2013, Gitelman 2013). The causality operative in the distance is computational, at least partially. Here, the debate’s legacy is shared by concerns that surrounded photography at its inception. Thus, Balzac wrote in *Cousin Pons* (1847): "If any man had come to Napoleon to tell him that a building or a figure is at all times and in all places represented by an image in the atmosphere, that every

existing object has a *spectral intangible double* which may become visible, the Emperor would have sent his informant to Charenton for a lunatic …Yet Daguerre's discovery amounts to nothing more nor less than this" (Balzac, 1847: 584-5). In the post-medium situation, the new media is data, and its genealogy is to be found in the science labs, scientific experiments and statistics, rather than the legacy of art: landscape and aura as distance (Benjamin, 1936). After decades of discussion about the status of indexicality in photography, there are new ontological concerns about the image following a change in the material production of a digital photograph, too (Osborne, 2013: 127). Osborne's argument here, for instance, is that technological specificity is always "fantasmatic" and its social uses of indexical signification are never secure (2013: 124).

While abstraction in data analytics is not linearly causal, it operates in the distance, obtaining models, deducing statements, making inferences. Apps offering to cross-match Instagram feeds and IP addresses, and then IP and street addresses orchestrate multiple alignments that are beyond representation. The ways in which they are aligned create digital subjects. Still, in posing the question in the manner of Benenson and Valla, there is a danger of constructing the (human) subjectivity and the digital subject categorically and in the manner of the one-to-one correspondence. To counter this tendency, the distance must be posited as non-dualistic by thinking through its forms of contraction and twisting as grounded in the specificity and variability of embodied techno-cultural and socio-political practices rather than any idea of essential qualities. It is also based on the wonder of the events of encounter with digital subjects similar to the one I had (Stengers, 2010). In such an event, I am confronted with what someone or something recognises or uses as me, a record and a model of whatever I could be. As digital subjects are constructed not only to sell products but also to imprison, medically treat or differentiate between individuals, the non-coincidence and spatiality of the distance become an urgent political matter.

Therefore, claiming representational, indexical, correlationist or reinventive distance is not only a matter of funding or disciplinary norm, but a question of politics. A humanist would claim the distance as the site of rivalry, while a media scholar might focus on the distance as the producer of authenticity. Hackers hide in the distance, while artists work it. Some sciences would accept the distance, - while others claim there is no distance, but equivalence (a human equals her Tweets; we can understand society by studying Twitter). Struggle over the distance is political through and through: to claim and maintain distance is a matter of keeping and developing alternative options of becoming and living with our digital subjects.

In this paper, the question of how my digital subject is produced in relation to me and what it produces in return is ultimately a question of distance. The distance is interrupted, recruited, intersliced. When one is abused online, the distance collapses, and when a security-mad enthusiast employs an echelon of tools to hide electronic footprints, it may become tense and dense. Distance can be maintained and manipulated: a technical disturbance or other processes can be disguised in the distance; hence it can serve as an artificial enhancement (for instance, a long iris scan performed by border control can disguise multiple other checks and cross-references). Playing with a camera's image filters, probing the limit at which the face would become non-recognisable, is maintaining the distance, pulling on it and relaxing it, - an important part of the technical-subjective maintenance of distance.

To sum up, the forms of production of distance, whether by neural networks, platform infrastructures, projects or engagements are key to the generation of digital subjects. The fluctuating distance is manufactured, created and played with. It is through the actualisation of the distance that digital subjects acquire value, service or seize aesthetic sense. It is also through the distance that digital subjects become more or less personalised or multiple, put together and disaggregated. The distance is a concept that specifies the new relationality that allows addressing both ontological and epistemological questions brought about by data modality,

which, allowing for a meeting point between disciplines, can be interrogated by problematic of various genealogies and open to questioning.

The distance must be established; it can then be claimed. Less "valuable" persons, such as women, people of color or poor people can be assigned digital subjects that continue discriminating them. Some people in the global south do not get to have digital subjects. The computational production of digital subjects is not "naturally flowing," "objective" or transparent. Claiming the distance brings attention to the processes of production of digital subjects and the ways in which their relation from and back to, their indexicality, is established. It then becomes something that can be interfered with, redirected, played with, and reinvented.

In what follows below, I focus on discussing this distance in relation to existing notions of indexicality. I then examine neural networks as a specific case of the production of distance. I will conclude by drawing upon art practices as a distinct form of maintenance of the distance and argue for hybrid forms of actualization of distance that posthumanities are uniquely positioned to develop.

