

Article

Uncertain Archives: Approaching the Unknowns, Errors, and Vulnerabilities of Big Data through Cultural Theories of the Archive

Daniela Agostinho

University of Copenhagen, Denmark
bkc683@hum.ku.dk

Catherine D'Ignazio

Massachusetts Institute of Technology, USA
catherine_dignazio@emerson.edu

Annie Ring

University College London, UK
annie.ring@ucl.ac.uk

Nanna Bonde Thylstrup

Copenhagen Business School, Denmark
nbt.msc@cbs.dk

Kristin Veel

University of Copenhagen, Denmark
Kristinv@hum.ku.dk

Abstract

From global search engines to local smart cities, from public health monitoring to personal self-tracking technologies, digital technologies continuously capture, process, and archive social, material, and affective information in the form of big data. Although the use of big data emerged from the human desire to acquire more knowledge and master more information and to eliminate human error in large-scale information management, it has become clear in recent years that big data technologies, and the archives of data they accrue, bring with them new and important uncertainties in the form of new biases, systemic errors, and, as a result, new ethical challenges that require urgent attention and analysis. This collaboratively written article outlines the conceptual framework of the Uncertain Archives research collective to show how cultural theories of the archive can be meaningfully applied to the empirical field of big data. More specifically, the article argues that this approach grounded in cultural theory can help research going forward to attune to and address the uncertainties present in the storage and analysis of large amounts of information. By focusing on the notions of the unknown, error, and vulnerability, we reveal a set of different, albeit intertwined, configurations of archival uncertainty that emerge along with the phenomenon of big data use. We regard these configurations as central to understanding the conditions of the digitally networked data archives that are a crucial component of today's cultures of surveillance and governmentality.

Introduction

The accumulation, storage, and management of information in today's big data repositories, while unprecedented, raises important questions that have long been at the heart of cultural theories of the archive. Indeed, big data repositories constitute highly networked archives of constantly growing data streams, which we here argue must be addressed as part of the history of archival phenomena. Moreover, the big data

Agostinho, Daniela, Catherine D'Ignazio, Annie Ring, Nanna Bonde Thylstrup, and Kristin Veel. 2019. Uncertain Archives: Approaching the Unknowns, Errors, and Vulnerabilities of Big Data through Cultural Theories of the Archive. *Surveillance & Society* 17(3/4): 422-441.

<https://ojs.library.queensu.ca/index.php/surveillance-and-society/index> | ISSN: 1477-7487

© The author(s), 2019 | Licensed to the Surveillance Studies Network under a [Creative Commons Attribution Non-Commercial No Derivatives license](#)

archives into which information is now gathered by means of surveillance and automation also call for a revisiting of the critique of archival reason, a critique that was made possible first by post-structuralist thought and later by feminist, queer, postcolonial, and critical race theories. In their own ways, these theoretical and political movements have challenged the authority of the archive as a reliable repository: its capacity to produce truth, offer evidence, and categorise human identities. From these cultural-theoretical perspectives, archives have always by definition been uncertain, and this uncertainty has not decreased as the archives have become larger. On the contrary, digital environments repurpose and refashion the logic of the archive under conditions of social, political, ecological, and technological uncertainty. Post-structuralist critiques of the archive, we argue, thus still provide valid and compelling entry points into the understanding of contemporary information storage and its logics, in conjunction with exciting new critical archive theories that both draw on and expand theoretical and political horizons beyond Jacques Derrida (1995), Michel Foucault (1995, 2007), and Gilles Deleuze and Felix Guattari (2013). This means drawing on feminist, queer, postcolonial, and critical race theories of the archive, such as those by Saidiya Hartman (2008), Diana Taylor (2003), Rebecca Schneider (2011), Ann Cvetkovich (2003), Ann Laura Stoler (2002), Sara Edenheim (2013), Jack Halberstam (2005, 2011), Michelle Caswell (2014), Marlene Manoff (2004), Marika Cifor (2015), and Tonia Sutherland (2017a), among many others. Although they come from different disciplinary vantage points, taken together these theories offer a critical approach to archival reason and practice in the present day.

Current practices of data production, collection, distribution, and consumption both build upon and draw from the cultural history of the archive, as well as raising pertinent new questions that exceed the horizon of physical archives. Thinking in this way allows us to recognise the historical roots of current practices of data hoarding, storing, leaking, and wasting, while also reminding us that today's seemingly streamlined interaction with digital files and folders is every bit as messy, porous, and generative as archival encounters have always been. In this article, we therefore address the large digital archives that have come to characterise our time as the latest instalment in a long negotiation between surveillance technology and its subjects (or objects), between control and uncertainty, order and chaos, and ultimately between power and knowledge. With this we hope to demonstrate that big data, while technologically new, in fact belong in a long historical trajectory, one that we show is haunted by historical and contemporary uncertainties familiar from critiques of archival reason. This article has the further function of outlining the conceptual and theoretical framework that has governed the Uncertain Archives research project, which has run at the University of Copenhagen from 2015 to 2019 and guides a forthcoming glossary of terms relevant to the conceptualisation and analysis of big data phenomena.

Our approach constitutes a critical methodology that is responsive to—and offers modes for addressing—archival subjects, movements, and affects. For these purposes it is essential not merely to regard artistic creation as a separate and objectified sphere for academics to critique, but to recognise and facilitate arts-based research as a productive mode of critique and knowledge production in itself (Dirckinck-Holmfeld 2011). This way of looking by no means implies that scholars should leave aside their criticism; rather, we aim to carve out an epistemic space for modes of inquiry that are motivated less by providing answers than by posing questions, exploring uncertainties, and offering material and speculative approaches to the challenges posed by big data archives.¹

In what follows, we address uncertainty through three related conceptual lenses—the *unknown/unknowable*, *error*, and *vulnerability*—which we propose as crucial concepts for exploring the uncertainties of the large data archives in which we are interested. We home in on the spatial and temporal connotations of each lens from a cultural-theoretical perspective, and we highlight several arts- and practice-based research projects that provide valuable contributions to understanding the unknowns, errors, and vulnerabilities implied by big data archives. These projects all call for more in-depth analysis, as indeed does each section of the typology we propose. However, we aim with this article to present a cultural-theoretical approach to big

¹ Katrine Dirckinck-Holmfeld, La Vaughn Belle, and Kristoffer Ørum are some of the artists who have been part of the research team and have contributed to the exploration of archival uncertainty through their artistic practice.

data archives that can open new avenues for thinking about big data and the research methodologies best suited to investigating them.

Researching the Uncertainties of Big Data Archives

Retrieving and repurposing the notion of the archive allows us to think of big data as a regime of knowledge, power, and control. As cultural theories of the archive have repeatedly noted, the function of archives exceeds the mere storage and preservation of data, as the epistemological decisions made in the compilation of an archive fundamentally define what is archivable, what is knowable, and hence what is deemed disposable and forgotten (Agostinho 2016; Ring 2014). The archival operations associated with big data therefore raise important political and epistemological questions that have been addressed by the “archival turn” in the humanities (Stoler 2002): questions about access, selection, exclusion, authority, lacunae, and silences.

In his etymological analysis of the word *archive*, Derrida foregrounds the important interplay between authority and interpretation. As he reminds us, *archive* derives from the Greek *arkhé*, which names at once a commencement and a “commandment” (Derrida 1995: 9), thus combining the principle according to history (where things begin) with the principle according to the law (where authority and social order are exercised). But the etymological meaning of *archive* also stems from the Greek *arkheion*: the residence of the archons, the superior magistrates who command and guard the law, and who possess the “hermeneutic right” and competence to interpret it. The “violence of the archive” (Derrida 1995: 12) is played out in its selective operation (storing some elements at the expense of others) and its hermeneutic privilege (ascribing the stored elements with meaning). This is possibly why, for Carolyn Steedman, “in Derrida’s description, the *arkhé*—the archive—appears to represent the now of whatever kind of power is being exercised, anywhere, in any place or time” (2001: 1,159). Nevertheless, the moment when information is captured by means of surveillance technologies has particular significance for understanding how the logic of archiving shapes the operation of power (Ring 2014). Our argument here is that both gestures that the archive carries out—selection and interpretation—gain new epistemological and political implications when we look at them in relation to surveillance in terms of the collection and use of big data.

