Research Article

Luke Munn*

Machine Readable Race: Constructing Racial Information in the Third Reich

Abstract: This paper examines how informational processing drove new structures of racial classification in the Third Reich. The Deutsche Hollerith-Maschinen Gesellschaft mbH (Dehomag) worked closely with the government in designing and integrating punch-card informational systems. As a German subsidiary of IBM, Dehomag's technology was deployed initially for a census in order to provide a more detailed racial analysis of the population. However the racial data was not detailed enough. The Nuremberg Race Laws provided a more precise and procedural definition of Jewishness that could be rendered machine-readable. As the volume and velocity of information in the Reich increased, Dehomag's technology was adopted by other agencies like the Race and Settlement Office, and culminated in the vision of a single machinic number for each citizen. Through the lens of these proto-technologies, the paper demonstrates the historical interplay between race and information. Yet if the indexing and sorting of race anticipates big-data analytics, contemporary power is more sophisticated and subtle. The complexity of modern algorithmic regimes diffuses obvious racial markers, engendering a racism without race.

Keywords: reich, race, dehomag, punch-card, information

The Nazi utopia of an Aryan nation required racial information. An immense project would be needed to tackle the current *Völkerchaos*, or chaotic mix of incompatible races, and replace it with the *Volkseinheit* of a racially pure Germany. By supporting the reproduction and flourishing of one race, and suppressing, excluding or annihilating others, a racially pure society would be established. But first one needed to know the precise composition of the existing society. As SS Obergruppenführer Reinhard Heydrich stated, this meant the "sorting out and the evacuation of the worthless races and the furthering of the Germanification process of the remaining good races"; an informational analysis that captured race would be the first step towards that goal: "A precondition for the critical examination of the population is an inventory of the people that once and for all registers all the persons of the Protectorate and sorts them according to certain aspects" (Aly and Roth, 2004, p. 52). To sort out society, it would first need to be sortable.

This paper argues that informational processing drove new structures of racial classification in the Third Reich. Information technologies, from the list to the punch-card and the census, required a more precise or even procedural definition of race. Only when race was formalized could it become race-as-information, integrated into technical systems and rendered capturable and calculable. To trace this trajectory, the paper draws upon a broad spectrum of published material: IBM archives, racial legislation, punch-card specifications, academic histories, and Nazi-era documents. Synthesizing these accounts provides new insights into the gradual codification of race.

The first three sections examine how political imperatives and technical affordances dovetailed together to drive the pursuit of a more fine-grained racial classification. Section four follows this quest for total information through to its apex in a single code for every individual in the Reich. The fifth section

*Corresponding author, **Luke Munn,** Western Sydney University, New Zealand, E-mail: luke.munn@gmail.com

demonstrates how these informational technologies created issues, anticipating "big data" problems. And the final section more explicitly links these prototypical technologies with present day technologies, putting punch-card machines in dialogue with modern platforms like Palantir. Taken together, these historic technologies demonstrate the longstanding intersection between race and information, while also foregrounding some of the new logics at work in contemporary regimes of racial classification.

As a contribution to knowledge, this work draws a line from early technologies through to contemporary platforms. Racist computation, in this sense, did not suddenly emerge with the microchip or the mainframe, but has a much longer lineage, one grounded in the informational regimes of the census, the list, the punch-card. These prototypical systems, even if slower and cruder, lay the groundwork for the more sophisticated technologies to come. But perhaps more subtly and importantly, the historical moments examined here show how *race itself* as a classifier emerges in part from a computational requirement. To operate at all, these technologies needed categories that could be punched, sorted, and counted. The vagueness of Jewish stereotypes founded on appearance or myth were insufficient.[1] Jewishness needed to be made machine-calculable, based on a kind of algorithmic logic that could be repeated over and over again, at scale. The ambiguous contours of a people group needed to be sharpened, to have hard-edged boundaries. The logic of information required a discrete category. "Confused and unsystematic race-thinking aspired to become something more coherent, rational, and authoritative," observes Paul Gilroy (2001, p. 31), "'race' began to create new possibilities and orchestrate new varieties of knowledge and power centered on the body." Gilroy consistently puts "race" in quotations, signalling that grouping people according to type is a powerful fiction. Rather than some predefined definition, this article too stresses how race is constructed over time and in tandem with particular devices and techniques. Attending to these early conditions demonstrates how race is not merely a pejorative ideology held by racist people, but one that draws on—and is amplified by—the imperatives of information technologies.