Distant indexicality of data

Constant data generation is the condition of today. While data is not direct measurement, it is unclear whether and what it is a representation of; and as data generation is not exactly an on-going occurrence such as photosynthesis, its "naturalness" and scale often invite biological metaphors. As things and humans go on about their business leaking data, this informational condition is to be considered a new ecology of computation. The key to understanding ecology is that there is no simple causality (Stengers, 2010: 32-33). Claiming, establishing, overcoming and doing without linear forms of causality, and the questions of correlation, remain one of the core themes in the critique of big data.

The discussions around data generation are often framed in terms of demography and the vocabulary of statistics. Rita Raley, for instance, writes about the "digital demographic self:" a

person with data produced by her that includes approximate age and gender, changes in bodily states (getting pregnant, getting ill), location, consumption behaviour, interests, the social networks he or she partakes in, and other aspects (Raley, 2013). Raley includes demographics and psychographics (interests and behaviours, often called "proprietary interest data" in advertisement) into the demographic category. This is a digital census citizen, a consumer, possibly a patient or a suspect.

There is a lineage in documentary modernity, as Ronald Day argues, from documentation that works as an indexical sign to prove a phenomenon, and that is “any type of ontological substance that acts as evidence” (Day, 2014: 5) to information: a transferal of knowledge from the document, a process of informing, a response to the need (for knowledge) (Day, 2014: 37, 42). Descriptions of the generation of digital subjects, which are themselves responses to requests, are well rehearsed. Only the recent, hot “data footprint” matters. If one has browsed an online store’s sofa department, they will be served ads of exactly those or different sofas on websites they visit next. As it is highly unlikely that the interest in sofas will persist indefinitely, the value of this interest data will decline with time. With only recent activity valuable in this context, a holistic picture is not strived for: this digital subject does not attempt a full psychological or economic portrait.

What kinds of digital subjects are there? One visits a sofa’s website and is assigned a unique number that is matched with a specific product or/and an interest category of sofas and living room furniture. This match is already a digital subject. An identifier could be an IP address, use Wi-Fi triangulation, or be a cookie. The matching expands from deductive to inductive. As more data is processed, correlations are found within it. These correlations are also digital subjects: interest in sofas, together with purchasing power, geographical position, age, preference of particular products and styles. This data is turned into profiles: models applied to correlations in the individual data, rules and data on others. Here, induction (and abduction) becomes possible: judging by similar behaviours, it is possible to infer that one might be inclined

to wish for a designer coffee table from a specific store or compute a prediction linking a computational model and an individual data. Profiles are also digital subjects. The long duration and consistency are gone here. The fact that one would need a reading light by the sofa will not necessarily be part of such a profile as it is not a spy's report on one's personality as a compulsive reader.

Rita Raley points out that data and subject are indexically and repeatedly linked: the digital subject is constantly produced, re-instantiated, re-engendered, refreshed (2013: 123). The digitally, demographically and psychographically centred digital subject is in fact a set of dynamic processes that have the structures of computational actions, models, and socio-political cultures. It is a process in which no exact or stable state is significant or valuable: what matters is algorithmic interpretation at the moments data can be used, sold or otherwise acted upon.

According to Day, documentary signs function within a broad socio-technical regime of information organisation that gives order, value and meaning to "texts and person as documents…" (2014: 20) For Day, indexes, a mode of modern documentary tradition, similar to Osborne's distributed photographic complex as a mode of visuality described above, changed from "explicit structures" to "implicit devices," where both documents and users merged into data (2014: 2) at the same time as the socio-technical regime or the infrastructure itself disappeared from view (2014: 38, 47). For Day to some extent too, as for Raley, social media act as "evidentiary fragments" in modern information infrastructures (Day, 2014: 29). Assigning an identity through a documentary process, he argues, is core to modernity, which today extends to individuals asserted through attributions, reduced to "conjoined data points" and indexed in the current computational information infrastructure (Day, 2014: 59-61). However, in his wonderfully argued book, Day also writes: "The problem … is that it may appear … that what is given by the information is a 'fact:' that the constellation of references is closed, that the text… is a document" (2014: 66). Here, it is clearly the problem of the distance.

