Big Data Biopolitics

Computing Racialised Assemblages in Terrorist Watchlist Matching

Gary Kafer

Abstract

This article considers the medial logics of American terrorist watchlist screening in order to study the ways in which digital inequities result from specific computational parameters. Central in its analysis is Secure Flight, an automated prescreening program run by the Transportation Security Administration (TSA) that identifies low- and high-risk airline passengers through name-matching algorithms. Considering Secure Flight through the framework of biopolitics, this article examines how passenger information is aggregated, assessed and scored in order to construct racialised assemblages of passengers that reify discourses of American exceptionalism. Racialisation here is neither a consequence of big data nor a motivating force behind the production of risk-assessment programs. Both positions would maintain that discrimination is simply an effect of an information management system that considers privacy as its ultimate goal, which is easily mitigated with more accurate algorithms. Not simply emerging as an effect of discriminatory practices at airport security, racialisation formats the specific techniques embedded in terrorist watchlist matching, in particular the strategies used to transliterate names across different script systems. I argue thus that the biopolitical production of racialised assemblages forms the ground zero of Secure Flight's computational parameters, as well as its claims to accuracy. This article concludes by proposing a move away from the call to solve digital inequities with more precise algorithms in order to carefully interrogate the forms of power complicit in the production and use of big data analytics.

Introduction

It is all too common now in our contemporary age of big data to hear warning of the ever-expanding range of monitoring, tracking and classifying programs that pervade our daily lives. Alongside media spectacles like the Snowden leak in 2013 and the Facebook-Cambridge Analytica breach in 2018, we have witnessed a resurgence of debates about the end of privacy and the inevitable death of democracy by the hand of algorithmic proxy politics. Today's surveillance society is one in which

DOI 10.14361/dcs-2019-0103

public and private activity can be quantified for inclusion in federal and commercial databases, moulded into various identity profiles, and circulated among data brokers for financial gain. At the centre of these debates is an untenable anxiety around the dissolution of the subject into vast arrays of quantifiable profiles circulating federal and commercial databases that aim to classify individuals and predict their future behaviour. While claiming accuracy, such profiles are only discrete approximations of membership within populations of data, "epistemologically fabricated" within algorithmically mediated versions of reality that authorise a range of governmental actions (Cheney-Lippold 2017: 45). Accuracy here is often championed as a solution to issues of privacy rights and access that continually plague the imbrication of computational systems within our lived reality.

However, while certainly everyone's data are continually made public in varying degrees, the impact of this transparency is not equally distributed across the social body. One of the central premises of big data that is often overlooked in debates on privacy and transparency is how most people are not targeted as individuals, but as aggregates. Because big data deprioritises the content of data relative to its structure, the significance that data accrues is in its connections, patterns and networked potential, not merely its representational meaning. A datafied "identity" is only ever defined through correlations with larger populations of data, correlations often resulting not from causal relations but use of proxies to determine class membership. And because proxies (like zip code or consumer preferences) encode categories of social difference, people of colour, migrants, sexual minorities, the poor and other oppressed populations are overwhelmingly more likely to bear the burden of population-based classification (Chun 2016: 58). Indeed, surveillance is not a totalising force suppressing the social body, but rather a variegated power matrix that is distributed along axes of social difference under signs of empire, settler colonialism and white supremacy.

This article extends debates on the production and circulation of algorithmically mediated identities by examining the ways in which digital inequities structure the specific design and operation of surveillance systems. My primary site of interest is Secure Flight, an automated prescreening program that identifies low- and high-risk passengers by matching their names against the no fly list and the selectee list. Both managed by the Terrorist Screening Center (TSC) and distributed to the Transportation Security Administration (TSA), the no fly list completely bans suspected persons from flying, while the Selectee List requires additional security checks at the airport before the passenger is allowed to board the aircraft. Initiated in 2009, Secure Flight replaced two previous versions of Computer Assisted Passenger Prescreening Systems (CAPPS I and CAPPS II), the latter of which was implemented following the attacks on 9/11 in order to place control of passenger screening in the hands of the government. By the end of 2010, Secure Flight met its goal of vetting one hundred percent of all domestic and international flights, reportedly prescreening on average two million passengers per day by 2012 (DHS 2012: 22).

Secure Flight and its attendant watchlists have been subject to harsh criticism, particularly for the ways in which they disproportionately and erroneously identify ethnic and religious minorities, immigrants and non-US citizens as perceived threats to the nation. However, while much can be said about the legal frameworks that attend Secure Flight and its unfair redress system, what I explore below are the specific medial logics that underpin discriminatory practices at airport security. By *medial logics*, I refer to the conceptual, discursive and rhetorical tendencies that structure mediations across historical and social formations. My use of this term is inflected by media materialism, an approach which considers technologies as active agents that structure how we know and experience the world (cf. Fuller 2005; Cubitt 2014; Parikka 2015). Importantly, media materialism understands the transmission and processing of cultural epistemologies not simply as medial effects, but rather as coming into being as matter within processes of mediation. Medial logics articulate how the mattering of mediation is formatted by specific discursive and conceptual expressions within historically situated contexts.

In the case of Secure Flight, medial logics flag how racial epistemologies structure inequities within the unique and productive formatting of computational expression subtending the automated name-matching system. I take cue here from Tara McPherson (2018), who insightfully examines how the development of UNIX in mid-century America employed a logic of modularity that encoded the modular logics of race simultaneously emerging in neoliberal projects of urban planning, managerial discourse, and university specialisation. Key to her argument is the claim that there is a sociocultural history to the mathematical and physical basis upon which computational systems are designed, that design itself is never without politics. In her words, "race, particularly in the United States, … fundamentally shap[es] how we see and know as well as the technologies that underwrite or cement both vision and knowledge" (ibid: 50). Attending to the medial logics of a particular system means that we need not bracket identity politics as pertaining only to cultural context. Rather, we must acknowledge how contemporary computational systems ground claims to accuracy by foreclosing an understanding of how technological formations are deeply bound up with logics of race.

In order to unearth the medial logics of Secure Flight, I turn to a range of federal reports conducted by the Department of Homeland Security, the TSA, the Government Accountability Office, and other Congressional committees. Many of these reports are openly accessibly by law. Others are declassified materials only made public after security measures have been taken to redact sensitive material. Due to limitations of confidentiality, such documents offer a narrow glimpse into the computational parameters employed in the name-matching process. Yet where they lack in substantial evidentiary power, they instead lay bare the discursive strategies that link the promise of big data analysis to the biopolitical project of national security.