1 The Holleriths and the Census

Almost immediately after coming to power in 1932, the National Socialist party ordered a new census to be taken (Aly and Roth, 2004, p. 16). The previous census conducted in 1925 was now woefully out of date, with 8 year old statistics. More importantly, it had failed to adequately classify the population, prying out the particular demographics that were now of such interest to the administration (Black, 2001, p. 54). A new census was urgently needed, a great ingestion of information that would capture and sort the population, and do it in record time. More data than ever before was required, and required rapidly; manual methods would simply not suffice. The solution was the Deutsche Hollerith-Maschinen Gesellschaft mbH, or Dehomag, the German subsidiary of IBM before and during WWII (p. 55).

Like IBM's other worldwide subsidiaries, Dehomag was centred on the same core technology—the Hollerith machine. The machine was originally designed specifically for the counting of population. The U.S. Constitution mandated a census every 10 years. But counting the 1880 census was so slow that it wasn't completed until 1888, making the information practically irrelevant (Aul, 1972). German-American engineer Herman Hollerith, a former employee of the U.S. Census Bureau, decided to address the problem by automating the count. To do this, he designed paper cards marked with specific categories: age, race, gender, location, and so on. Holes within these categories delineated male from female, black from white, soldier from civilian, with each "bearing a specific relation to each other and to a standard" (Hollerith, 1889). More detailed information on race went along with this more mechanized process. As Paul Schor notes (2017, p. 109), enumerators for the upcoming 1890 census received these instructions:

> Be particularly careful to distinguish between blacks, mulattoes, quadroons, and octoroons. The word "black" should be used to describe those persons who have three-fourths or more black blood; "mulatto," those persons who have from

---

[1] For example, the antisemitic caricatures found in weekly tabloids like *Der Stürmer* or the chapter "How To Recognize a Jew" in the virulently antisemitic book *Der Gilfpilz* (The Poisonous Mushroom). Ernst Hiemer, *Der Giftpilz* (Nuremberg, Stürmerverlag, 1938).

three-eighths to five-eighths black blood; "quadroon," those persons who have one-fourth black blood; and "octoroon," those persons who have one-eighth or any trace of black blood.

Already then, in tandem with Hollerith's calculation of population, there is a calculation of race. Thresholds are established that constitute certain races. In order to record the correct information for a specific individual, the human enumerator must essentially step through a basic algorithm. IF the percentage of black blood is larger than three-fourths, the subject is "black"; ELSE IF the percentage is larger than three-eighths, the subject is "mulatto," and so on. Once race had been mentally "computed" in this way, it could be recorded on a punch-card, stored and retrieved. Though crude, these early calculational technologies point to the historical co-emergence of race and information. Indeed the reproduction of systematic racism on a vast scale is often enabled by informational architectures that provide it with the necessary technical affordances. As Jonathan Beller reminds us, inequality is "organized and enforced in a matrix of valuation that tracks and weights factors of whiteness" (2018, p. 3). The census, like more contemporary informational architectures, is a technology that decides what counts. And some things count more than others.

Upon receiving a completed census form, an operator would translate the information by punching the relevant holes. Automation was therefore by no means total. But the rationalised process of recording the data into the system and the mechanisation of sorting and counting it was vastly more efficient than anything that had come before. Rather than the manual clerical work usually employed, a Hollerith machine could tabulate thousands of cards per hour. Despite the larger population (62 million vs 50 million) and the more detailed information requested, the machines completed the computation two years earlier than expected, in 1896, with significant cost savings compared to the manual approach (Aul, 1972). The "Holleriths" had aced their first real-world test, demonstrating their worth to potential public and private clients.