While individual data points can be tendentially evidential, they are not documents or evidence, and profiles even less so. Digital subjects are values, dynamically re-instantiated correlations, rules, and models, - overall, shreds of actions, identities, interests, engagements, which are put into relation to each other, disaggregated, categorised, classified, clustered, modelled, projected onto, speculated upon and made predictions about. Profiles are future oriented. The digital subjects: values, patterns, models, - unconnected and entangled - are distributed and distributive. As models can be relaxed or tightened, profiles are evaluated in terms of probability. They spawn different computationally spatio-temporal scales: in terms of length of alphanumerical strings, proximity, complexity, units, forms of composition and proposition, frequency and type of future, - all aspects of the distance in operation.

For assessing risk in insurance, for instance, with its specific periodicity and temporality, an individual doesn't live long enough to correspond to enough points at which data is acquired to offer accuracy. While with other kinds of data (Facebook data or supermarket buying patterns) the degree of correlation can be more accurately decomposed and thus they can be more meaningful for specific products, the individual scale doesn't de facto and universally give better indications. The digital subject here is not personalized; it corresponds to an aspect of a population. A digital subject in fact rarely corresponds to a classically constituted individual: it is always more and less than a human. I would argue that digital subjects aren't ever computationally aggregated into an order and composition that corresponds to the classical subject of modernity. Such aggregation is possible, though laborious, and relies on turning data into evidence and establishing indexical links (more on this below).

While promising extreme individualization, and indeed capable of granularity, digital subjects are processes of scaling up and down in their distributed-ness. The momentary request of web-based advertisement and permanent data aggregation of the Five Eyes programmes pose radically different temporalities and kinds of threat, relying on the some of the same sets of computational procedures. Ranging from a value to other data, rules and the model making a

proposition about a profile or a proportion of the population, digital subjects change form or identity: they become different things. Yes or no in a row; the table itself; the model; are not a spatial scaling of digital subjects, but forms of recording, aggregating, establishing relationality, and prediction where each has its own spatio-temporal framework. These are the processes and frameworks that comprise and characterise the distance though they are not able, on their own, to exhaust or explain it away.

Digital subjects offer new forms of singularity and multiplicity. The basic dialectic of today is the promise of individualisation, of standing out via the creation of a singular occurrence; yet such a singularity can only be produced through constant aggregation, comparison, sorting and re-arrangement of other singularities in the operation of multiplicity ([redacted]). Singularities can only exist through their relation in the multiplicity: they are determined from the interactions of multiplicities. The oscillation between singularity and multiplicity is not circular. In the case of profiles, correlations in the data linked to an individual are used in a model that refers to a group of people, from which something can be inferred about the individuals whose data points might be missing. Such individuals will not only not be the people whose data is crunched, but they don't need to be specific persons altogether (van Otterlo, 2013: 44). They might be *likely* nameless candidates for fitting a profile. Determination of singular events (evaluating, targeting) is enacted from such an operation of multiplicities, in which singularities pulsate in an out of existence. The likelihood becomes the condition under which a singularity might be inferred from multiplicities, so it doesn't even need to exist or be proven right. Yet, a multiplicity is aggregated from singular points, rules and models.

On request, digital subjects are pulled together, into plastic aggregates of subjects, associations between them, a set of propositions and probabilities. While digital subjects are produced, distributed by software, they are also pulled together within the computational socio-political infrastructure. They are flecks of identity (Fuller, 2005), distributed computational processes, – with a capacity to become concentrated, singular, linked to individuals, whether

precisely or likely. The subject here has been lost in abstraction, reduction, scales and directionality, and yet constantly arranged to present forms of correspondence, correlation, and association. There is a performative and processual action that dynamically and constantly produces, compares, updates and manages digital subjects: they are co-constructed by the computational infrastructure, the way in which software records, marks and structures activity.

How do digital subjects map onto the producing, data generating individuals? Digital subjects enact realities and forms of action. What do they enact: subjects born every time anew, which, at their distance, still manage to maintain (do they?) continuous relations with the data generating individual or her multiple non-unitary subjectivity? If the beginning point of this data generation is a body, demographic self, citizen, psychic, cognitive or communicative processes, and aggregates of those and the end point is a constantly changing collection of identifiers, values, patterns, models and profiles, how are both parts continuously affected by each other and reinstantiated through the distance that links them? Digital subjects arise out of data generated about something, and become active in the computational infrastructure that enacts something else in return. Do these three somethings manage to coincide through a distance? Are they connected tightly or loosely and through which mechanisms?