These central archival gestures have been scrutinised by scholars across different fields who have examined the limitations and possibilities of the archive. Within performance studies, thinkers such as Taylor (2003) and Rebecca Schneider (2011) have questioned archival logic and its exclusion of (or failure to integrate) repertoires of embodied knowledge formed by gestures, voices, movement, flesh, and bone. At the same time, these theorisations have reconceptualised the archive to propose that embodied practice offers alternative perspectives to those derived from conventional archival inscription. Such critiques are taken further by feminist and queer theories of the archive that have pointed to how archival reason has overlooked the experiences of women and queers, and how these histories are often obscured within existing sources, or discarded altogether (Stone and Cantrell 2016). Literary theorist Ann Cvetkovich (2003) has famously argued for an “archive of feelings,” for the need to preserve everyday queer experiences that are difficult to chronicle through the materials of a conventional archive; historian Sara Edenheim (2013), in response to this claim, argues that the traditional archive is often a queer body of knowledge in and of itself, a disorderly, contingently organised place rather than a site of systematic order where information gives itself to easy retrieval. For feminist and queer scholars, the archive often emerges as a place for the recovery of suppressed or marginalised histories, a recuperative project of moving from silence to productive discourse. Similar debates have taken place within African-American, Caribbean, transatlantic, and postcolonial studies in relation to the archives of slavery and colonialism. Here, too, scholars have questioned both the capture and exclusion of people of colour in and from archives and the kind of knowledge that can be gleaned from the archives of the ruling classes, archives that dehumanise those under colonial rule (Hartman 2008; Copeland and Thompson 2011; Helton et al. 2015; Edwards 2016). At the same time, these studies have reminded us that the archives of slavery and colonialism continue to inform who counts as a human subject today (Gikandi 2015), while also urging us to understand and actively mobilise archives as instruments for social and restorative justice.

Alongside those critical approaches, the field of critical archival science has questioned the metaphorical use of the concept of the archive in humanities scholarship, wherein “the archive” emerges “as an abstract, depopulated space, untouched by human labor and laborers” (Whearty 2018). Critical archival scholars have urged humanities researchers to consider “actually existing archives,” as well as to acknowledge the intellectual contribution of archival science scholars, in order to advance critical work on archival reason (Caswell 2016). Such critical work would benefit highly from mutual dialogue between these disciplines. Critical archival science has been pivotal in questioning archival praxis through the lenses of feminist, queer, post-, and decolonial studies, drawing attention to affective responsibilities in archival practice, the often invisible labour of archivists, the materiality of digital archives, the ethical challenges of archiving sensitive material, the need for collective work with marginalised and vulnerable communities, and the relevance of archives for human rights, social justice, and care (Caswell and Cifor 2016; Cifor and Wood 2017; Caswell, Punzalan, and Sangwand 2017; Sutherland 2017a). These critiques of archival reason and practice can be mobilised towards a critical analysis of big data’s repositories: they show that the piecing together of information is not a neutral pursuit, both capture and exclusion have important ethical consequences, and archives are always contested sites of power, knowledge, possibility, and aspiration. In many ways, as we shall see, big data archives extend and transfigure the problematics entailed by traditional processes of archival reason.

Although these critiques of archival logic acknowledge and address uncertainty as inherent to concepts of the archive, the emergence of datafication—and its embeddedness in regimes of neoliberal global governance and increasing precarity—means that uncertainty and risk have become functions of disruption complicit with power. Our present times are thus imbued with a “spirit of uncertainty” that animates the cultural backdrop of social life (Appadurai 2012: 7). Ulrich Beck (2009) famously imagined the risk society as teetering on the brink of catastrophe and disaster, in a manner that prompted social scientists to invent and deploy ever more managerial technologies intended to minimise risk. Uncertainty, he suggested, implied the government of the incalculable, a regime in which one could only ever govern through estimation. From this perspective, uncertainty emerges both as a feature of the event (which cannot be calculated statistically) and as a way of imagining the future by non-statistical methods (O’Malley 2004). Uncertainty, then, exists at the crux of today’s governance dilemma between security and freedom: if modernity’s governance systems had demanded scientific predictability, universality, and rationality, today’s globalised economies also celebrate uncertainties (O’Malley 2009).

Nowhere is this tension between risk and uncertainty, and the presence of the two conflicting desires for control and uncertainty, more visible than in today’s big data archives (Amoore and Raley 2017). On one hand, data companies and governments across the globe welcome big data as an effective solution to deal with informational uncertainty, risks, and unknowns. Its promise of accurate calculations, precise predictions, and pre-emptions speaks to contemporary concerns about the taming of social, economic, financial, environmental, and political risks. On the other hand, as N. Katherine Hayles (2017) has demonstrated in her work on high-frequency trading, the very same companies and governments also frame big data as drivers of creativity and high-gain opportunity. Uncertainty and control are thus equally embraced by techno-capitalism.

If we situate a rethinking of the archive within this regime of uncertainty—which is simultaneously and never unambivalently political, technological, and cultural—we can understand big data archives not simply as rational apparatuses, but also as reflections of a political and social reality where uncertainty is profoundly feared and yet simultaneously embraced as potentially disruptive (Thylstrup 2018). Increasingly, however, we also see moments of resistance when individuals and collectives turn towards uncertainty, despite its co-optation by risk regimes, as an engine of creativity and innovation that may be remobilised to undermine archival authority. Uncertainty, if viewed from this angle, has the potential to change the way information is held and employed and to provide marginalised individuals and groups with opportunities to contest dominant organisations of information. In the following sections we demonstrate how this remobilisation of uncertainty works, analysing it in turn through the conceptual frames of the unknown or unknowable, error, and vulnerability.

Mapping: Unknowns/Unknowables

When we look at big data and the apparently insatiable desire they represent for mapping everything (including the self), we are reminded of André Breton's ([1924] 1969: 9) criticism of realism in his "Manifesto of Surrealism": "our brains are dulled by the incurable mania that consists in reducing the unknown to what is known, to what can be filed." The surrealists were concerned with the high degree of structure found in unconscious objects such as dreams and subconscious actions, as well as with the impossibility of capturing and showing it fully. For them, archival practices were devoted not so much to imposing order on contingencies, but rather to how to work with the existing order of the unconscious (Spieker 2008). Issues concerning the unknown and the unknowable are classic archival problems, as well as contemporary problematics, that are no less pertinent to the calculative realm of large data archives. The unknown and unknowable are the ultimate generators of uncertainty. This is a spatial as well as a temporal problematic, and it represents the other side of the coin from the extensive mapping practices of information-intensive environments. A working hypothesis for us has therefore been that the unknown and unknowable are key concepts in today's data landscapes, the archives they generate, and the speculations to which they give rise.

Thinking about unknowns invariably leads on to thinking about maps. We suggest that mapping is a fundamental archival–technological function and indeed is a drive to make unknowns known and render them available for detection and domination. A historical example of the unknown's speculative forces and imaginary landscapes is the depiction of uncertain and speculative territories by ancient Roman and medieval cartographers, occasionally with the phrases *hic sunt leones* (here be lions) or *hic sunt dracones* (here be dragons). Rather than leaving unknown territories blank, these cartographers populated them with fantastical creatures that would instil in the map's readers both a fear of the unknown and a desire to colonise it. Indeed, the inscription of the unknown with the visual vocabulary of fairy-tale monsters served as an attempt to justify the history of violence in the conquest of those lands. The colonial history of mapping practices offers a productive insight into contemporary mappings of the unknown. At the end of the seventeenth century, sea monsters gradually disappeared from maps. As technology changed our understanding of the oceans and navigation advanced, more emphasis was placed on the colonial imaginary of humans' ability to master the water: to sail on it, conduct trade on it, and colonise it. Thus, images of the sea's unknown dangers became less frequent over time, while images of the ships that conquered the ocean and everything belonging to it became more common. The monsters that had once been feared were turned into commodities to be colonised, managed, and mastered.

Today, thanks to the arrival of big data technologies, the science of mapping has reached new levels. There are few areas of the world—including human life itself—that are not constantly in the process of being mapped, archived, and calculated with a view to detecting patterns and eliminating or colonising unknowns. But now the tools have expanded from cartographic instruments to encompass algorithmic models, thus introducing new kinds of uncertainties (Kwan 2016). And the motivations have expanded from colonial conquest to encompass societal securitisation, civic management, capitalist exploitation, healthcare advancement, the handling of climate change, and crisis prevention (Halpern and Günel 2017). In addition, algorithmic mappings chart not only spaces but also temporalities (Parisi 2013). They create new archival folds in time: the future is folded back into the past, insofar as the future that an algorithmic system can predict is limited by the historical data used to train that system. If the input data are racist and sexist, the output will most likely also be racist and sexist. Yet the conquest of unknowns is a highly generative and speculative act that gives rise to spatial and temporal imaginaries and a wide variety of coping and exploitation strategies (Kwan 2002). These imaginaries function as mirrors of our societies: they tell us something about how to cope with unknowns and about our affective relationship with the unknown. Artist Mimi Onuoha (2016) points directly at the unknown in her project *The Library of Missing Data Sets*, a list of data that go uncollected because of lack of political will, discrimination, or systemic neglect. The list includes topics such as "unarmed civilians killed by police" and "unpaid internships." She exhibits the work as a set of tabbed file folders, which upon opening reveal themselves to be empty (Figure 1).