Dehomag held out that same promise of big-data to its clients in the Reich. In a speech to government officials (Deutsche Hollerith Maschinen Gesellschaft m.b.H, 1934, p. 39, translation mine), Dehomag president Willy Heidinger enthusiastically proclaimed the capabilities of the company's technologies using nationalist Nazi terminology:

> Like the doctor, we dissect the national body (*Volkskörper*) down to the individual cells. We specify the individual characteristics of every single national comrade (*Volksgenosse*) on a card. These cards are not dead but rather display an eerie life when, at a speed of 25,000 cards per hour, the national body is grouped according to certain characteristics in our sorting machine, and the relevant values calculated and determined in our tabulating machine. We're proud to be permitted to participate in this work, a work which provides the doctor of our national body the material for his examination, so that he can determine whether the calculated values are harmonious in terms of public health, i.e. standing in healthy relationship to one another, or if these morbid conditions must be medically corrected through interventions.

In this vision, the nation was sick, weakened by its impurity. But diagnoses had so far failed to locate the disease. The crude categories of previous population surveys were no longer good enough. More articulated information with more refined demographics was now necessary. A population would be encapsulated and the subjects within it characterised and clustered. In doing so, computation would parse out the heterogeneous pockets of difference concealed in a previously monolithic public.

While certainly infused with corporate grandstanding, Heidinger's speech is nonetheless chilling in its anticipation of the "dissection" of the national body and the devastating interventions the Nazi regime would implement in order to correct its "morbid conditions." Indeed the very idea of using information as a pincer to isolate and extract a pathologic race from the broader social body suggests genocidal concepts. As Andreas Kranebitter observes (2012, p. 96, translation mine): "Counting and capture, with its attendant separation, selection, and finally concentration of the societal-biological alien—the 'sick portion' of the 'healthy national body'—was an important accessory to extermination ideologies, or even its core ideology." Yet if the concept of a physician extracting cancerous tissue was ingrained in Nazi rhetoric, the transformation from mere metaphor to massive social engineering project required a modern mindset and technical amplification. As Zygmunt Bauman suggests (2001, p. 177), if it is difficult to conceive of extermination without medicine, and "its model of health and normality, strategy of separation and technique of surgery" is almost impossible to imagine it without "the engineering approach to society"

and the "practice of scientific management." Holleriths would be the critical instruments for slicing into the social body in a more articulated fashion, and they would do so at the scales required of the sovereign.

Dehomag's promise was underpinned by its technical capabilities. For the 1933 census, Dehomag was commissioned to manage the count of Prussia, Germany's most populous state comprising three fifths of the total population. Along with personal and work information, the census requested details of religious affiliation and birthplace, with a note that these were "in the interests of investigating race-biological questions" (Domscheit-Berg, 2018). To mechanize this racial investigation, the Dehomag solution was put to work—machines were leased, custom punch cards printed, operators and editors hired. "About 900 Dehomag keypunch operators and editors, punching and verifying around the clock in three 7.5 hour shifts, transferred data on about 40 million Prussian citizens onto 60-column punched cards at a rate of 150 cards per person per hour," recount David Luebke and Sybil Milton in their incisive account of tabulation technology and persecution in Nazi Germany (1994, p. 25): "These cards were then processed with 35 Dehomag sorters with a capacity of 24,000 punched cards per hour, 13 combined Dehomag sorter-tabulators with a capacity of 20,000 punched cards per hour, and 25 Dehomag tabulators with a capacity of 12,000 punched cards per hour." The census was processed rapidly, a fact not only due to the immense amount of personnel involved, but also to the efficiency of machinic processing. The Holleriths had once again proven their worth.