In my assessment of digital inequities, I locate biopolitics as the production of racialised assemblages of low- and high-risk populations: those who are presumed capable of participating in discourses of American exceptionalism and those who are deemed threatening to national security. Racialisation here is neither a consequence of big data nor a motivating force behind the production of risk-assessment programs. Both positions would maintain that discrimination is simply an effect of an information management system that considers privacy as its ultimate goal, which is easily mitigated with more accurate algorithms. Rather, processes of racialisation format the specific computational techniques embedded in terrorist watchlist screening, in particular the transliteration strategies used to represent names across different script systems. As such, I argue that the biopolitical production of racialised assemblages forms the ground zero of Secure Flight's medial logics, as well as its claims to accuracy. Situating the computational parameters of terrorist screening within biopolitical discourse not only reveals how digital inequalities are structurally formatted within algorithmic mediation, but also how big data takes as its basis a project of racialisation in its attempt to quantify and manage populations within the scope of a white and gendered American exceptionalism.

The Biopolitics of Terrorist Watchlisting

Every time an individual makes an airline reservation, arrives at a US port of entry, submits an application for a visa or is stopped by state or local police, the frontline screening agent initiates a name-based search of the individual against applicable watchlists (Fig. 1). The dedicated program within the TSA that assumes the function of watchlist matching for civil aviation is Secure Flight. As outlined in the Intelligence Reform and Terrorism Prevention Act (108th Congress 2004), Secure Flight is an automated prescreening program that allows the TSA to assume from aircraft carriers the role of checking passenger information against the watchlists generated from the Terrorist Screening Database (TSDB) (Fig. 2). Airlines, as well as third-party entities such as travel agencies and booking websites, are required to submit to the TSA specific passenger information known as Secure Flight passenger data (SFPD), which include name, birthdate, gender, flight information and passport number (DHS 2008: 64021). Secure Flight cross-checks SFPD with the no fly and selectee lists, as well as searches in the Terrorist Identities Datamart Environment (TIDE) database and the Treasury Enforcement Communications System to discern whether or not there is a positive match (DHS 2012: 24). Notifications are then sent back to aircraft operators in the form of a boarding pass printing result with one of three possible outcomes: the passenger is cleared to fly, selected for additional screening or prohibited from boarding the airplane.[1]

[1] Prior to the 9/11 attacks, risk-based passenger screening did not exist as an automated computational system. In past decades, terrorist watchlists were composed of a heterogeneous assemblage of media usually involving human oversight. The

Fig. 1: The general set of procedures for matching individuals to terrorist watchlists used across federal and local agencies. U.S. Government Accountability Office (2017): "Terrorist Watch List Screening," GAO-08-194T, p. 5.

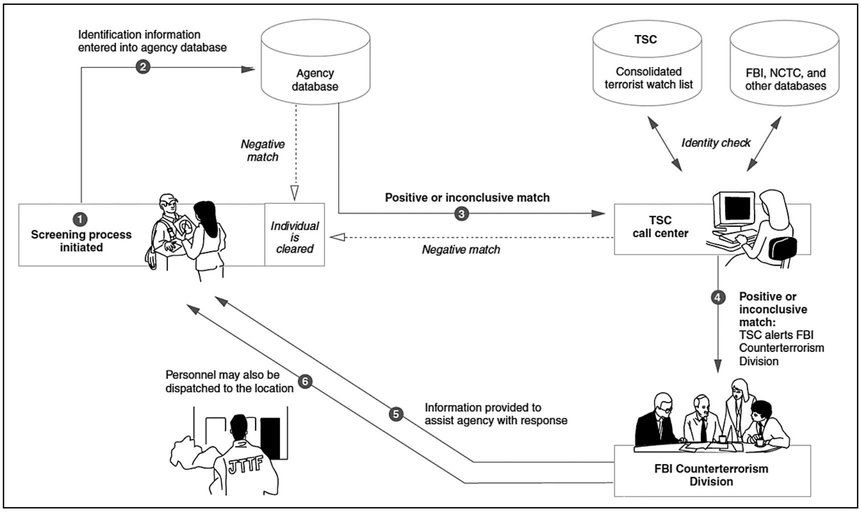
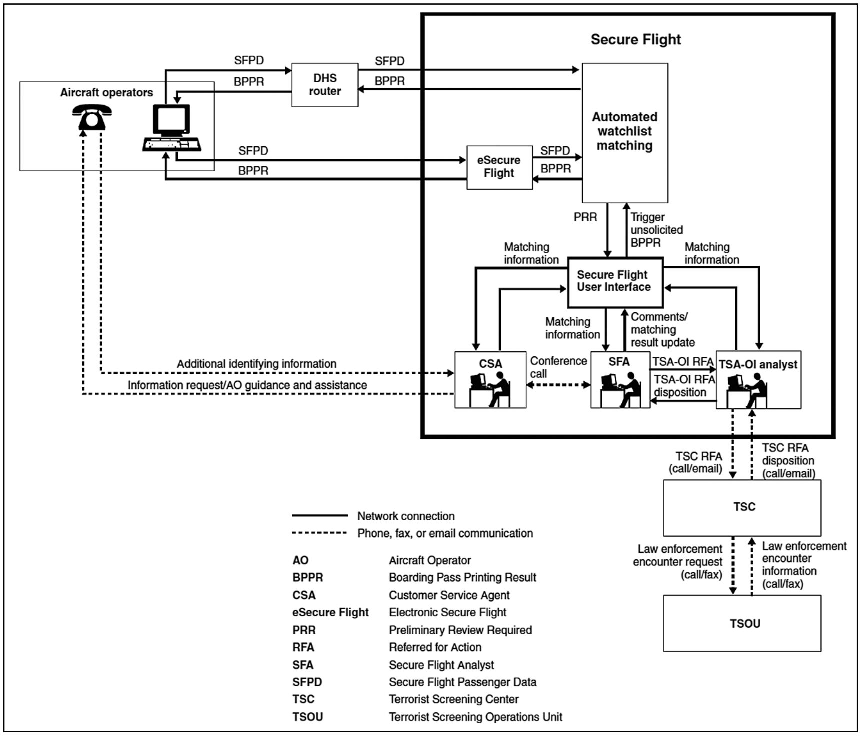

Identification information entered into agency database
2
Agency database
TSC
Consolidated terrorist watch list
FBI, NCTC, and other databases
Negative match
Identity check
Positive or inconclusive match
3
1
Screening process initiated
Individual is cleared
TSC call center
Negative match
4
Positive or inconclusive match: TSC alerts FBI Counterterrorism Division
Personnel may also be dispatched to the location
6
5
Information provided to assist agency with response
JTTF
FBI Counterterrorism Division

Fig. 2: Flow of operations in the Secure Flight watchlist matching process. U.S. Government Accountability Office (2009): "Aviation Security," GAO-09-292, p. 7.