Or take one of the most persistent tropes of data discourse: the black box. In mid-twentieth-century metaphorical terms, the black box is the paramount symbol of that about which we can only know the input and the output. It has come to represent what we do not—or cannot—know. The black box expresses not only the unknown and unknowable, but also the human desire to know more, to make the opaque at least predictable, if not transparent. It is a question of how we cope with the invisible—with that which has a presence, yet which is not present or comprehensible with the tools at our disposal (Steiner and Veel 2018). Viewed in this way, the question of the unknown is also a question of methodologies—of the tools available to us to understand the world. Paradoxically, the new archival mapping mechanisms embedded in big data have not only created new forms of knowing and predictability, they have also given rise to new modes of unknowing (Rosenberg 2014; Murakami Wood 2015).



Figure 1: Mimi Onuoha (2016), *The Library of Missing Data Sets*. Courtesy of the artist.

Moreover, the archival mechanisms embedded in big data make apparent a significant challenge in the use of big data: namely, that the empirical past, no matter how comprehensively collected, can never preclude the arrival—or attack—of the unknown and the unexpected. The question, of course, is why a fascination, indeed obsession, with unknowns and unknowables has arisen at a time when we supposedly track ourselves more than ever before (Schüll 2016). Perhaps this fascination can be understood by turning to psychoanalysis, which inspired the surrealists and has unlocked and mapped the human psyche—and in the process has configured the psyche in the realm of the unknown or unknowable.

Melanie Klein (1931), following Freud's model of the constituent ego instincts of scopophilia (pleasure in looking) and epistemophilia (pleasure in knowing), saw the epistemophilic instinct as exploratory and necessary, but also aggressive. She located the emergence of epistemophilia—and thus the seeds of intellectual life—in the child's phantasies of getting inside the mother to explore and, as often, take over (as it were, to search and destroy) (Klein 1931; see also Spillius et al. 2011). As Toril Moi (1989) notes, epistemophilia is self-defeating, always frustrated by the limitations of the body. The Freudian drive for knowledge is structurally incapable of achieving total insight or perfect mastery. Moreover, and perhaps

just as interestingly, Moi suggests that if reason is always already shot through with the energy of the drives, the body, and desire, being intellectual cannot be theorised as simply the opposite of being emotional or passionate. This speaks to the gendered dimensions of issues of the unknown. Often what can be known is presented as rational, while what cannot be known is characterised as emotional/subjective/irrational and associated with female subjectivity.

Lastly, we need to consider the temporal coding of the affective reactions of the unknown and the unknowable: the fear or allure of unknown future knowledge that we do not know whether we desire or hate because we do not yet know what it entails. Or the knowledge that in hindsight we regret having come to know and that we cannot make unknown again. This is brought out, for example, in architect Laura Kurgan's (2015–2017) interactive mapping project *Conflict Urbanism: Aleppo*, which seeks to map the devastation wrought on the city during the current civil war, while simultaneously archiving the cultural sites and neighbourhoods that were present before the conflict began (Figure 2).



Figure 2: Laura Kurgan (2015–2017), *Conflict Urbanism: Aleppo*. Two satellite images of Aleppo's Citadel and surrounding neighbourhoods on March 23, 2016 (left) and September 3, 2015 (right). Yellow squares indicate damaged sites identified by the United Nations in May 2015. Additional unmarked damaged sites are also visible. Copyrights: WorldView-2 X 2012 DigitalGlobe, Inc.; Pleiades-1A X CNES_2015, distribution AIRBUS DS, France; WorldView-2 X 2016 DigitalGlobe, Inc.

Indeed, archival memory today appears as a paradox. On one hand, there has never been a greater desire for inclusive archives, archives that can include more historical voices than that of the familiar white male protagonist (Borgen, Thylstrup, and Veel 2016). An example of this is the found-footage documentary by directors Manu Luksch, Martin Reinhart, and Thomas Tode (2015), *Dreams Rewired/Mobilisierung der Träume*, which digs deep into the vaults of film archives from the 1880s to the 1930s and brings out lesser-known stories of technological utopianism and connectivity—in particular, the place of women in this

history of technological progress (Figure 3). On the other hand, if archive theories have questioned archival logic and its exclusion of (or failure to integrate) the repertoires of embodied experience formed by gestures, voices, movements, and affects, it has become clear that big data archives increasingly record and sort such embodiments for the purposes of surveillance and profit.



Figure 3: Manu Luksch, Martin Reinhart, and Thomas Tode (2015), *Dreams Rewired*. The archival image in the still is from the 1925 black-and-white short film *KIPHO*, directed by Guido Seeber. Courtesy of the artists.

Such all-encompassing archives have never sounded more creepy and intimidating, as the subjects of large-scale surveillance increasingly recognise that being archived and mapped comes with new and unforeseeable risks. A new human right has even given a juridical name to this uncertainty: “the right to be forgotten,” which suggests that not all memory is desirable, and that it also matters in what context a memory is presented and into whose hands it falls (Sutherland 2017b). We may wish to unknow things, or we may wish for others to unknow things about us and our communities, as in recent European Union and Argentinian legislation. Archivists have been aware of this problematic for a long time, working under conditions that oblige them to black-box their material for seventy-five or one hundred years (Bowker and Star 1999).² Yet, as archival spaces expand beyond the tomes of state archives, and as even state archives are faced with new working practices in the digital realm, there is a need to problematise and reframe the stakes of archival memory and its unknowns. When do we wish to be remembered, by whom, and in what contexts? Who decides the terms of remembrance? What are the tensions between the permanence of the digital sphere and the right to be forgotten (Sutherland 2017b; Caplan-Bricker 2018)? These questions pose practical challenges as much as they represent a theoretical quandary.

Detecting: Error

A crucial concept for considering big data from the perspective of cultural theory, and one that is helpful for imagining possible responses to the problems posed under the heading of the “unknown,” is provided by the notion of *error*. Reflecting on the role of error in big data archives permits us to shift our focus from

² See also the recent discussion in *Electronic Book Review* on datafication and open access (Rockwell and Berendt 2017; Schram and Whitson 2018).

the question of how material is gathered to that of how the content of big data archives is classified and ordered. The central issues that emerge through the perspective of error are the processes for detecting the lacunae, slippages, or resistant material that haunt archival holdings, as well as potential errors in archival functioning in the form of technical malfunctions, information anomalies, and human wrongdoing. A valuable cultural-theoretical approach to these topics is provided by psychoanalysis, as Wendy Hui Kyong Chun notes. Chun suggests that network analytics can be read as “the bastard child of psychoanalysis,” (2018: 69) in the sense that relationships and meanings are constantly being constructed out of seemingly disparate activities carried out by surveilled subjects. Chun’s diagnosis playfully invokes the subjectivity of contemporary archival logics, attributing to modes of data analysis a familial structure from which contemporary network analytics can be seen to diverge as a “bastard.” But her approach does more than just toy with the sensibilities of archiving as a patriarchal practice. In fact, this provocation suggests that psychoanalysis, far from being an outdated mode of thinking in the digital age, can prompt scholars to think anew about big data and their analytics. We therefore welcome Chun’s proposal to maintain the presence of psychoanalysis in the study of contemporary archives, echoing previous works outlining the relations between archival logics and psychoanalytic perspectives (Spieker 2007). We propose thinking about errors in the contemporary archive not only in terms of the technical problems they can suffer, but also giving consideration to the psychopathological states of mind that digital archives might reflect or even display as cognitive agents themselves. We argue in this section that this model for thinking psychoanalytically about archival states of mind can take forward our understanding of the epistemological and political implications embedded in big data regimes, where uncertainty exists and indeed often thrives on different scales and with varying ethical results.