2 Constructing Race

But the 1933 census was just a start. Indeed, it was quickly discovered that this ambitious acquisition of information was entirely inadequate. In the census, Jewishness had been defined as being a *glaubensjuden*, or religiously observant Jew. However as Luebke and Milton point out (1994, p. 30), this meant that non-practicing Jews or those of mixed Aryan/Jewish descent had not been isolated; the same problems applied to Roma and Sinti, who marked Christian as religion and thus fell back into the general population pool. Contrary to Heydrich's wishes, these "worthless" demographics had not been sufficiently sorted out. The sociotechnical problem of adequately defining what race meant (blood? marriage? religion?) and integrating that delineation into the spatial constraints of a punch card system had proven too difficult. Rather than providing penetrating insights into the makeup of the population, the survey instead revealed the complexity of accurately drawing out the racial categories that so interested the National Socialist government.

More exhaustive information was required. Firstly, this meant more complete data. The previous census had been only partially tabulated by machine. But as Luebke and Milton note (1994, p. 28), the entire 1939 census would now be mechanised, it would now include all of Germany (including the newly annexed Austria) and it would be consolidated under a single new authority—the Statistisches Reichsamt. Rather than Prussia's 40 million inhabitants, this would entail counting and sorting the fields of 80 million citizens. Additionally, this more complete understanding of the population would be matched with a more complete understanding of the subject. In the previous years, the IBM cards had now switched to using rectangular holes. This acted as a type of physical compression, allowing more categories to be included on a card. For the 1939 census, the machines were all retrofitted to include the latest 80 column design, rather than the previous 60 column version (Luebke and Milton, p. 28). By doubling the remit of Dehomag and swapping to the new cards, the data captured effectively increased by 233% between 1933 and 1939. To be sure, the census was only one capture operation of many. Indeed the competition between various state agencies meant that divergent modes of data collection proliferated. Throughout the 1930s, as Luebke and Milton observe (p. 30), information capture "was neither centralised nor systematic. Competition between state and Nazi party agencies resulted in administrative confusion and the proliferation of institutions charged with the same general tasks." On a practical level, even the expanded 1939 census could not capture all of the information required, but instead functioned as an important supplement to other forms of data. But the expansion of subjects and fields evident between the census of 1933 and that of 1939 demonstrates the asymptotic quest for an exhaustive repository of data—a totalisation of information.

Secondly, this meant more sources of data. As suggested, the census was far from being the only point of supply for information. In fact, the Nazi regime was highly aware of the importance of detailed, up-to-date intelligence in order to complete its objectives. The result was a great eruption of information from a vast array of entities that touched on every facet of society: law, education, medicine, reproduction. Information from all of these sources should be regularly collected and centralised into a single, all-encompassing index. As one directive outlined (Aly and Roth, 2004, p. 107):

> data from all departments, all of which must contribute to the effort, must be included in this collection: health welfare institutions of all kinds, economic welfare, youth and education, court rulings, special foster care cases, cases of infertility, absolutely all sentences in criminal cases regarding drug matters, repeat offenders, and all other sentences requiring character evaluations, medical examiners' reports, medical reports from companies dealing with health, retirement, and accident insurance, reports from the Hitler Youth, sports clubs, the labor service, the military service, medical evaluations for civil servant applicants, marriage applicants, college examinees, examinations of the party organisation and their health-welfare work... health reports from hospitals, mental institutions, and so forth. All of these offices, organisations and institutions will receive information as long as they contribute to the collection of information.

Effective governance depended on a more comprehensive understanding of the populace. And this understanding could only be attained by integrating a constellation of divergent, overlapping informational sources.

Thirdly, this meant more detailed data. For the Nazi party, the Jew was always the target. But what, exactly, constituted a Jew? A satisfactory answer had never been found. As Holocaust historian Raul Hilberg noted (1985, p. 65), the inability to precisely define the race had been a "stumbling block for an earlier generation of anti-Semites." Of course there were existing notions of Jewishness. As many scholars have shown, Europe was home to an extended history of antisemitism, a long-standing quest to identify Jews within a population, to objectify them in often devastating ways, and to exclude them from circuits of social and cultural capital (Arendt, 2012; Braun, 2000; Lindemann and Levy, 2006). This vilification became more formalized by the pseudo-scientific discourse of eugenics at the turn of the 20th century, first in the USA, then as an export to Germany (Whitman, 2017). By 1927 research programmes and centers like the Kaiser Wilhelm Institute for Eugenics in Berlin would open, concentrating "on a comprehensive project on racial variation as indicated by blood groups" (Kühl, p. 21). The scientific discussion of racial hygiene would eventually be turned into social engineering with the German sterilization law mandating higher reproduction rates for Aryan women alongside forced sterilization for the "inferior races" and those who might produce diseased or deformed offspring (Reichsgericht, 1933).