Secure Flight
SFPD
BPPR
DHS router
SFPD
BPPR
Aircraft operators
Automated watchlist matching
SFPD
BPPR
eSecure Flight
SFPD
BPPR
PRR
Trigger unsolicited BPPR
Matching information
Matching information
Secure Flight User Interface
Matching information
Comments/ matching result update
Additional identifying information
Information request/AO guidance and assistance
CSA
Conference call
SFA
TSA-OI RFA
TSA-OI RFA disposition
TSA-OI analyst
TSC RFA (call/email)
TSC RFA disposition (call/email)
TSC
Law enforcement encounter request (call/fax)
Law enforcement encounter information (call/fax)
TSOU
Network connection
Phone, fax, or email communication
AO Aircraft Operator
BPPR Boarding Pass Printing Result
CSA Customer Service Agent
eSecure Flight Electronic Secure Flight
PRR Preliminary Review Required
RFA Referred for Action
SFA Secure Flight Analyst
SFPD Secure Flight Passenger Data
TSC Terrorist Screening Center
TSOU Terrorist Screening Operations Unit

When announced by the TSA in 2004, the automated prescreening program was said to represent "a significant step in securing domestic air travel and safeguarding terrorism related national security information" (TSA 2004: 57343). Alongside other technologies of border control that regulate the mobility of populations, Secure Flight participates as a key site for the biopolitical management of national security standards. This network of security practices, which Didier Bigo (2008) terms the "ban-opticon," articulates a biopolitical terrain that wields the normative imperatives of national exceptionalism in order to intern certain groups while granting free movement to unmarked populations rendered legible to surveillant systems. Famously defined by Foucault in his 1976 lectures "Society Must Be Defended," biopower "consists in making live and letting die" (Foucault 2003: 247), a form of control that modulates populations according to their relative capacity to propagate life or their risk of death. Here, the quantified individual is measured against a range of normative assessments like vitality, ability and morbidity, producing categories of risk that are then distributed on the population level. In aviation security, risk is most often assessed through the interaction of three factors: perceived threat posed by individuals, vulnerabilities in layers of security protocol and the potential consequences of security tactics (DHS 2011: 20). Technologies of border control, like Secure Flight and its attendant watchlists, mobilise risk-assessment strategies in order to contain bodies perceived as dangerous for the health of the national body. The primary function of terrorist screening programs

first federal government watchlist devoted to identifying suspected terrorists was TIPOFF, created in 1987 and managed by the state department. A deck of three-by-five index cards that were kept in a shoebox, TIPOFF consisted of names, birthdates, nationalities and passport numbers for foreigners suspected of terrorism who should not receive visas for entry across the US border (Kahn 2013: 10). The list was eventually computerised in 1993, but still distributed on paper to airlines around the time of the first World Trade Center attack in that year. In the early 1990s to 2001, the Federal Aviation Administration assumed the responsibility of terrorism prevention through security directives, which were issued to airlines in order to deny boarding to targeted individuals. While the current orthodox for inclusion in the TSDB operates upon an edict of "reasonable suspicion," the standard for issuing security directives required "specific credible threats" to civil aviation (Department of Transportation 1991: 4322). The FAA thus required clear evidence that a threat was immanent to a particular aircraft or carrier in order to bar travel. At the time, there was no automated system for cross-referencing names issued by the FAA. Security directives were distributed on paper by fax machine to aircraft operators and manually checked for any positive hits (Kahn 2013: 133). On average, about 20 to 30 security directives were issued per year, and among these only a handful of people were named. Days before 9/11, FAA security directives prohibited only twelve named individuals from boarding commercial aircrafts (TIPOFF at the time had about 60,000 entries of those who would be denied visas for entry into the United States).

is to demarcate boundaries of the abnormal, perverse and sexually deviant Other, which in turn normalises those bodies perceived as innocent through discourses of American exceptionalism (Puar 2007: 38).

However, in the preemptive politics of security programs like those deployed in terrorist watchlists, populations are not always available as legible, stable or measurable entities. Terrorism rather emerges as an unspecified threat within some undefined futural scenario. Indeed, the TSC, which houses the TSDB, publicly defines a "suspected terrorist" on their website as "an individual who is reasonably suspected to be engaging in, has engaged in, or intends to engage in conduct constituting, in preparation for, in aid of, or related to terrorism and/or terrorist activities" (Federal Bureau of Investigations 2017: 3).[2] As the "unknown unknown," Brian Massumi diagnoses terrorism within an accelerating temporality of crisis, "an ever-presence of indiscriminate threat, riddled with the anywhere-anytime potential for the proliferation of the abnormal" (Massumi 2009: 157). In this case, the distribution of risk upon suspect populations is mobile, contingent and unfolding. Foucault (2007) makes evident here the way in which risk is differentially applied to suspicious populations in the security state: risk is "not the same for all individuals, all ages, or in every condition, place or milieu," but instead is differential, revealing "zones of higher risk and, on the other hand, zones of less or lower risk" (89). In the contemporary political moment, biopower renders bodies into "data derivatives," wherein risk is preemptively distributed in order to justify the exclusion of populations from the national political community (Amoore 2011). Security apparatuses, like the TSA's risk-based passenger screening strategies, do not simply locate stable identity categories in the social field, but rather produce non-normative bodies through differential matrices of risk that become activated within everyday sites of surveillance: low-threat passengers who have access to TSA pre-check and high-risk passengers who are included on no fly lists.