The silent workings of today’s big data archives—constantly analysing and diagnosing data in order to detect new connections, trace new patterns, and generate new insights in the form of computational outputs—make Freud’s hermeneutic method of “evenly suspended attention” especially pertinent for a psychoanalytic reading of big data analytics. In “Recommendations to Physicians Practicing Psychoanalysis,” Freud described a special way of listening that consisted “simply in not directing one’s notice to anything in particular and in maintaining the same ‘evenly suspended attention’ [...] in the face of all that one hears.” ([1912] 1956–1974: 111–112). Warning psychoanalysts not to let their ambitions overdetermine what they heard, he was aware of the risk that too narrowly focused attention might fall prey to error, missing some crucial analytical clue. Freud’s proposed method helped the analyst “avoid a danger, which is inseparable from the exercise of deliberate attention,” namely that of focusing on some parts of the information given by the patient and leaving out others, because “as soon as anyone deliberately concentrates his attention to a certain degree, he begins to select from the material before him.” He argued that this “rule of giving equal notice to everything” was the “necessary counterpart to the demand made on the patient that he should communicate everything that occurs to him without criticism or selection.” It was thus the necessary method, he argued, for preventing the analyst from selecting from the material “following his expectations or inclinations,” thereby putting himself “in danger of never finding anything but what he already knows” and falsifying “what he may perceive” (Freud [1912] 1956–1974: 112). Thus, Freud propounded this mode of attention to try to eliminate errors arising from biases in the mind of the analyst.

Errors are just as important in the calculative realm of big data as they were to the metaphorical archives of the unconscious with which Freud was working. Although Freud’s ideal of evenly suspended attention is historically situated and characteristic of the way the attentive faculty was perceived in his time, it is nonetheless fruitful to consider how today’s big data operations seem to aspire to that ideal in a no less utopian way. Evenly suspended attention seems to promise to capture everything and help navigate the monstrous excess of information generated by today’s digital apparatuses, removing the risk of human error and asserting instead that the data speak—or even sing—for themselves (Anderson 2008; Veel 2018). At first glance, big data appear to offer a mechanical corrective, a concrete and technical method to analyse, detect, eliminate, and neutralise human error. As such, big data discourses on error belong to a long tradition of error studies as an applied science practised in a wide variety of fields, ranging from psychology and economics to engineering and business analytics, which address error as a practical question of detecting and preventing (or exploiting) unwanted occurrences. At the same time, big data archives fundamentally

transform how errors are thought about in those fields, in so far as these archives' volume and scale supposedly make them less vulnerable to individual errors. Yet, the puzzle that troubles us remains: how can we conceptualise the epistemological and political implications embedded in these types of error regime, where uncertainty thrives on various scales? The problem at stake is what constitutes an error in the context of big data archives. What does the mode of detecting them entail? Which errors are tolerated in big data archives and which are feared? Which subjects are deemed errors by these archives? While posing such questions, we also need to remember the potential of errors to carry a complicit function within big data archives. Archival errors can give rise to fabrications, including quasi-life forms such as glitches and viruses. But the ends to which such errors are employed depend upon how the errors are inscribed in various strategic programmes, most clearly in the case of neoliberalism and its subversive counter-replies, in which failure can be employed to produce critique, but not necessarily a critique that subverts the system.

These problematics, while extremely important for analysing new data technologies, also inhere in a long historical trajectory of philosophy engaged with the interrelation between knowledge and truth. Moreover, to think about this area means being particularly attuned to the *experience* of errors—the sensory mark that they make for statistical outliers in particular. Freud's method of evenly suspended attention, and our use of the analyst's couch as a metaphor for data analytics, draws attention to the sensory mode in which error is a signal, the disruption of a flow, and a bodily experience all at once. As such, we may simultaneously desire and fear it. The difficulty of locating the political implications of error reflects how difficult it is to separate error detection from the political knowledge regimes of which those errors, whether productive or troubling, are a part. One pertinent example of this is Sarah T. Roberts's (2018) work on content moderation, which highlights the ideological framework of naming something an error as it plays out between humans, machines, and political regimes. As a form of uncertainty, error is thus simultaneously a spatial, structural, temporal, and political problematic.

Our claim is therefore that error is a key concept for understanding today's data landscapes, the archives that they generate, and the speculations to which they give rise. We regard what counts as an error in big data as an unresolved and complex question that pertains as much to the interdisciplinary realm of big data as to the conceptual landscape of error itself. In this way we think of errors in terms not only of technical problems but also of political conditioning and psychopathological states of mind that produce their own logics, which in turn have their own aesthetics and modes of resistance. These logics become clearer if we think about the spatial connotations of error.

In the realm of etymology, the term *error* incorporates the mobile notion of erring, producing an understanding of error as a deviation from a route, a departure from a principle. In the medieval era, the notion of error was often tied to questions of morality, to the operations of a higher order looking down with a surveillant gaze from above. Thus, St. Augustine linked the question of erring to the question of sin and morality, imagining God's path as orderly and straight, and any deviation from it as sin. With the Enlightenment, a more complex—albeit still religious—understanding of knowledge and its relation to error emerged, whereby error was increasingly understood not in terms of sin, but rather as an inevitable part of progress. Yet the notion of error maintained its spatial dimension, albeit now in the form of a labyrinth rather than a straight line. As historian David W. Bates (2002) notes, knowledge in the Enlightenment was often imagined as a journey, and so too was the method for detecting errors. These journeys produced a framework of knowledge in which error and knowledge were productively linked. Catherine D'Ignazio (2014) works with such issues in a contemporary framework in her browser app *Terra Incognita: 1000 Cities of the World*, which aims to help users step out of their personalised media filter bubbles. Each new browser tab the user opens presents her with a city that she has not read about and highlights local media articles she can read to learn more (Figure 4).



Figure 4: Catherine D'Ignazio (2014), *Terra Incognita: 1000 Cities of the World*.

Just as ideas about truth and error have been increasingly secularised, big data corporations are troubled by the same questions as Enlightenment thinkers: how to distinguish unproductive or aimless wandering from an adventure that might lead somewhere worth going? Today this becomes a question of how machines learn to detect information-bearing patterns from distracting noise.

There are also ethical questions at stake. The notion of error, wherever it arises, invokes a sense of normativity and an anxiety about its transgression. Throughout history, the detection of error was designed to both produce and contain deviant subjectivities—perhaps most infamously in the history of psychoanalysis, the deviancy of the hysterical woman. Furthermore, the categorisation of error in knowledge production remains gendered and racialised: femininity and blackness are rendered as irritations to the norm, as something to be corrected or simply repressed. For instance, the notion of “outliers”—data points that do not conform to the rest—can be problematised insofar as these “errors” produce a norm, particularly in the categorisation of non-binary gender (D'Ignazio, forthcoming).

With their methods of statistical aggregation, the analysis of big data archives can end up excluding outliers, extending and reifying existing norms, and producing marginalised subjects as error. Buolamwini and the Algorithmic Justice League have demonstrated that structural racism and unrepresentative data sets impact on how facial recognition software mis-genders women and people of colour, as a result of an over-representation of white men in benchmark data sets (Buolamwini 2017). As the project *Gender Shades* describes: “in the worst case, the failure rate on darker female faces is over one in three, for a task with a 50 percent chance of being correct. In the best case, one classifier achieves flawless performance on lighter males: 0 percent error rate” (Gender Shades 2017). The question here is not so much how to improve such technologies in order to attain more accurate results, but to ask whether such technologies should be developed in the first place and to whose benefit. As Anna Lauren Hoffmann puts it: “In fact, if facial recognition worked flawlessly, it would only make matters worse. It would simply ‘perfect’ unfair and stifling patterns of targeting and abuse aimed at historically vulnerable populations” (Hoffmann 2019: np). These concerns have been theorised by Ramón Amaro (2016), who considers how machine learning functions as a “threshold” through which blackness enters, often unwittingly, into social, political, and economic experience, and how racism is seen as an error in algorithmic intentions or human judgement, instead of as a systemic reality from which the machine directly derives. Of fundamental interest to us is what counts as error in big data archives, how errors produce and reiterate norms, and how errors are often framed as malfunctions when they actually signify systemic features. In Mushon Zer-Aviv’s (2013) experiment *The Turing Normalizing Machine*, reductive data sets and algorithmic bias are explored to look

at precisely how machine learning builds an extensive database of what is deemed social normalcy and what counts as deviation from it (Figure 5).