For this dilemma, Raley suggests "the data double." For Raley, data doubles are discrete and constantly aggregated data bodies, Deleuzian dividuals, rather than subjects, the "figures and products of modernity" (Raley, 2013: 127). In this line of thinking, we are becoming our data, granulating our subjectivities in the data worlds we inhabit.

Yet the double is like a shadow. The shadow arguably belongs to the representational paradigm in understanding science. Pickering writes: "The representational idiom casts science as, above all, an activity that seeks to represent nature, to produce knowledge that maps, mirrors and corresponds to how the world really is. …Within the representational idiom, people and things tend to appear as shadows of themselves" (Pickering, 1995: 5-6). One could argue that a radical critique in the tradition of the humanities conceptually corresponds to the well-critiqued

representational paradigm in the philosophy of science. Worse than that, such a paradigm is put into service by big business and politics, assigning the "usefulness" of science and data regimes on the grounds of the modern quest to conquer by abstracting, reducing and representing. Confirming what it is arguing against, critical humanities are facing an acute interdisciplinary challenge: the need to develop a notion of distance relying on the performative idiom of science.

Digital subjects do not have absolute representational correspondence to humans and their collectives. They are not doubles. *Indexicality comes from elsewhere*. Certainly, Facebook advertising service insists that it gives access to "real people," maintaining that there is a relation back from the double, an indexing relation. It is however clearly a speculative relation, one that is beyond indexing. The return and blossoming of behaviourism is undeniable, and the questions of privacy and human rights are urgent. With the equivalence habitually claimed between the "digital demographic selves" and the data generating individuals, there is definitely a danger, a question of control. But with indexicality appointed from outside of data relations, it must also become a question of destabilising it, with the notion of the distance being pushed to the fore.

Distance is a topologically twisted time-space. A human, or a certain solidarity across scales, and digital subjects, or particular computational formalisations, form a Moebius strip: one follows it, but always ends up on the wrong side. We are writing a story about ourselves and in the story are the subjects we become, the subjects of the story as we write it. We're the subjects of the book that acquires concrescence through writing. This is the subject as a position, as abstraction, construction: it subordinates, controls and acts. Except there is no writing: the process is computational. The analogy with writing can be very misleading: the story is patterns, similarities, models, clusters, which are sorted, re-arranged, stored, sold. There is a movement, we write ourselves by generating data that is worked upon and then produced as digital subjects, non-consistent and not very coherent, and for different requests: some for advertisement, some for secret services, some for our own consumption. These digital subjects do not coincide with any originating we: they are at a distance. Yet, as I return to this point again and again, there is a

legal, industrial, and techno-scientific pull for the computed digital subjects to map onto human beings.

It is easy to discern why the techno-scientific complex or jurisprudence would talk about information being attached to a specific identity. After all, an identifiable person can be assigned debt or a prison sentence. Is there something more in the pull? Data values can have some evidential relationship but they are not absolute. Moreover, as I've shown, models abstracted from data or, in some cases, coinciding with data (such as k-nearest neighbour algorithm) are not factual at all. Data can be actualised as pointers or witness through meticulous work – as, for instance, in the projects of the *Forensic Architecture* group. Here, it requires bespoke methodologies, which Eyal Weizman and his team call techniques of the interrogation of images, meta- and social data. Such techniques interpret and position data (images of a drone attack, video footage, aerial images and maps, reports) as evidence that is supplied to the UN, International Criminal Court, high courts and commissions. Here, distance can be actualised into an evidential relation, producing, as Day argues, "the documentary … representation that … transforms the possibility of identity … into … fact," (2014: 3), but data is not a priori testimonial.

Indexicality comes from elsewhere. Forms of indexicality may include the collapsing of differences, assigning sites of authenticity, and stitching together data worlds, via tropes of unity and persistence. Key to such rhetoric of the digital subject is that it is formed in the distance that, while distributing the individual, can and sometimes does simultaneously establish it. Indexicality looms upon us, creating individuals, because modernity calls upon individuals. Using Althusser's interpellation (subjects being constructed through being hailed by authority in the manner of "Hey, you there!") and algorithmic interpolation, Day suggests that it is because of the training of modernity to become subjects via a response to the call, that such data singularities are pulled together into an individual (2014: 78). To function within orders of modernity, we must respond to the call upon us: "It is a sociocultural, technological, modern

documentary system that is calling one, as one… First of all, it demands that one has formed a notion of the documented subject within ones' self… One must have internalised the law of the subject before the call of the law or any other moral order" (Day, 2014: 80). It is upon this training from birth that one responds to being hailed by the algorithms.