In order to demarcate risk differentials of data aggregates, Secure Flight's pattern analysis system mobilises logics of racialisation. Biopolitics crucially configures populations not simply as terrains of governance, but as what Alexander Weheliye (2014) terms "racialised assemblages" which demarcate those who may enjoy the status of human and those who are consigned to the position of not-quite-human or non-human (43). Racialisation does not locate fixed elements of phenotypic, biological or cultural difference. Rather, it names the zero degree of biopolitical control, a process of imposing classificatory regimes upon populations in order to produce categories of human difference. In Weheliye's words,

[2] Under the awning of reasonable suspicion, the TSC can place persons in the TSDB based simply on intel that suggests one might pose an undefined threat in the future. What has result is a massive influx of entries seemingly absent of discretion. From 2009 to 2013, approximately 1.6 million individuals were designated for inclusion in the database, while only about 1 percent of nominations (just over 14,000) were rejected (Bjelopera, Elias, & Sisken 2016: 6).

racialisation drives "a set of sociopolitical processes of differentiation and hierarchisation, which are projected onto the putatively biological human body" (ibid: 5). In the production of racialised assemblages, political subjugation is masked by relations of biological substance. At work is a generalised calculus of risk assessment that recruits biology in order to justify the eviction of populations from the political community. As a biopolitical strategy, racialisation finds its end in techniques of capture, incarceration and deportation by appealing to perceived differences of species classifications at the sub-human level (Mbembe 2017: 35). One need not look further than Donald Trump's comments from May 2018, in which he referred to Mexican immigrants as "animals," in order to witness the way contemporary security apparatuses racialise populations in order to justify state violence.

Within the post-9/11 era, normative mechanisms of airport security practices operate through the biopolitical production of racialised assemblages. The goal of Secure Flight alongside other tracking, monitoring and scanning technologies is to determine who is secure for travel and those who are "irredeemably opaque" to transparency efforts (Hall 2015: 79). Throughout the intensely consolidated network of airport security, religious and ethnic minorities who carry "racial baggage" (Browne 2015: 154), specifically American Muslims, Arabs and South Asian communities, are subject to a litany of profiling practices based on skin colour, clothing, behaviour, gender, age, class, ability and much more. Names in particular, circulating on passports, boarding tickets, passenger manifests and terrorist watchlists, locate a particular racial difference that works to situate bodies adjacent to the figure of the Muslim terrorist. In these cases, pre-existing interpretive schemas of race congeal around bodies caught within an economy of suspicion circulating the highly charged scene of surveillance, which often results in excessive monitoring and potential detention of South Asian, Muslim and Arab passengers (Selod 2018).

Secure Flight, however, is unique among other technologies of airport security, such as full-body x-rays, biometric scanners and passports, because the program operates *before* passengers arrive to the airport. Like no fly lists, Secure Flight maneuvers a logic of preemption that militarises national and global big data infrastructures. As Kenneth Werbin (2009) writes of no fly lists, "the logic is to preempt terrorist threats, and such security practices are largely represented as turning on [...] complex computer algorithms and technological networks [that] are understood as the critical security arrangement mitigating the inevitable future of terrorist threats" (621). The reliance of computational systems to aggregate and identify passenger profiles makes possible a set of strategies for security agencies like the TSA to absolve responsibility in misidentifying persons as suspected terrorists. In this case, claims of computational objectivity obfuscate the discriminatory basis of the list as a "foundational inclusionary and exclusionary political form through which the circulation of dangerous people is reified and policed" (ibid: 617). Werbin goes to lengths to demonstrate the way in which technological systems alone are not enough to account for the way fear locates the figure of the terrorist within actual bodies. As he states:

> normative mobility is not enacted through techno-scientific practices alone, for it is not the digital profile, computer algorithm, or risk-assessed score that ultimately marks, concretizes, and reifies the threat: it is the list. Where profiles can be understood as indicators of risk, it is ultimately the name on the list that calls the state of exception into being, operationalizing the ban and invoking the looming omnipresent camp. (ibid: 621)

At work here is a tension between the computational systems that assess passenger information for possible risk and the watchlists that produce terrorist threats within the traversed space of the airport.

In what follows, I extend to Werbin's methodical examination of no fly lists as they intersect normative aviation security practices in order to advance the ways in which we understand computational systems as imbricated within the biopolitical production of racialised assemblages. Certainly, we may observe how computational systems inherit sociopolitical relations of difference that then actualise in the citation of bodies as risky in airport security. Similarly drawing upon Weheliye in his study of big data analytics, Ezekiel Dixon-Román (2016) asserts that "regardless of the initial code of the algorithm, as it intra-acts with myriad persons and algorithms and analyses and learns from the data, the ontology of the algorithm becomes a racialised assemblage" (488). For Dixon-Román, algorithms draw upon hierarchies of social difference that then reify racial understandings when made operational in formats for disciplining human bodies, such as no fly lists. In his words, "the algorithm is not inherently racialised but becomes racialised through the analysis of data assemblages" (ibid: 489). I offer an opposing view of racialised assemblages in computational systems, arguing that the statistical analysis of risk always already operates upon logics of racialisation that are encoded into specific computational parameters. Algorithms do not become racialised when encountering data imbued with elements of sociopolitical difference, but rather *mobilise logics of racialisation in order to process data assemblages.*

As the TSA's dedicated name-matching program, Secure Flight is a crucial site to consider how logics of racialisation inhere within the specific medial logics underpinning big data computational analysis. One of the principle difficulties with data analysis in terrorist watchlist matching is what happens when data belonging to the TSDB is used to cross-reference no fly lists. In a report on the impact of automated selection on privacy and civil rights, the Department of Homeland Security (2006) conceded that even if the lists themselves were able to flawlessly index all alleged terrorists, the algorithms that are used to recall specific information would still produce certain risks of misidentification because "no matching technology works perfectly" (ii). While we might locate discrimination within the production of terrorist identities on no fly lists – as well as those "mistakes" of the database that are then included on lists – it can also be detected in the methods of computation that activate passenger data within risk-assessment algorithms. As such, it is crucial that we expand the biopolitics of terrorist screening beyond the citation of no fly lists within airport surveillance practices to the computational parameters through which names are identified as risks to national security.

Naming Risk

When computing matches between names in SFPD and no fly lists, Secure Flight makes use of a scoring threshold function to determine the relative accuracy of the potential match. This score is determined by three factors: the relative importance of each piece of information, such as name versus date of birth or flight number; the criteria used to specify how aspects of the SFPD might register as potential hits (which include, for example, the range of birth dates that the system would consider a match); and the numeric threshold over which the SFPD will determine a positive hit (GAO 2009: 8).[3] The use of a scoring threshold in the name-matching process indicates that Secure Flight employs probabilistic computational methods. As opposed to deterministic matching, which simply assesses unique identifiers for an exact comparison, probabilistic matching applies statistical analysis to establish relationships between data elements in order to account for errors, complex data structures and multiple databases. Outcomes of probabilistic computation are then assigned a percentage to indicate the probability of a successful match.