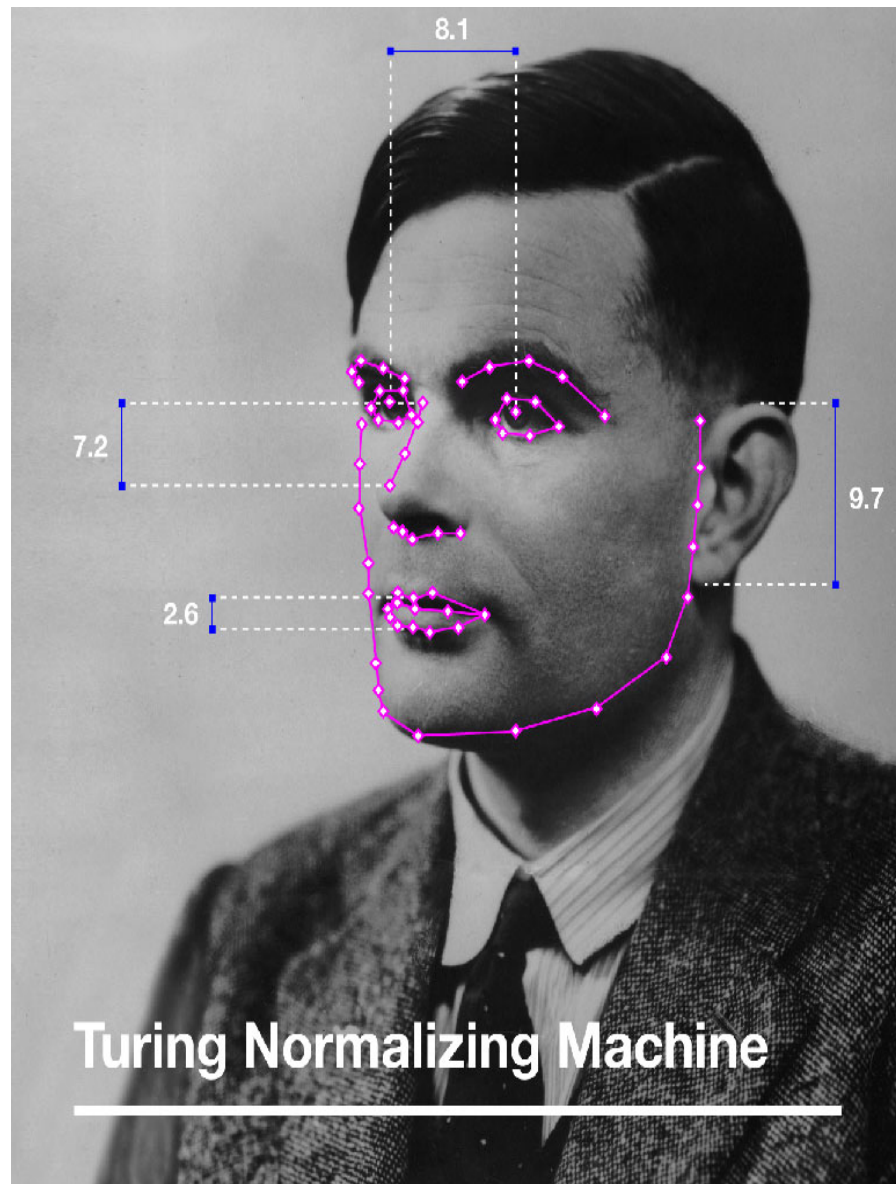


Figure 5: Mushon Zer-Aviv (2013), *The Turing Normalizing Machine*. Courtesy of the artist.

Thus, errors can be regarded as an inroad into critical engagement with big data archives as political sites of information distribution, rather than as objective statements of truth. Theoretical and artistic approaches that do this often treat those archives as holding subversive potential that can be used against the prevailing epistemological order. In this way, works such as those by Joy Buolamwini and the Algorithmic Justice League show that alternatives are embedded in the dominant culture—with the result that the power gathered by surveillance and stored in big data archives is never total, consistent, or omnipotent. Thanks to a subversive approach to error that recasts the potential for errors in the archive as a world of possibilities, a potential escape from calculation and predictability is established. As Jack Halberstam suggests, failure—an error that has become fundamental—can offer different kinds of reward: “under certain circumstances failing, losing, forgetting, unmaking, undoing, unbecoming, not knowing may in fact offer more creative, more cooperative, more surprising ways of being in the world” (2011: 3). Kara Keeling, meanwhile, notes that “errantry introduces relatively unpredictable contact between different organisms thereby raising

possibilities for new forms, intermixtures, knowledges and sensibilities to be crafted through the pressures, pains and pleasures characteristic of those contacts.” (2014: 51).

In relation to contemporary surveillance practices, we can observe instances of system error as productive interruptions of seemingly liberal environments that nonetheless closely surveil human behaviour in order to control attempts at resistance. Annie Ring (2016) argues that such productive interruptions are found in recent documentaries by Harun Farocki and Carmen Losmann, who explore new regimes of surveillance by documenting the constitutive disruptions that threaten their persistence. Indeed, as Mark Nunes notes, errors—in their failure to communicate—signal “a path of escape” from the “predictable confines of informatic control: an opening, a virtuality, a *poiesis*” (2011: 4). Glitch artist Rosa Menkman appears to agree with him when she notes that glitches offer “a not yet defined break from a procedural flow, fostering a critical potential” (2011: 27). The productivity of error becomes apparent not least in the digital sphere, where error itself has become both an aesthetic practice and an object of study. Thus, there are clusters of scholarship that focus on the moments when digital media go awry, on moments of breakdown, glitches, errors, or bugs. Rebecca Schneider (2011), for instance, has discussed Carolee Schneeman’s 2000–2001 video installation “More Wrong Things” as a production of error that interrupts but also shows how normalcy (of identity, patriarchy, structures of racism, war, waste, and exploitation) operates. Whether interruption or business as usual, error as a category can reveal to us the historicity of current archival operations of big data, helping us to draw attention to messy, wrong, glitchy archival imaginaries and to analyse their resistant potential.



Figure 6: Rosa Menkman (2010–2012), *Collapse of PAL*. Courtesy of the artist.

These desires for the glitch challenge the Enlightenment narrative of history as the overcoming of errors and the future as progress away from the mistake-ridden past, turning error instead into a powerful *poiesis* that challenges regimes of control. Yet we should remain mindful that, not least in recent years, error is also an approach that can find itself in alliance with political economies in which performances of critique are praised as ventures whose pursuit promotes and strengthens the economy. Silicon Valley’s fetishisation of failure is a case in point (Harvard Business Review 2011). Errors can even be seen as a crucial computational function whereby machines used for algorithmic securitisation can learn more accurate methods of control. From this perspective, the subversions that lend agency to those subjected to the uncertainties of big data

archives may at the same time be co-optable by the regimes of knowledge that those very subjects set out to resist.

Subjecting: Vulnerability

The category of vulnerability permits examination of the archival practice of *subjecting* and the conceptualisation of data subjects. Subjecting is taken here as the process of becoming subordinated, as well as the process of becoming a subject (Butler 1997: 2). As Judith Butler (following Foucault) explains, not only is a subject formed in subordination, but this subordination provides the subject's continuing conditions of possibility. The archival practice of subjecting thus echoes the notion that archival power does not merely act on a subject but also forms and enacts the subject into being. The subject thus emerges both as the effect of a prior gesture and as the condition of possibility for a contingent form of subjectivity: she is both acted on and enacted by. If one considers this form of subjection to be both external to and the very location of the subject, Butler suggests, this means that subjects are always constituted in primary vulnerability, acted on from the start by norms they never chose. If subjectivity is conferred from elsewhere, vulnerability or exposure to external forces—social norms, institutions, technologies—is necessarily a condition for becoming a subject.

While the primary state of vulnerability is recognised by both psychoanalysis and Foucauldian theories of power, for Foucault this process of subjectification occurs primarily through the body. In *Discipline and Punish* (Foucault 1995), the prisoner's body is the site of regulation and normalisation, but also the site of becoming a juridical subject. Foucault famously argues that the individual is formulated by his or her discursively constituted identity as prisoner. This theory of subjection, based as it is on subjects' internalisation of regulatory mechanisms, might thus provide a valid entry point to conceptualise the process through which individuals today readily subject themselves to the tracking, monitoring, and measuring whereby they become data subjects. But the Foucauldian perspective does not fully account for the material, psychic, and affective life of this data generation process. If vulnerability is the essence of being acted on by norms we never choose, what kinds of vulnerability emerge from our quotidian voluntary and involuntary (or semi-voluntary) engagements with data, and through what vulnerabilities do we become data subjects? What kinds of vulnerability emerge when subjection no longer occurs only through the bodily encounter with the institution, but increasingly through data? What kinds of vulnerability are amplified by data, or in the form of data doubles, a virtual and new site of vulnerability?

While emphasising that vulnerability is a shared condition of human (and non-human) life, theories of vulnerability also stress that not all subjects are equally vulnerable. Whereas the ontological notion of vulnerability foregrounds a common embodied humanity and equal susceptibility to suffering, a context-specific and politically inflected approach highlights the ways in which inequalities of power, capacity, or need render some subjects more vulnerable to harm (Taş, 2020). This is especially true for those who are marginalised in different circumstances due to their gender, race, sexuality, class, immigration status, ability, and more. How are subjects and communities differently affected by the encounter with big data archives?