These practices, while devastating in themselves, slowly developed a typology of race. Whether through the pseudoscience of physiognomy or the burgeoning new field of familial research, these fields began stitching together a social consensus regarding Jewishness. Certain biological and physical markers were associated with being a Jew. Racism's goal, as Bauman suggests (2001, p. 169), was to postulate a "systematic, and genetically reproduced distribution of such material attributes of human organism as bore responsibility for characterological, moral, aesthetic or political traits." If this matrix of key factors could be "discovered," then individuals or even whole groups could be tested against it in order to determine their race, and by extension, whether they would contribute to the new national people's body or infect it. Yet even as late as 1933, for example, the so-called *Arierparagraph* was dividing the world into a highly simplistic binary: the Aryan and the non-Aryan (Reichsgericht, 1933).[2] To properly isolate and exterminate those races deemed worthless, demographics would need to be subdivided and definitions honed. More fine-grained information would identify and amplify racial difference.

---

2 Aryanism was based on ancestry. As Raul Hilberg notes, to prove one was Aryan, no less than seven documents were required: a birth or baptismal certificate, the certificates of the parents, and the certificates of the grandparents. Raul Hilberg, *The Destruction of the European Jews* (New York: Holmes & Meier, 1985), 73.

3 From the non-Aryan to Nuremberg

For the Nazi regime, a precise definition of the Jewish race was still needed. Indeed this racial ambiguity had led to vigilante violence against those suspected of being Jews; to restore order, the government decided that "unambiguous laws" were required (Whitman, 2017, p. 83). In one sense, then, this formalization was a move to reestablish the juridical process through a clearer definition. Yet the eradication of ambiguity is also a crucial preparatory move for informational capture. "By imposing a mathematically precise form upon previously unformalized activities, capture standardizes those activities and their component elements" states Phil Agre (1994, p. 120). For optimal processing, vagueness must be reduced to a minimum. Informational architectures instead thrive on gapless definitions that leave no liminalities: zero or one, male or female, Jew or non-Jew. In this regard the Nazi informational apparatus, even in its comparatively crude mix of lists, handwritten documents, and punch-cards, anticipates a more contemporary algorithmic system, which "has no tolerance for the emergent or half-seen figure" (Amoore, 2014, p. 111).

By the beginning of 1935, as Hilberg notes (p. 69), the problem of defining Jewishness was beginning to be tackled; in September of that year, instigated by Hitler himself, the "Law for the Protection of German Blood and Honor" was drafted in just two days, and on November 14th, the "Reich Citizenship Law" was amended. These Nuremberg Race Laws, as they're commonly known, legally formalised what it meant to be a Jew. A full Jew was anyone who had three or more grandparents who were Jewish; those with one or two Jewish grandparents were classified as "mixed" 2nd or 1st degree Jews (*Mischling 2.Grades oder Mischling 1.Grades)*; and finally there were four legal conditions for catching so-called *Geltungsjude*, or Jews by law, which consisted of being enrolled in a Jewish congregation, being married to a Jew, or being an offspring of Jews, either by marriage or by extra-marital affair (Erste Verordnung Zum Reichsbürgergesetz, 1935).