Within the context of terrorist watchlist screening, probabilistic computation is necessary to account for name variations in records across the databases to which Secure Flight has access. While TSDB is the primary resource for watchlist screening, Secure Flight also cross-references TIDE, a database maintained by the National Counterterrorism Center that serves as the primary source of non-US

[3] The scoring threshold adjudicates the relative yield of false-positive versus false-negative cases. Raising the scoring threshold would result in fewer matches on the watchlists, which would increase the possibility of false-positives. On the other hand, if the scoring threshold is lowered, name variants and birth date entries would be made more comparable, thus increasing the risk of false-positive matches (GAO 2009: 8). From the perspective of national security, false-negatives must absolutely be avoided, while false-positives are a small price to pay for preventing possible terrorist threats. Ultimately, however, the security trade-off must not compromise the operation of Secure Flight's automated program for national security efforts. As the U.S. Government Accountability Office (2006) reported, "any policy trade-off considerations regarding use of algorithms likely will favor ensuring homeland security over minimizing inconveniences to travelers" (43).

citizen passenger records in the TSDB (Elias 2014: 9). Not all names, however, can be accurately represented across different language systems, particularly when forced to conform to formatting procedure across database record-keeping structures. We see this in the way that non-English names are transliterated into conventional Western naming conventions and Latin script. For example, Latin American names may contain multiple surnames, while Chinese names place the surname at the beginning, which might also be complemented by a Western name. Arabic names in particular confound transliteration practices, as there is no single methodology in place for representation in Latin script. The name of Libyan revolutionary Muammar Gaddafi can be written more than one hundred ways in Latin text, including Muammar Qaddafi, Moamar Gaddafi, Mouammar Kadhafi and Mu'ammar al-Qadafi. As many computer scientists have shown, name-matching algorithms, particularly of the phonetic class,[4] reach certain limitations when names entered into the database are represented in a script different from their native system, as there may be many alternatives for encoding a string of phonemes from one language into another (cf. Christen 2006; Freeman, Condon, & Ackerman 2006; El-Shishtawy 2013). A congressional report from 2009 revealed that while the TSDB at the time contained over 400,000 individual identities, the number of actual records within the database exceeded one million due to aliases and name variants from across languages (Krouse & Elias 2009: 3–4).

Perhaps not surprisingly, the TSDB is calibrated upon Western English naming conventions and Latin script, which means that name variation most often maps onto the incorporation of non-English names into the TSDB. Following the attacks on 9/11, the National Commission on Terrorist Attacks speculated in their official report of the events that the absence of a standardised transliteration practice was one of the key faults through which terrorist measures were able to manifest that day, as perhaps the attackers were able to obfuscate their identities by adopting an alternate spelling of their names. The NCTA (2004) recommended a universalised transliteration protocol as the way forward for preventing future crises: "While the gradual introduction of biometric identifiers will help, that process will take years, and a name match will always be useful. The ICAO [International Civil Aviation Organization] should discuss the adoption of a standard requiring a digital code of all names that need to be translated into the Roman alphabet, ensuring one common spelling for all countries" (565, note 40). Following the publication of this report, Congress encouraged the President's office to enter into international negotiations in order to institute universal transliteration standards

[4] Name-matching algorithms can be generally divided into two main categories: orthographic algorithms compare units and strings of values across items, while phonological algorithms relate entries with respect to phonetic representation. According to a report filed by the Department of Homeland Security (2006), multiple algorithms may be used in performing matches against terrorist watchlists, such that the differences that results from these two classes may be used to verify one another.

based on the Latin script and English semantics. Importantly, universal transliteration forecloses a range of many other computational methods for processing names rendered in non-Latin script. I take seriously Simone Browne's (2015) contention that "when particular surveillance technologies, in their development and design, leave out some subjects and communities for optimum usage, this leaves open the possibility of reproducing existing inequalities" (162–163). A modified N-gram algorithm, for example, has been shown to process more robustly the unique features of the Arabic language (cf. Al-Sanabani & Al-Hagree 2015; Alsurori, Al-Sanabani, & Al-Hagree 2018).

In the post-9/11 era, the NCTA advanced a model of name-matching that operated upon norms defined by the phonetic patterning of the English language. Non-English names – "ethnic names" or "foreign names" – were intentionally included within databases as *exceptions* to the rule, rather than as constitutive elements of its computational logic. As passenger screening technologies developed over the following decade, the United States continued to serve as the cultural-computational background for risk assessment in terrorist watchlist matching, as well as the foundation for a range of other systems for border control and national security. Analysing the medial logics of biometric passports, in which names, dates and numbers must be represented in Arabic numerals and Latin characters, Liv Hausken (2017) refers to this process as "cultural homogenization," where "the reduction in cultural diversity [is] the result of processes by which local cultures are transformed or absorbed by a dominant culture" (272–273). The globalisation of communication systems forces local standards to conform to international regulations, thus constructing an informational economy in which abnormal citizenship is systematically measured against a template provided by the US empire. In the case of terrorist watchlist matching, transliteration practices not only consolidate standards of efficiency and accuracy in the medial logics of Secure Flight, but also attempt to render legible human identities within discourses of American exceptionalism.

Racialising Data

The citation of "suspicious sounding" names has long been a casual marker of difference or threat in the American context (e.g. German-sounding names during WWII or Soviet names during the Cold War). Within the post-9/11 American cultural imaginary, names perceived as Muslim, Arab and South Asian overwhelmingly bear the burden of racialisation within discourses of terrorism, while other cultural or ethnic names become racialised in contexts like border security and policing. Secure Flight intervenes in the biopolitical distribution of risk by elevating the racialisation of names to the level of data science. Within TSA's automated prescreening program, names become racialised as data points that are assessed within pattern analysis algorithms to detect matches with no fly lists. It is thus imperative to understand the name-matching system as a function of racialisation within the biopolitical management of population mobility.