For example, Stephanie Dinkins' *Project al Khwarizmi* (2017) stages conversations with communities of colour about the impacts of algorithms and artificial intelligence systems on their everyday lives, tapping into the differential distribution of vulnerability often exacerbated by the archives of big data (Figure 7). In her book *Dark Matters*, Simone Browne (2015) traces biometric surveillance back through the surveillance of black bodies during slavery—for instance, the accounting of bodies through ledgers, the branding of enslaved bodies, or the institutionalisation of lantern laws in the eighteenth century, which subjected the enslaved to heightened visibility. Through her reading of the archives of transatlantic slavery, Browne shows that current surveillance practices are part of a historical continuum of commodification and violence upon black bodies that render these subjects more vulnerable to current data archives.



Figure 7: Stephanie Dinkins (2017), *Project al Khwarizmi*. Courtesy of the artist.

These questions recentre the archival tension between capture and exclusion, that is, the dangers of being both included and excluded by data gathering processes, how datafication amplifies the commodification and visibility of specific bodies (Sutherland 2017b; Noble 2018), and how big data archives are embedded in historical relations of racial capitalism. Such issues are addressed by La Vaughn Belle's (n.d.) artistic work, which uses the framework of the ledger to add alternative records to the archives of colonial history in the US Virgin Islands, while also exploring material remnants of colonial times that shed light on current capitalist extractive practices (see Navarro 2018). The temporal continuum between the archives of the past and the data archives of the present is also explored by Katrine Dirckinck-Holmfeld (2017), whose video installation "The Christmas Report & Other Fragments" engages with the temporalities of the colonial archives of the US Virgin Islands once they become digital, while trying to attune to the bodies, voices, and movements of those recorded by the archives through the Afrofuturist figure of the Data Thief (Figure 8).

The differential distribution of vulnerability is key if we are to conceptualise and understand how data archives produce subjectivities by generating new vulnerabilities or reinforcing existing ones. In regard to the practice of "slut-shaming"—the release and public circulation of photographs and videos of women engaged in consensual and non-consensual sexual acts—Wendy Chun (2016) emphasises that technological infrastructures premised on the continuous exchange and flow of information ultimately subject the already vulnerable to new forms of vulnerability. Significantly, instead of arguing for a more private, contained internet, Chun proposes to rethink vulnerability as a way to disavow gendered violence and to fight for the right to be vulnerable, to be in public (both online and offline) and not be attacked (see also Agostinho and Thylstrup, forthcoming).



Figure 8: Katrine Dirckinck-Holmfeld (2017), *The Christmas Report & Other Fragments* (installation view). Courtesy of the artist.

Theorising the human body as dependent on and a part of infrastructure, technologies, organisations, life processes, and institutions demands a reconceptualisation of dominant ontological understandings of the embodied subject. How can we account for the datafied production of bodies without eviscerating the fleshiness of material lives and without reiterating soma-technical dualism? How are we to account for affect, desire, and imagination, and how are these organised, channelled, and transformed through datafied processes? Data motility—the process through which the data one generates increasingly move out of one’s control as they are mobilised for specific ends and values (alongside function creep and social sorting [Lyon, forthcoming])—can be diagnosed as a mode of vulnerability affecting the archival subject. Notions such as *data shadows* or *data doubles* articulate this anxiety about subjectivation and loss of control. Mark Andrejevic (2014), updating the digital divide of the 1990s, has noticed the “big data divide” that separates subjects from their data, a symptomatic condition of datafied subjectification. While we are responsible for generating the largest part of our data out there (by logging in, signing up, agreeing to, profiling), we are rarely granted access to the data, or to the purposes for which they are used, once they have been stored and rendered actionable by governments, institutions, companies, or acquaintances. Kristoffer Ørum’s (2017) “Ambiguous Physiognomy,” a live facial-tracking lecture, engages with these data shadows and the trail of information one leaves behind through navigation, which often gives rise to data subjects that we no longer recognise as ourselves. Similarly embracing the “creepiness” (Chun 2016) of data motility, Erica Scourti’s book *The Outage* (2014) is a ghostwritten memoir based entirely on her digital footprint. Profiles, intimate data, email correspondence, and other information, gathered with help from professionals in the fields of cybersecurity and data privacy, were given to a ghostwriter to construct a fictional memoir of the artist’s life.

In addition to having diminished access to our own data, we often lack the means and expertise to analyse them, make sense of them, or even recognise them as our own. This then creates new power differentials between those who (willingly and unwillingly) provide data and those who have the means to collect,

analyse, and deploy them. This divide is far from unprecedented, but it generates new vulnerabilities due to the shape-shifting nature of the data we hand over to often unknown recipients (Keyes, forthcoming). Such a process spawns an anxious feeling that our data gain a life of their own, circulate freely without our knowledge, and generate ever new data and unintended effects, ultimately creating a deeper and often imperceptible gulf between our embodied selves and our data lives.

As Judith Butler (1997) reminds us, none of us is quite commensurate with the norms that govern who we are supposed to be. Her notion of performativity—which posits that identity is enacted through performative actions, behaviours, and gestures beyond constitutive norms—indicates that subjection can be negotiated, that there is room for acting upon the processes of subjection that form us, and that identity can be shaped within a contingent space of subversion. But we are also witnessing how performativity is being co-opted and rendered profitable by political economies and regimes of knowledge under late capitalism. As many social and digital media scholars have argued, digital archives are archives in motion, running on dynamic processes of data collection and reproduction. The data they accumulate do not stay the same, and nor are those data fixed in time; rather, the data are put into flux, in a continuously generative process for profit and control. Companies can now engage in pre-emptive personalisation, letting us know in advance what we might want to become and possess in the future. This capacity of data to move on from the original signature and produce new possibilities echoes, but also threatens, the performative strategies of individuals in their engagements with data processes, giving rise to unintended performativities that widen the self-estrangement of the big data divide.

These kinds of vulnerability have prompted legal attempts to define and regulate the data subject, leading to the acknowledgment and formulation of the data subject's rights, such as the right to know if an institution or body is processing data about a subject, the right to information about the particular uses of that processing, the right to object to processing, and the right to be forgotten. Legal discourse thus attempts to intervene upon the conditions of possibility for subjects to engage with data, turning data vulnerability and risk into the very essence of online life by granting it legal protection. Yet, vulnerability as a touchstone for liberal ethics and legal frameworks has been questioned and challenged by intersectional feminist legal theory: there is a universalist quality to its implicit account of human nature, and biopolitics similarly depends on and profits from circulating narratives of risk and protectionism that are essential for legitimating discrimination and violence. These criticisms are often concerned with the conditions of possibility of becoming a vulnerable subject, with who is deemed vulnerable and worthy of protection, and with questions about who grants vulnerability status and how that vulnerability is reinforced rather than mended. Expanding Chun's discussion of vulnerability, we thus ask if vulnerability, rather than a shared ontological condition, can be recognised as an unevenly distributed and politically inflected condition of online and offline existence—and, hence, can be seen as a driver of responsibility and intervention that potentially shapes ethical and political existence within emerging data worlds.

Going Forward

This article demonstrates the ways in which cultural theories of the archive can be meaningfully applied to the empirical field of big data. In particular, we have argued, cultural theories of the archive help us to consider and conceptualise the new epistemological, empirical, and ethical uncertainties that have followed in big data's wake. Approaching big data archives in terms of unknowns, errors, and vulnerabilities, we draw on a broad range of contemporary theory as well as arts- and practice-based research, and we point to specific projects that promise to take this thinking further. Giving an overall outline of this typology of uncertainties of big data archives, our work here foregrounds the necessity for further cultural-theoretical inquiry into big data. In our forthcoming glossary, we bring together a multitude of scholarly, activist, and artist voices to unfold their approaches to the uncertainties of the big data archives that permeate our present cultural moment. In this article we make the argument for a critically attuned, multidisciplinary approach to the unknowns, errors, and vulnerabilities that pertain to big data environments. This approach, grounded in cultural theory, implies a sensibility to the experience of the signs, flows, and bodies marked by uncertainty that inhabit and define contemporary big data archives. That is an experience which requires us to revisit

the archives that form the basis of our own scholarly work and to engage with other voices and forms of knowledge production, as we move forward in understanding and living among the archives produced by big data in our current era of surveillance.

Acknowledgments

We thank the Danish Research Council for the research grant that enabled the Uncertain Archives project, as well as Naja le Fevre Grundtmann for her contribution to the editing of this article and Merl Storr for her meticulous proofreading.