It is then with the proposition of the posthuman and posthumanities, which dispose of the subject as the metaphysical centre of the modern order, that the ascertaining of the distance can begin with. Citizen, body, demographic self, identity and individual are usually terms of unity: subjectivities as assemblages of complex and intensive forces are a form of resistance to the logic of modernity. The generation of the digital subjects in the multimodal distance that is itself a site of contention incorporates resistance to The Subject, continuing the project of Deleuze and Guattari, Braidotti and other scholars. Digital subjects can be established as evidence, or they can be used to re-construct the modern subject, collapsing the distance, and calling upon digital surrogacy. But the options are multiple. Digital subjects are illusory subjects, imagined by networked machines as profiles, likelihoods, probabilities. The non-coincidence of the digital subjects either with data sources or with the actions they enact ascertains the distance. Sometimes it is vast, at other times it collapses. Sometimes it is evidential, at other times it is purely speculative. It is misleading and yet in this very fog lie the possible routes to escape.

What do digital subjects look-alike? On neural networks

The specificity and range of operations generating digital subjects require close reading of modelling, data structures, services and infrastructures. To understand digital subjects is to develop posthumanist descriptions and cross-disciplinary vocabularies of computational socio-political operations bringing them about. In the short section below, I focus in particular on *Lookalike Audiences*, a Facebook advertising service, and on some of the neural networks that, together with other machine learning engines, run these kinds of products.

Facebook offers numerous advertising services. The data it holds on its users include mail order and online purchase frequency, credit card balance and "tendency to make indulgent purchases," level of spend, technology at home and interest in fashion, life-stage, and affluence, among others. The Lookalike Audiences offers a potential expansion of the customer base. For example, a company sells products online. Facebook places a "pixel" on the purchase webpage and tracks customers' behaviour elsewhere on the web, building a database. It then runs it against its own databases, looking for profiles with behaviours that might look like the company's customers, subsequently targeting them. Here, a group of attributes (Facebook claims that it has over 100,000 attributes) is compared and matched with a different degree of proximity, using different algorithms. More matching attributes will offer a good fit (small scale campaigns) and less will increase the lookalike yield while decreasing the lookalikeness.

How does distance operate here producing the digital subject of the lookalike? The "lookalike" moment is most likely resolved by a range of neural networks. Neural networks are part of machine learning, a branch of computer science that studies algorithms and systems that improve their knowledge or performance with experience" (Flach, 2012: 3). Models that see what looks alike are results produced by machine learning algorithms as they are trained on and as they work through data. Neural network models are indeed called models, and yet they differ from the mathematical statistical modelling that is intended to produce accounts that provide causal explanations that could be mapped onto "real systems". David Byrne, for instance, argues that modelling "is formalising by abstracting," where abstracting is the ability "to manipulate mathematically" (2002: 113). Neural networks are computing systems that, like simulations, can deal with complex non-linear cases. Byrne claims: "… when we deal with non-linearity and emergence we can't turn to any formal modelling because we won't be able to establish an analytical solution to equations. Differential calculus will fail us in a non-linear case. Neural nets can handle non-linearities" (2002: 131, 136). Whereas variables in models were causes, in neural networks they are cases. Neural networks are case-centred, and "what emerges from the

procedures are sets of cases rather than models" (Byrne, 2002: 95). Neural networks' models can be seen as an aggregation of individual cases and they act as propositions and probabilities. In one lookalike construct filed for patenting by Facebook in 2015, such propositions and probabilities are described in terms of "likelihood," "expectation," "possibility," and "possibility of preference" (Cheng at al., 2017).

Continuously optimised neural networks are abductive, prospective, and predictive (Mackenzie, 2017), prescribing a new form of action. As neural networks specify a mode of learning, rather than a mode of action, it is understated. Here, the new military techniques are a withdrawal of the categorically understood military action; just as stated in the introduction to this special issue, there is no specific difference between the state of war and the state of peace. With patterns and correlations instead of linear causality, action is sourced opportunistically and dynamically. With tactics and strategies merged together, a desired effect can be reached through a combination of influences none of which is decisive or openly decided upon. Results-driven data science corresponds to results-driven interventions, lacking in strong causality, but rapid in anticipatory self-optimisation.