One might object by observing that English names too are susceptible to false-positive matches in the name-matching process. Infamous among such cases is Senator Ted Kennedy who was mistakenly identified on the no fly list in 2004 by CAPPS II due to the similarity of his name to an alias used by someone suspected of terrorist affiliations. This is reflected in official discourses on Secure Flight's design and infrastructure, which is careful to mitigate concern that categories of social difference, including race, gender and citizenship, might somehow be implicated within the program's computational design. In a notice outlining the process and aims of Secure Flight, the TSA reports that in addition to training the automated prescreening program on information from the TSC, it will also use data from commercial aggregators who provide services to banking, home mortgage and credit industries in order to determine if such data are able to identify incorrect passenger information. Among other legal assurances stipulating that proper erasure and security safeguards are employed, the TSA (2004) indicates that commercial data "would not result in inappropriate differences in treatment of any protected category of persons" (57343). In a 2012 progress report on the implementation of Secure Flight, the Department of Homeland Security further observes that Secure Flight Operations Center employees receive operational training to align with the impartial standards inscribed in the program policies, including training in cultural naming conventions and nominal gender signifiers.

Yet as the computational basis of Secure Flight's probabilistic matching system reveals, there is a greater margin of error for non-English names. The mobile norm of risk management is only mobile in relation to the relative stability of a gendered and sexualised whiteness as a cultural-computational basis for the pattern recognition system of name-matching. As a technique of biopower, statistical pattern analysis is immediately racialised in attempting to determine risk through the parsing of data assemblages with respect to larger populations of information. In network science, the identity of a data element is produced by correlations bounded by the differentiation of norm and anomaly. Of course, in the drive to amass more and more information, there is no longer a functional difference between targets and non-targets, as intelligence about both groups is needed in order to isolate their relevant differences (Andrejevic & Gates 2014: 190). The identification of risk thus becomes a matter of an *exception* from a data population that may be regarded as typical or unremarkable. Exception here does not refer to the content of data, but the inclusion of a value precisely through its necessary difference from an informational background considered as the norm.

The kinship between racialisation and statistical assessment is certainly far from new. The relation between the two concepts derives from the historical construction of risk as a classification system. The history of life insurance policy in the United States, particularly in the Reconstruction era through the early 1900s, serves as a key example of the way in which risk, race and debility became imbricated within a nationalist biopolitical agenda. As Michael Ralph (2012) demonstrates, the collapse of the slave insurance industry in the postbellum period did not necessitate the equal distribution of life insurance across racial lines. As conceptions of race shifted from chattel status to a formal property buttressed by a priori figurations of biological difference, the alleged expertise of medical examiners and the apparent statistical correlation between race and criminality offered by social scientists forged a crucible of objectivity that apprehended black populations as greater risks to the economic stability of the nation. As such, risk in actuarial science is grounded in what Dan Bouk (2015) terms a "white data politics" (185), wherein whiteness serves as the default neutral category for statistical models that are used to assess those populations deemed mobile, indeterminate and substandard. Racialisation emerges as a process of hierarchisation in which biological difference is displaced by a statistical racism rooted in the notions of black inferiority and pathology measured against the ascendance of other "foreign-born" immigrants, like the Irish, Slavic and Italian, to the category of whiteness (Muhammad 2010: 9).

Secure Flight exemplifies the ways in which a white data politics not only extends into the contemporary moment, but also how it undergoes certain transformations performed by algorithmic mediation. No longer a matter of trained judgment or expertise, computation shifts the epistemic foundation of objectivity towards a digital economy bracketed by the density and speed of information transmission (Halpern 2014). It would seem then that in the era of big data, racialisation is powered by the subsumption of biological or cultural difference within the capacity of networked relations. Writing on the use of "smart intelligence" social media platforms in US security agencies, Werbin (2011) asserts that the threatening Other is not a population marked by consistent biological, national or cultural formations, but emerges as "a series of amorphous discriminatory profiles derived through a series of patterns of interaction, and factored in a series of imagined security scenarios" (1263). In the context of Secure Flight, the racialisation of certain populations would hinge upon the proximity of assets within a database, the parsing of search elements, the calibration of the scoring threshold function and the management of probabilistic computation.

However, it is important that we do not completely remove biological, social and cultural formations from our understanding of automated national security systems. Rather, Secure Flight reveals a more complicated view of the relation between statistical analysis and the biopolitical logics of racialisation that continues the project of risk assessment from pre-digital eras. I argue that the sub-human and the quantified data profile work coterminously within Secure Flight's automated system as names are not simply abstracted patterns of data, but are naturalised as indices of racial difference. The inclusion of gender in SFPD alongside passenger name, passport number and redress information is an important component here, as it maps the statistical process of racialisation along categories of sexual difference. Not unlike the regulation of gender conformity in full-body scanner and passport documentation, assumptions of biological sex structure everyday state surveillance practices that attempt to position "a variety of bodies, behaviors, and identities – not only those explicitly identified as transgender – as gender-nonconforming" (Beauchamp 2019: 7). While categories of race might be absent from SFPD, binary sex categories serve to animate ontological imaginaries of biological difference. Resonating with Kyla Schuller's (2018) writing on the biopolitics of 19th-century sentimental science, Secure Flight sets in motion binary sex "to accomplish the work of racial differentiation" (17). Gender is rendered computational here as a variegation of racial biopower in order to assess the accuracy of probabilistic computation in the name-matching process.

Racial difference is naturalised as an a priori feature of names when the computational background of dominant cultural classifications recedes under the guise of accuracy. Consequently, "variance" only makes sense as an epistemic category when located as an attribute of the name itself, rather than as a penalty of the particular pattern analysis that makes use of gender when assessing the result of probabilistic matching. Name variation is a marker of threat within US national security, which is most clear in the way that terrorist aliases were identified by the NCTA as the primary motivation for advancing universal transliteration standards. However, in the use of the English language as the computational background for the Secure Flight program, the opposite seems to be the case. Name-matching algorithms are not designed to detect risk within names, but rather produce certain names as risky through the technological inscription of accuracy as a prerequisite for statistically parsing assemblages of data elements.

Of course, within the specific context of airport security, names are not racialised in the same way. Certain culturally specific names might be algorithmically interpreted as white, while names in non-Latin script are most vulnerable to processes of racialisation through universal transliteration standards. In preemptive national security strategies, markers of nonnormativity, such as name and gender in SFPD, are not "fixed, ahistorical, or easily read markers of deviance," but are rather "active interpretations that [...] can shift according to context" (Beauchamp 2019: 77). Risk does not simply map onto stable identities of "us" versus "Other," but is differentially applied across shifting definitions of ethnic, religious and cultural status to produce a perceived adjacency to American exceptionalist constructs of a gendered whiteness. In terrorist screening, data populations are relationally computed in order to secure whiteness as a universal norm.