References

- Agostinho, Daniela. 2016. Big Data, Time, and the Archive. *Symploke* 24 (1–2): 435–45.
- Agostinho, Daniela, and Nanna Thylstrup. Forthcoming 2019. “If Truth Was a Woman”: Leaky Infrastructures and the Gender Politics of Truth-Telling. *Ephemera: Theory & Politics in Organization* 19 (2).
- Amaro, Ramon. 2016. Paper presented at Uncertain Archives Workshop on Error, Copenhagen, November.
- Amoore, Louise, and Rita Raley. 2017. Securing with Algorithms: Knowledge, Decision, Sovereignty. *Security Dialogue* 48 (1): 3–10.
- Anderson, Chris. 2008. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*, June 23, 2008. <https://www.wired.com/2008/06/pb-theory/> [accessed July 6, 2018].
- Andrejevic, Mark. 2014. Big Data, Big Questions: The Big Data Divide. *International Journal of Communication* 8 (1): 1673–89.
- Appadurai, Arjun. 2012. The Spirit of Calculation. *Cambridge Journal of Anthropology* 30 (1): 3–17.
- Bates, David W. 2002. *Enlightenment Aberrations: Error and Revolution in France*. Ithaca, NY: Cornell University Press.
- Beck, Ulrich. 2009. World Risk Society and Manufactured Uncertainties. *Iris* 1 (2): 291–99.
- Belle, La Vaughn. n.d. *La Vaughn Belle: Visual Artist*. <http://www.lavaughnbelle.com/info/> [accessed July 7, 2018].
- Borgen, Maibritt, Nanna Thylstrup, and Kristin Veel. 2016. Introduction. *Nordisk Tidsskrift for Informationsvidenskab og Kulturformidling* 5 (1): 3–8.
- Bowker, Geoffrey, and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Breton, André. (1924) 1969. Manifesto of Surrealism. In *Manifestoes of Surrealism*, 1–48. Ann Arbor, MI: University of Michigan Press.
- Browne, Simone. 2015. *Dark Matters: On the Surveillance of Blackness*. Durham, NC: Duke University Press.
- Buolamwini, Joy. 2017. *Algorithmic Justice League*. <https://www.ajlunited.org> [accessed July 7, 2018].
- Butler, Judith. 1997. *The Psychic Life of Power*. Palo Alto, CA: Stanford University Press.
- Caplan-Bricker, Nora. 2018. Preservation Acts: Toward an Ethical Archive of the Web. *Harper's Magazine*, December, 2018. <https://harpers.org/archive/2018/12/preservation-acts-archiving-twitter-social-media-movements/> [accessed January 8, 2019].
- Caswell, Michelle. 2014. *Archiving the Unspeakable: Silence, Memory, and the Photographic Record in Cambodia*. Madison, WI: University of Wisconsin Press.
- Caswell, Michelle. 2016. “The Archive” Is Not an Archives: On Acknowledging the Intellectual Contributions of Archival Studies. *Reconstruction* 16 (1). <https://escholarship.org/uc/item/7bn4v1fk> [accessed January 8, 2019].
- Caswell, Michelle, and Marika Cifor. 2016. From Human Rights to Feminist Ethics: Radical Empathy in the Archives. *Archivaria* 81: 23–53.
- Caswell, Michelle, Ricardo Punzalan, and T-Kay Sangwand. 2017. Critical Archival Studies: An Introduction. *Journal of Critical Library and Information Studies* 1 (2). <http://libraryjuicepress.com/journals/index.php/jclis/article/view/50/30> [accessed January 8, 2019].
- Chun, Wendy Hui Kyong. 2016. *Updating to Remain the Same: Habitual New Media*. Cambridge, MA: MIT Press.
- Chun, Wendy Hui Kyong. 2018. Queering Homophily. In *Pattern Discrimination*, edited by Clemens Apprich, Wendy Hui Kyong Chun, Florian Cramer, and Hito Steyerl, 59–97. Minneapolis, MN: University of Minnesota Press.
- Cifor, Marika. 2015. Presence, Absence, and Victoria's Hair: Affect and Embodiment in Trans Archives. *TSQ: Transgender Studies Quarterly* 2 (4): 645–649.
- Cifor, Marika, and Stacy Wood. 2017. Critical Feminism in the Archives. *Journal of Critical Library and Information Studies* 1 (2). <http://libraryjuicepress.com/journals/index.php/jclis/article/view/27/26> [accessed July 14, 2018].
- Copeland, Huey, and Krista Thompson. 2011. Perpetual Returns: New World Slavery and the Matter of the Visual. *Representations* 113 (1): 1–15.
- Cvetkovich, Ann. 2003. *An Archive of Feeling: Trauma, Sexuality, and Lesbian Public Cultures*. Durham, NC: Duke University Press.
- Deleuze, Gilles, and Félix Guattari. 2013. *A Thousand Plateaus: Capitalism and Schizophrenia*. New York: Bloomsbury.
- Derrida, Jacques. 1995. Archive Fever. *Diacritics* 25 (2): 9–63.
- D'Ignazio, Catherine. 2014. *Terra Incognita*. <http://www.kanarinka.com/project/terra-incognita/> [accessed July 6, 2018].
- D'Ignazio, Catherine. Forthcoming 2020. Outlier. In *Uncertain Archives*, edited by Nanna Thylstrup, Daniela Agostinho, Annie Ring, Catherine D'Ignazio, and Kristin Veel. Cambridge, MA: MIT Press.