As causality in models gives way to analogies, continuous measurement and the production of sets of cases, their validity is established by their usefulness. Jimmy Lin in particular argues that when big data substitutes causation with correlation, it acts within a paradigm of engineering (what works best) rather than science (with its regime of objectivity) (Lin, 2015). Better engineering rather than better (scientific) understanding also means that what works does not need to correspond precisely to anything too specific: this is certainly the case with Lookalike. Byrne summarises similar findings as: "while neural nets are complex systems, we really have no way of establishing if they correspond in any meaningful way to any real complex system … no correspondence can be proven between outputs of neural nets and the real world." (2002: 139 -140). The distance build by neural networks is particularly stretched;

operations in the world of their own, they have a "working" relationship with something they model and simulate.

Byrne chooses to describe neural networks as "idiot savants," tools and "icons" (2002: 141), sharing in what they stand for. Digital subjects produced by such operations, the set of cases, the matches between attributes, seem to have various amounts (from some to very little) of relationship to us: representative correspondence is not on the menu. They are more separate and further away than demographic selves, doubles and shadows. Yet, these almost separate entities have a form of partaking in us. Like our icons coming alive, "sealed angels" (Leskov, 1873), they are also engineering tools. The digital subject, as itself not an entity or a scientific fact, cannot be understood on the basis of identity, casually determined from the living being. If there is indeed an iconicity correlatively produced in the stretching and twisting distance, it is only glimpsed through the moments of exclamations: "It works!" Such iconicity has no relation to the angels, but to the process of its own manufacturing or even fabrication.

Performing a Digital Subject

In the *Politics of Aesthetics*, Rancière briefly writes about the logic of fact and the logic of fiction (2006: 45). If the logic of fact is established via the apparatus of the production of objectivity and is operative in science-related work, the logic of fiction guides the creation of literary or art work and is no less rigorous. Literary scholarship, for instance, offers analytical tools of high value that can be employed to understand the construction and power of a literary *oeuvre*. The logic of fiction guides writing that becomes operative and established in a field of text beyond authorial power as much as the production of scientific fact depends on the infrastructure of science to enliven it. Both constructions of forms of engagement with and production of reality, which still somewhat escape analysis (Stengers, 2010), the logic of fact and the logic of fiction operate in domains that have previously been distinctly different. Science and art, model and novel, data and narrative enter into new kinds of partnerships today. The logic of

fact arranged by data science carry some forms of fiction and the logic of fiction within which contemporary digital artists operate call upon the logics of fact.

In art, it is the digital media and feminist work, where some aspects of the digital subject resurface quite spectacularly today as digital identity or persona. A self to be over-performed, a distance to be filled in with such baroque eagerness that it becomes a place to hide are characteristic of digital feminist artists. In what follows, I will discuss one particular project, an Instagram performance by Amalia Ullman, *Excellences & Perfections* (2014) that propelled her to stardom while triggering significant debate.

The project consists of photographs posted over 5 months that tell a story of a 25-year-old provincial girl. She moves to Los Angeles, breaks up with her high school boyfriend and wants to be a model, dying her hair blond. She then runs out of money, "gets a sugar daddy," has a lavish lifestyle, gets depressed and starts doing drugs. Then Amalia's alter ego "gets a boob job" at the request of her lover and has a nervous breakdown. After that, she apologizes to her followers, dies her hair dark and goes back home, posting photographs of avocado toasts and hinting at a caring boyfriend. The project uses the real name and feed of the author and consists of her photographs of herself in carefully chosen environments, yet is totally fictional. Both the narrative and the visuals are beyond stereotypical: they correspond to the patriarchal regime of female positioning in society. Ullman is presented as very young, vulnerable, and powerless in her glamour. Yet, it is documented that many young women see Ullman's project as a reflection on how they're compelled to manufacture their digital persona through digital network portraiture.

The project has caused significant debate precisely because it re-enacts the dominant repressive visual regime whereas its satiric nature rests solely in the author's claim that appoints it an art project. Ullman claims to have taken time and effort to master the craft of the construction of feminine beauty online and that it is in this very exposure through re-enactment that the critical message lies. Mixing registers of art discourse and cliché, Ullman's project

creates a certain digital persona as a figure, a position of sensibility, transposing and creating what it senses. Ullman's fake identity functions within a logic of fiction that employs the visual terror of the stereotypical.