The Secure Flight program is of course not blind to the problems attending its scoring threshold. However, the solutions advanced in federal reports do not seem to address the underlying issue at stake when white data politics becomes imbricated in the name-matching protocol. One proposed solution is the use of approximate string matching, or fuzzy string searching, which computes similarities between various elements of data in order to roughly estimate certain values rather than identifying a precise name. The sophistication of fuzzy logic has been regarded as an asset for updating phonetic algorithms across federal agencies, particularly as a way of dealing with the issues that emerge with the entry of non-English names in databases (OIG 2004: 22–23). Within the context of national security, the advantages of such algorithms far outweigh the negatives: fuzzy logic computation can more readily identify variants of non-English names, despite the possibility that such algorithms might also yield names not included on the watchlists. Nonetheless, approximate string searching does not quite rectify the racialised logics upon which watchlist matching operates. Without addressing how non-English names are considered exceptional, or mobile with respect to some accepted norm of pattern assessment, we overlook the ways in which the racialisation of accuracy safeguards risk-assessment protocol in terrorism prevention efforts. In a world witnessing the collapse of traditional nation-state boundaries and increasing degrees of cultural conservatism, name-matching score values simultaneously work to displace anxieties around the accommodation of cultural hybridity at the same time that they efface the production of taxonomies of racial difference within the militarisation of computational systems.

Conclusion

In locating the biopolitical distribution of risk in contemporary security systems within the medial logics of name-matching score values employed in terrorist watchlist screening, I risk a perilous conclusion – that discrimination is solely the result of technical calibration. Such an argument would fall prey to the "techno-deterministic and discriminatory thinking behind these post-9/11 security measures – that the right technological arrangement, deployed in the right way, can invariably solve any governmental problem, including terrorism" (Werbin 2009: 615–616). However, if error is not an incidental effect of computation but inimical to the purported efficiency of name-matching algorithms in Secure Flight, then we cannot simply hold the tools of big data analytics accountable for digital inequities. More accurate algorithms will not achieve more objective results, but only perturb, mutate and reify nationalist projects of racialisation. To address the discriminatory practices of big data, we must grapple with the political, medial and social formations that justify the unfair distribution of risk across minority populations. The trade-off between accuracy and security in probabilistic name-matching is certainly not balanced. If all populations were equally at risk of being prone to false-positives, then surely adjustments would be made. Since only certain religious, ethnic and immigrant communities are most vulnerable to being cited as terrorists, the discriminatory logics that would otherwise be unacceptable for more privileged travellers are retained.

In imagining more equitable relationships through digital media, we must take responsibility for the way in which we invest technological advancements as efficient, capable and objective. Conceiving of alternative algorithms must not aim to solve digital inequities with more accurate parsing strategies, but rather engage from the beginning the way in which race matters, that way that race is made actionable in the medial logics of computational systems. Simply furnishing accuracy as a standard of measure falls into the trap of a liberal humanism in which one's identity is imagined to be stable enough to cohere across numerous datafied representations. Our systems are never accurate, but only ever *accurate enough* for the preservation of cultural norms. Rather than insisting on greater degrees of granularity and precision in our computational models, we must challenge forms of power that wield accuracy as an indefensible tactic for collecting, sorting and classifying populations under the stratagem of risk. Uncovering the political decisions that constitute the threshold between what is unacceptable and what is accurate enough, we find a familiar story: our deeply held convictions of freedom, privacy and identity are continually computed along old lines of gender, race, class and citizenship.

References

108th Congress (2004): "Intelligence Reform and Terrorism Prevention Act." Public Law, pp. 108–458. Retrieved from https://www.gpo.gov/fdsys/pkg/PLAW-108publ458/pdf/PLAW-108publ458.pdf.

Alsurori, M./Al-Sanabani, M./Al-Hagree, S. (2018): "Design an Accurate Algorithm for Alias Detection." International Journal of Information Engineering and Electronic Business 10(3), pp. 36–44.

Al-Sanabani, M./Al-Hagree, S. (2015): "Improved An Algorithm For Arabic Name Matching." Open Transactions on Information Processing, pp. 2374–3778.

Amoore, L. (2011): "Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times." Theory Culture & Society 28(6), pp. 24–43.

Andrejevic M./Gates, K. (2014): "Big Data Surveillance: Introduction." Surveillance & Society 12(2), pp. 185–196.

Beauchamp, T. (2019): Going Stealth: Transgender Politics and U.S. Surveillance Practices. Durham: Duke University Press.

Bigo, D. (2008): "Globalized (in)Security: the Field and the Ban-opticon." In: D. Bigo/A. Tsoukala (eds.), Terror, Insecurity and Liberty: Illiberal practices of liberal regimes after 9/11. Abingdon: Routledge, pp. 10–48.

Bjelopera, J./Elias, B./Sisken, A. (2016): "The Terrorist Screening Database and Preventing Terrorist Travel." Congressional Research Service Report, R44678. Retrieved from https://fas.org/sgp/crs/terror/R44678.pdf.

Bouk, D. (2015): How our Days Became Numbered: Risk and the Rise of the Statistical Individual. Chicago: University of Chicago Press.

Browne, S. (2015): Dark Matters: On the Surveillance of Blackness. Durham: Duke University Press.

Cheney-Lippold, J. (2017): We are Data: Algorithms and the Making of Our Digital Selves. New York City: New York University Press.

Christen, P. (2006): "A Comparison of Personal Name Matching: Techniques and Practical Issues." Joint Computer Science Technical Report Series TR-CS-06-02, The Australian National University.

Chun, W. (2016): Updating to Remain the Same: Habitual New Media. Cambridge: MIT Press.

Cubitt, S. (2014): The Practice of Light: A Genealogy of Visual Technologies from Prints to Pixels. Cambridge: MIT Press.

Dixon-Román, E. (2016): "Algo-Ritmo: More-Than-Human Performative Acts and the Racializing Assemblages of Algorithmic Architectures." Cultural Studies ↔ Critical Methodologies 16.5, pp. 482–490.

Elias, B. (2014): "Risk-Based Approaches to Airline Passenger Screening." Congressional Research Service Report R43456.

El-Shishtawy, T. (2013): "A Hybrid Algorithm for Matching Arabic Names." International Journal of Computational Linguistics Research, 4(2). Retrieved from https://arxiv.org/pdf/1309.5657.pdf.

Federal Bureau of Investigations (2017): "Terrorist screening center: frequently asked questions." Retrieved from https://www.fbi.gov/file-repository/terrorist-screening-center-frequently-asked-questions.pdf/view.