- Dinkins, Stephanie. 2017. *Project al Khwarizmi*. <http://www.stephaniedinkins.com/project-al-khwarizmi.html> [accessed July 7, 2018].
- Dirckinck-Holmfeld, Katrine. 2011. What Is Artistic Research? Notes for a Lip-Synch Performance. In *Investigação em Arte e Design: Fendas no Método e na Criação*, edited by João Quaresma, Fernando Paulo Rosa Dias, Juan Carlos Ramos Guadix, 159–69. Lisbon: Centro de Investigação e de Estudos em Belas-Artes, Universidade de Lisboa.
- Dirckinck-Holmfeld, Katrine. 2017. The Christmas Report & Other Fragments. *Kopenhagen*. <http://kopenhagen.dk/artguide/single/article/the-christmas-report-other-fragments/> [accessed July 6, 2018].
- Edenheim, Sara. 2013. Lost and Never Found: The Queer Archive of Feelings and Its Historical Propriety. *Differences: A Journal of Feminist Cultural Studies* 24 (3): 36–62.
- Edwards, Elisabeth. 2016. The Colonial Archival Imaginaire at Home. *Social Anthropology* 24 (1): 52–66.
- Foucault, Michel. 1995. *Discipline and Punish: The Birth of the Prison*. New York: Vintage.
- Foucault, Michel. 2007. *The Archaeology of Knowledge*. London: Routledge.
- Freud, Sigmund. (1912) 1956–1974. Recommendations to Physicians Practicing Psycho-Analysis. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud Vol. 12*, 109–20. London: Hogarth Press.
- Gender Shades. 2017. *Algorithmic Bias Persists*. <https://www.media.mit.edu/projects/gender-shades/overview/> [accessed July 14, 2018].
- Gikandi, Simon. 2015. Rethinking the Archive of Enslavement. *Early American Literature* 50 (1): 81–102.
- Halberstam, Jack. 2005. *In a Queer Time and Space: Transgender Bodies, Subcultural Lives*. New York: NYU Press.
- Halberstam, Jack. 2011. *The Queer Art of Failure*. Durham, NC: Duke University Press.
- Halpern, Orit, and Gökçe Günel. 2017. FCJ-215 Demoing unto Death: Smart Cities, Environment, and Preemptive Hope. *Fibreculture Journal* 22 (1): 1–23.
- Hartman, Saidiya. 2008. Venus in Two Acts. *Small Axe: A Journal of Criticism* 11 (2): 1–14.
- Harvard Business Review. 2011. The Failure Issue. Special issue, *Harvard Business Review*, April, 2011. <https://hbr.org/archive-toc/BR1104> [accessed July 13, 2018].
- Hayles, N. Katherine. 2017. *Unthought: The Power of the Cognitive Nonconscious*. Chicago: University of Chicago Press.
- Helton, Laura, Justin Leroy, Max A. Mishler, Samantha Seely, and Shauna Sweeney, eds. 2015. The Question of Recovery: Slavery, Freedom and the Archive. Special issue, *Social Text* 33 (4).
- Hoffmann, Anna Lauren. 2019. The Privacy Risks of Unchecked Facial-Recognition Technology. *The Seattle Times*, March 22, 2019. <https://www.seattletimes.com/opinion/the-privacy-risks-of-unchecked-facial-recognition-technology/> [accessed June 3, 2019].
- Keeling, Kara. 2014. Electric Feel: Transduction, Errantry and the Refrain. *Cultural Studies* 28 (1): 49–83.
- Keyes, Os. Forthcoming 2020. (Mis-)Gendering. Schneider, Rebecca. In *Uncertain Archives*, edited by Nanna Thylstrup, Daniela Agostinho, Annie Ring, Catherine D’Ignazio, and Kristin Veel. Cambridge, MA: MIT Press.
- Klein, Melanie. 1931. A Contribution to the Theory of Intellectual Inhibition. *International Journal of Psycho-Analysis* 12 (1): 206–18.
- Kurgan, Laura. 2015–2017. *Conflict Urbanism: Aleppo*. <http://c4sr.columbia.edu/conflict-urbanism-aleppo/> [accessed July 6, 2018].
- Kwan, Mei-Po. 2016. Algorithmic Geographies: Big Data, Algorithmic Uncertainty, and the Production of Geographic Knowledge. *Annals of the American Association of Geographers* 106 (2): 274–282.
- Kwan, Mei-Po. 2002. Introduction: Feminist Geography and GIS. *Gender, Place and Culture* 9 (3): 261–62.
- Luksch, Manu, Martin Reinhart, and Thomas Tode, dirs. 2015. *Dreams Rewired*. DVD. New York: Icarus Films.
- Lyon, David. Forthcoming 2020. Sorting. In *Uncertain Archives*, edited by Nanna Thylstrup, Daniela Agostinho, Annie Ring, Catherine D’Ignazio, and Kristin Veel. Cambridge, MA: MIT Press.
- Manoff, Marlene. 2004. Theories of the Archive from Across the Disciplines. *Portal: Libraries and the Academy* 4 (1): 9–25.
- Menkman, Rosa. 2010–2012. *Collapse of PAL*. <http://rosa-menkman.blogspot.com> [accessed July 14, 2018].
- Menkman, Rosa. 2011. *The Glitch Moment(um)*. Amsterdam: Institute of Network Cultures.
- Moi, Toril. 1989. Patriarchal Thought and the Drive for Knowledge. In *Between Feminism and Psychoanalysis*, edited by Teresa Brennan, 189–205. London: Routledge.
- Murakami Wood, David. 2015. Vanishing Surveillance: Ghost Hunting in the Ubiquitous Surveillance Society. In *Invisibility Studies: Surveillance, Transparency and the Hidden in Contemporary Culture*, edited by Henriette Steiner and Kristin Veel, 281–301. New York: Peter Lang.
- Navarro, Tami. 2018. Beyond the Fragments of Global Wealth. *Periscope*. https://socialtextjournal.org/periscope_article/beyond-the-fragments-of-global-wealth/ [accessed January 8, 2019].
- Noble, Safiya. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- Nunes, Mark. 2011. Error, Noise and Potential: The Outside of Purpose. In *Error: Glitch, Noise and Jam in New Media Cultures*, edited by Mark Nunes, 3–23. London: Continuum.
- O’Malley, Pat. 2004. *Risk, Uncertainty and Government*. London: Taylor and Francis.
- O’Malley, Pat. 2009. “Uncertainty Makes Us Free”: Liberalism, Risk and Individual Security. *Behemoth: A Journal on Civilization* 2 (3): 24–38.
- Onuoha, Mimi. 2016. *The Library of Missing Datasets*. <http://mimionuoha.com/the-library-of-missing-datasets/> [accessed July 6, 2018].
- Ørum, Kristoffer. 2017. Ambiguous Physiognomy. *Digital Creativity* 28 (4): 308–317.
- Parisi, Luciana. 2013. *Contagious Architecture: Computation, Aesthetics, and Space*. Cambridge, MA: MIT Press.
- Ring, Annie. 2014. The (W)hole in the Archive. *Paragraph* 37 (1): 387–402.

- Ring, Annie. 2016. System Error: Complicity with Surveillance in Contemporary Workplace Documentaries. *Seminar: A Journal of Germanic Studies* 52 (4): 465–489.
- Roberts, Sarah T. 2018. Digital Detritus: “Error” and the Logic of Opacity in Social Media Content Moderation. *First Monday* 23 (3). <http://firstmonday.org/ojs/index.php/fm/article/view/8283/6649> [accessed July 6, 2018].
- Rockwell, Geoffrey, and Bettina Berendt. 2017. Information Wants to Be Free, Or Does It? The Ethics of Datafication. *Electronic Book Review*, December 3, 2017. <http://www.electronicbookreview.com/thread/technocapitalism/datafiction> [accessed July 6, 2018].
- Rosenberg, Daniel. 2014. Stop, Words. *Representations* 127 (1): 83–92.
- Schneider, Rebecca. Forthcoming in 2020. Glitch. In *Uncertain Archives*, edited by Nanna Thylstrup, Daniela Agostinho, Annie Ring, Catherine D’Ignazio, and Kristin Veel. Cambridge, MA: MIT Press.
- Schneider, Rebecca. 2011. *Performing Remains: Art and War in Times of Theatrical Reenactment*. London: Routledge.
- Schneider, Rebecca. Forthcoming 2020. Glitch. In *Uncertain Archives*, edited by Nanna Thylstrup, Daniela Agostinho, Annie Ring, Catherine D’Ignazio, and Kristin Veel. Cambridge, MA: MIT Press.
- Schram, Brian, and Jennifer R. Whitson. 2018. Why a Humanist Ethics of Datafication Can’t Survive a Posthuman World. *Electronic Book Review*, July 1, 2018. <http://www.electronicbookreview.com/thread/internetnation/datafication> [accessed July 6, 2018].
- Schüll, Natasha Dow. 2016. Sensor Technology and the Time-Series Self. *Continent* 5 (1): 24–29.
- Scourti, Erica. 2014 *The Outage: Her Story*. London: Banner Repeater.
- Spieker, Sven. 2008. *The Big Archive: Art from the Bureaucracy*. Cambridge, MA: MIT Press.
- Spillius, Elizabeth B., Jane Milton, Penelope Garvey, Cyril Couve, and Deborah Steiner. 2011. Epistemophilia. In *The New Dictionary of Kleinian Thought*, edited by Elizabeth B. Spillius, Jane Milton, Penelope Garvey, Cyril Couve, and Deborah Steiner, 323–25. London: Routledge.
- Steedman, Carolyn. 2001. Something She Called a Fever: Michelet, Derrida, and Dust. *American Historical Review* 106 (4): 1159–80.
- Steiner, Henriette, and Kristin Veel. 2018. Towering Invisibilities: A Cultural Theoretical Reading of the Eiffel Tower and the One World Trade Center. *Qualitative Inquiry* 25 (4): 407–16.
- Stoler, Ann Laura. 2002. *Carnal Knowledge and Imperial Power: Race and the Intimate in Colonial Rule*. Berkeley, CA: University of California Press.
- Stone, Amy L., and Jaime Cantrell, eds. 2016. *Out of the Closet, Into the Archives: Researching Sexual Histories*. Albany, NY: State University of New York Press.
- Sutherland, Tonia. 2017a. Archival Amnesty: In Search of Black American Transitional and Restorative Justice. *Journal of Critical Library and Information Studies* 2. <http://libraryjuicepress.com/journals/index.php/jclis/article/view/42/27> [accessed January 8, 2019].
- Sutherland, Tonia. 2017b. Making a Killing: On Race, Ritual, and (Re)Membering in Digital Culture. *Preservation, Digital Technology and Culture* 46 (1): 32–40.
- Taş, Birkan. Forthcoming 2020. VulnerAbility. In *Uncertain Archives*, edited by Nanna Thylstrup, Daniela Agostinho, Annie Ring, Catherine D’Ignazio, and Kristin Veel. Cambridge, MA: MIT Press.
- Taylor, Diana. 2003. *The Archive and the Repertoire: Performing Cultural Memory in the Americas*. Durham, NC: Duke University Press.
- Thylstrup, Nanna Bonde. 2018. *The Politics of Mass Digitization*. Cambridge, MA: MIT Press.
- Veel, Kristin. 2018. Make Data Sing: The Automation of Storytelling. *Big Data and Society* 5 (1).
- Whearty, Bridget. 2018. Invisible in “The Archive”: Librarians, Archivists, and The CaswellTest. Paper presented at 53rd International Congress of Medieval Studies, Kalamazoo, MI, May, 2018.
- Zer-Aviv, Mushon. 2013. *The Turing Normalizing Machine: An Experiment in Machine Learning & Algorithmic Prejudice*. <http://mushon.com/tnm/> [accessed July 7, 2018].