However, what digital subjects of Ullman would data analytics produce, following her data in their logic of fact? Ullman used her previous hospital photographs and medical tape to pretend to be recovering from breast-enlargement surgery. She was repeatedly at the different locations to where she claimed to be, hunting for spots with good lighting, luxurious hotels and shops, employing a battery of tools to manicure her fictional photographic persona. According to the logic of fact, her digital subjects would be partly inaccurate (for instance, the author would not need the products advertised to her on the basis of her performance), though they would not be limited to the Instagram feed. The answer, however, lies altogether elsewhere. For data analytics, the real Amalia Ullman has no more importance than the figure with a particular sensibility that she produced. Ullman constructed the "girl next door," "sugar baby ghetto girl" and "life goddess" as three dominant mainstream trends in how women present themselves online (Kinsey, 2016; Corbett, 2014). The girl, the baby and the goddess function well enough in the symbolic regime of American and other western societies to be successfully employed in the logic of fact and become streams of digital subjects: the data to be trained on, to be fed into models, to be produced as profiles. For the scenarios of probability, real or fictional, authentic or imagined are equally useful.

The logic of fact and the logic of fiction are arranged very differently in Ullman's project than in my encounter with my digital subject. And yet, the logic of fiction is fed into the logic of fact without any restrictions. While the logic of fact promises factuality, in many registers it is filled with and thrives on fiction, even if this fiction is mostly culturally boring and aesthetically bad. A call for exciting fiction due and aside, this is also a condition in which the question of the real is reinstated.

Discussed under "authenticity" in relation to the questions of digital identity or enhancing the stereotypical scripts in pattern-recognition-based culture, the question of the real is decided differently depending on the discipline. On the one hand, the scientific real of the human as physical reality that can be known is itself a metaphysical claim. On the other hand, the real that operates through culture and its narratives, figures, and metaphors, a symbolic real, is today merged with the "real" of other domains. Stengers writes: "For each practice, it is on the basis of the definition of what is designated as 'reality' and what will be asserted as 'value' that the scope, implications, and problems of requirements and obligations can be specified" (2010: 53). Data regimes don't distinguish between bodies and novels, nature and culture - it is a process of recording and manufacturing about everything. The symbols or representation of reality were always contested constructions. The peculiar character of data is a bizarre situation of having to defend established designations of reality in science, law, linguistics or psychology or to offer a blanket critique of data management of the real. As the data modalities threaten traditions of defining reality, it is now the turn of data to invent its own modes of knowing the real. What it will be and how it will be constructed is something that must be accepted as an invitation to take part.

Conclusion

The digital subject is neither necessarily an extension of the human into digital networks (it is not a self), nor a representation of the I, but comes into being at a distance between the living being and the data pointers, profiles, models and active propositions that it may prompt. Such distance doesn't have the power to create indexical digital subjects, as our data shadows, or map onto humans precisely, and may not even be concerned exclusively with human selves – either particularly or in general. There is no given or natural factuality or evidentiality in digital subjects. The first part of the paper explored such distance against representational claims, whilst

paying close attention to the shape of some digital subjects, dispelling any idea of their unity, persistence or human scale. The second part looked more specifically into some products that operate using neural nets, among other machine learning approaches. Constructing digital subjects in terms of possibility is also a form of operation of distance that is not linearly causal.

Another level of the paper is an arts and humanities encounter with data sciences, and contribute to the digital media theory's forms of enquiry. As such, it is a methodological proposition; one that concerns itself with forms of humanity while working with the changed notion of the human. Both the subject and technology are the focus of this paper as a contribution to this posthumanities special issue.

If modernity calls upon us, seeking to claim indexical relations between data, models, profiles and living beings, it is not the only force. By operating in the distance, producing digital subjects that enact realities, there are other calls, other passions, knowledges, organisational forms and layers. Different purposes and uses will construct different digital subjects, with a possibility of reformulation. As we are in the process of self-calibrating to the new digital data forms that are being introduced, the posthumanities can become a form of alliance that claims the distance and calls upon digital subjects to manifest beyond man and into new practices of reality formation.

Bibliography

Balzac, Honore de (1847) *Cousin Pons.* In French available at: https://fr.wikisource.org/wiki/Le_Cousin_Pons, accessed 7.09.2016.

Benjamin, Walter (1936) The Work of Art in the Age of Mechanical Reproduction. Available at: https://www.marxists.org/reference/subject/philosophy/works/ge/benjamin.htm, accessed 7.09.2017.