Foucault, M. (2003): Society Must Be Defended: Lectures at the Collège de France, 1975–76. London: Penguin.

Foucault, M. (2007): Security, Territory, Population: Lectures at the Collège de France, 1977–1978. Houndmills, Basingstoke, Hampshire; New York, NY: Palgrave Macmillan.

Freeman, A./Condon, S./Ackerman, C. (2006): "Cross Linguistic Name Matching in English and Arabic: A 'One to Many Mapping' Extension of the Levenshtein Edit Distance Algorithm." Proceedings of the Human Language Technology Conference of the North American Chapter of the NAACL. New York: Association for Computational Linguistics, pp. 471–478.

Fuller, M. (2005): Media Ecologies: Materialist Energies in Art and Technoculture. Cambridge: MIT Press.

Hall, R. (2015): Transparent Traveler: The Performance and Culture of Airport Security. Durham: Duke University Press.

Halpern, O. (2014): Beautiful Data: A History of Vision and Reason since 1945. Durham: Duke University Press.

Hausken, L. (2017): "The Archival Promise of the Biometric Passport." In: I. Blom/T. Lundemo/E. Røssak (eds.), Memory in Motion: Archives, Technology, and the Social. Amsterdam: Amsterdam University Press, pp. 257–284.

Kahn, J. (2013): Mrs. Shipley's Ghost: The Right to Travel and Terrorist Watchlists. Ann Arbor: University of Michigan Press.

Krouse W./Elias, B. (2009): "Terrorist Watchlist Checks and Air Passenger Prescreening." Congressional Research Service 7-5700, RL33645. Retrieved from https://fas.org/sgp/crs/homesec/RL33645.pdf.

Massumi, B. (2009): "National Enterprise Emergency: Steps Toward an Ecology of Powers." Theory, Culture & Society 26(6), pp. 153–185.

Mbembe, A. (2017): Critique of Black Reason. Trans. Laurent Dubois. Durham: Duke University Press.

McPherson, T. (2018): Feminist in a Software Lab: Difference+Design. Cambridge: Harvard University Press.

Muhammad, K. (2010): The Condemnation of Blackness: Race, Crime, and the Making of Modern Urban America. Cambridge: Harvard University Press.

National Commission on Terrorist Attacks Upon the United States (2004): The 9/11 Commission Report. New York: W. W. Norton & Company.

Parikka, J. (2015): A Geology of Media, Minneapolis: University of Minnesota Press.

Puar, J. (2007): Terrorist Assemblages: Homonationalism in Queer Times. Durham: Duke University Press.

Ralph, M. (2012): "'Life… in the Midst of Death': Notes on the Relationship Between Slave Insurance, Life Insurance and Disability." Disability Studies Quarterly 32(3). Retrieved from http://dsq-sds.org/article/view/3267/3100.

Schuller, K. (2018): The Biopolitics of Feeling: Race, Sex, and Science in the Nineteenth Century. Durham: Duke University Press.

Selod, S. (2018): Forever Suspect: Racialised Surveillance of Muslim Americans in the War on Terror. Brunswick: Rutgers.

United States Department of Homeland Security (2006): "Report Assessing the Impact of the Automatic Selectee and No Fly Lists on Privacy and Civil Liberties as Required Under Section 4012(b) of the Intelligence Reform and Terrorism Prevention Act of 2004, Public Law 108–458." Retrieved from https://www.dhs.gov/xlibrary/assets/privacy/privacy_rpt_nofly.pdf.

United States Department of Homeland Security (2008): "Secure Flight Program – Final Rule," 49 CFR Parts 1540, 1544, and 1560, pp. 64018–64066. Retrieved from https://www.gpo.gov/fdsys/pkg/FR-2008-10-28/pdf/E8-25432.pdf.

United States Department of Homeland Security (2011): "Risk management fundamentals: homeland security risk management." Retrieved from https://www.dhs.gov/sites/default/files/publications/rma-risk-management-funda mentals.pdf.

United States Department of Homeland Security, Office of the Inspector General (2004): "DHS Challenges in Consolidating Terrorist Watch List Information," OIG-04-31. Retrieved from https://www.oig.dhs.gov/assets/Mgmt/OIG-04-31_Watch_List.pdf.

United States Department of Homeland Security, Office of the Inspector General (2012): "Implementation and Coordination of TSA's Secure Flight Program (redacted)," OIG-12-94. Retrieved from https://www.oig.dhs.gov/assets/Mgmt/2012/OIGr_12-94_Jul12.pdf.

United States Department of Transportation, Federal Aviation Administration (1991): "Flight and Cabin Crew Notification Guidelines." 14 CFR Part 108, in Federal Register 56.23, pp. 4322–4326.

United States Government Accountability Office (2006): “Terrorist Watch List Screening: Efforts to Help Reduce Adverse Effects on the Public,” Report to Congressional Requesters GAO-06-1031. Retrieved from http://www.gao.gov/new.items/d061031.pdf.

United States Government Accountability Office (2009): “Aviation Security: TSA Has Completed Key Activities Associated with Implementing Secure Flight, but Additional Actions Are Needed to Mitigate Risks,” GAO-09-292. Retrieved from https://www.gao.gov/assets/290/289632.pdf.

United States Government Accountability Office (2017): “Terrorist Watch List Screening: Recommendations to Enhance Management Oversight, Reduce Potential Screening Vulnerabilities, and Expand Use of the List,” GAO-08-194T. Retrieved from https://www.gao.gov/assets/120/118217.pdf.

United States Transportation Security Administration (2004): “Reports, Forms, and Recordkeeping Requirements: Agency Information Collection Activity Under OMB Review; Secure Flight Test Phase,” Docket No. TSA-2004-19160, in Federal Register 69. 185, 57342-57345. Retrieved from https://www.govinfo.gov/content/pkg/FR-2004-09-24/pdf/04-21478.pdf.

Weheliye, A. (2014): Habeas Viscus: Racializing Assemblages, Biopolitics, and Black Feminist Theories of the Human. Durham: Duke University Press.

Werbin, K. (2009): “Fear and No-Fly Listing in Canada: The Biopolitics of the War on Terror,” The Canadian Journal of Communication 34(4), pp. 613–634.

Werbin, K. (2011): “Spookipedia: Intelligence, Social Media and Biopolitics.” Media, Culture & Society 33(8), pp. 1254–1265.