Blackman, L., Cromby, J., Hook, D., Papadopoulos, D., Walkerdine, V. (2008) Creating Subjectivities, *Subjectivity* 22: 1. doi:10.1057/sub.2008.8

Braidotti, Rosi (2008) *Nomadic Subjects. Embodiment and Sexual Difference in Contemporary Feminist Theory*. Columbia University Press, 2008.

Byrne, David (2002) *Interpreting Quantitative Data.* London: Sage.

Cheng, Haibin, Xu, Xian and Yang Pei (Inventors) (2017) *Loolalike Evaluation. United States Patent Application 20170140283.* Assignee: Facebook, Inc. Filed 13 November 2015. Publication date: 18 May 2017. http://www.freepatentsonline.com/20170140283.pdf, accessed 7.09.2017.

Corbett, Rachel (2014) How Amalia Ullman Became an Instagram Celebrity. In: *Vulture*, December 14. Available at: http://www.vulture.com/2014/12/how-amalia-ulman-became-an-instagram-celebrity.html, accessed 7.09.2017.

Day, Ronald E. (2014) *Indexing It All: The Subject in the Age of Documentation, Information and Data.* Cambridge, MA: MIT Press.

Deleuze, Gilles (2006) *The Fold. Leibniz and the Baroque*. London – New York: Continuum.

Flach, Peter (2012) *Machine Learning. The Art and Science of Algorithms That Make Sense of Data.* Cambridge: Cambridge University Press.

Foucault, Michel (2005) *The Hermeneutics of the Subject. Lectures at the College de France, 1981-1983.* New York: Palgrave Macmillan.

Fuller, Matthew (2005) *Media Ecologies. Materialist Tendencies in Art and Technoculture.* Cambridge, MA: MIT Press.

Gitelman, Lisa (2013) (ed.) *"Raw Data" Is an Oxymoron*. Cambridge, MA: MIT Press.

Henriques, J., Hollway, W., Urwin, C., Venn, C. and Walkerdine, V. (1984*). Changing the Subject: Psychology, Social Regulation and Subjectivity*. London: Methuen.

Kinsey, Cadence (2016) The Instagram Artist Who Fooled Thousands. In: *BBC Culture*, 7 March. Available at: http://www.bbc.co.uk/culture/story/20160307-the-instagram-artist-who-fooled-thousands, accessed 7.09.2017.

Kittler, Friedrich (1990) *Discourse Networks 1800/1900.* Stanford: Stanford University Press.

Leskov, Nikolai (1873) The Sealed Angel. *Russian Messenger*, January.

Lin, Jimmy (2015) On Building Better Mousetraps and Understanding Human Condition: Reflections on Big Data in the Social Sciences. In: *The Annals of the American Academy,* 659, May.

Mackenzie, Adrian (2017) *Machine Learners. Archaeology of a Data Practice.* Cambridge, MA: MIT Press.

Osborne, Peter (2013) *Anywhere or Not At All. Philosophy of Contemporary Art.* London – New York: Verso.

Pickering, Andrew (1995) *The Mangle of Practice. Time, Agency and Science*. Chicago and London: The University of Chicago Press.

Raley, Rita (2013) Dataveillance and Countervailance. In: Gitelman, Lisa (2013)

Rancière, Jacques (2006) *The Politics of Aesthetics*. London – New York: Continuum.

Rouvroy, Antoinette (2013) The End(s) of Critique: Data Behaviourism Versus Due Process. In Hildebrandt, Mireille and de Vries, Katja. *Privacy, Due Process and the Computational Turn*. London – New – York: Routledge.

Stengers, Isabelle (2010) *Cosmopolitics I.* Minneapolis: The University of Minnesota Press.

Ullman, Amalia (2014) *Excellences & Perfections*. Available at: https://instagram.com/amaliaulman, accessed 7.09.2017.

Valla, Clement and Benenson, A.E. (2014) *Some Sites and Their Artifacts: 123D Catch*, In *Rhizome,* 8 May. Available at: http://rhizome.org/editorial/2014/may/8/some-sites-and-their-artifacts-123d-catch, accessed 7.09.2017.

van Otterlo, Martijn (2013) A Machine Learning View on Profiling. In: Hildebrandt, Mireille and de Vries, Katja. *Privacy, Due Process and the Computational Turn*. London – New – York: Routledge.

V&A Digital Futures – Self Versioning: Exploring the Potentials of the Digital Self (2015) Announcement. Available at: http://www.spacestudios.org.uk/art-technology/va-digital-futures-self-versioning-exploring-the-potentials-of-the-digital-self, accessed 7.09.2017.