Thus if antisemitism has a long lineage, Nuremberg logic nevertheless demonstrates a new systematization of this trajectory into machine-readable race. The definition possessed a kind of procedural quality, calculable by anyone. In fact Hilberg, even in his straightly historic account, says an "automatic sorting method separated the 'non-Aryans'" (p. 72). The Nuremberg Race Laws established what contemporary information systems would call conditional statements: IF the subject has 3 or more Jewish grandparents, THEN the subject is a full Jew; ELSE IF the subject attends a synagogue OR is married to a Jew THEN the subject is a Jew-by-law. The fact that this logic is circular (a Jew is a descendant of a Jew) was ignored. Even if it was tautological, the key thing was that it could be made operational, transformed into a procedure or routine that could be run on any population set. The Jewish race was finally rendered rule-based. Together these conditions took the smooth space of ethnicity, the complex details of ancestry, and the rich lineage of religion, and distilled them into the hard-edged category of race. Jewishness became rationalised into a mark, a rectangular void, recognisable by machine.

The Nuremberg Laws also demonstrate how information itself makes new racial definitions possible. The proliferation and consolidation of data in the Third Reich, from handwritten synagogue lists through to police registers, family trees and punch-cards, opened up novel variations of Jewishness. Disparate information could be assembled together and parsed through the logic of Nuremberg, a logic that established the model of a race (*volljude)*, but also allowed for subcategories (*mischling)*, and provided a catch-all container (*geltungsjude).* More detailed information thus allowed the Reich to move past the crude all-or-nothing designation of the "full" Jew. The additional definitions of "mixed" Jews or Jews-by-law constitute a more articulated definition of race, but also include those who might have escaped from a too narrow container. The result is a double move that constructs sub-categories of race while simultaneously casting more subjects into a racial category deemed to be inferior. As Judy Scales-Trent observes (2001, p. 263) because race is a construct, its rules can be redefined, allowing an oppressor group to "vary the size and composition to suit its needs." Supplied with a greater variety of more detailed data, the Reich could weave a wider net with a finer mesh.

The codification of Jewishness arrived at by the Nazi government foregrounds how race emerges from an interplay of social definitions and technical conditions. Stefano Harney and Fred Moten defined governance as "the order that collects differences" (2013, p. 54). But this is not simply a process of gathering up existing distinctions, like plucking pebbles from the ground. Technical governance does not just apply

pre-existing categories to larger population sets and filter them faster. Instead of differences being merely collected, the census demonstrates how they are actively *constructed* via a social/technical circuit.[3] As Bowker and Star stress (1999, p. 34), we must recognize the deep "interdependence of technical networks and standards, on the one hand, and the real work of politics and knowledge production on the other." The implementation of new technologies creates the need for a new set of social categories. Drawing on historical antisemitism and racial hygiene along with the pseudoscience of phrenology and physiognomy, social operations go to work, locating these categories, rationalising them and delivering them up to technologies. Once technically legible, these categories become formalised, fine-tuned and extended, fed back into understandings of self and society. I am a Jew because the punch card says I am a Jew. "Something was not known; then information arrives, namely that these, and none other, are the facts of the matter," Niklas Luhmann argues: "Now one has knowledge and, as a consequence, one cannot help orientating one's subsequent operations by means of this knowledge" (2006, p. 40).

Once these labels were in place, the Hollerith machines could get to work with characteristic speed and efficiency, parsing the population into discrete categories. Punch-card "techniques assisted in calculating the number of Jews, how many had emigrated, and the location of those remaining," writes Paul Weindling (2010, p. 326), "they calculated how many full, half, and quarter Jews still lived in the Reich." Calculation here is a technical operation, but also an epistemological one, establishing the boundaries of what is to be known, thought, and actioned. Ethics are no longer in scope; values are irrelevant to the task at hand. In this sense, race-as-information is not just an optimisation but an anaesthetisation, deadening empathies and prepping the space of the effective procedure: ("given the data, filter the following..."). As Zygmunt Bauman noted (2001, p. 235): "Technical responsibility differs from moral responsibility in that it forgets that the action is a means to something other than itself." Stress is shifted to the efficiency of techniques, the optimisation of procedures, and the accuracy of results. Focus is ratcheted down to the engineering of informational structures—the craft of classification.

Indeed this blinkered engineering perspective is precisely what emerges from a Hollerith newsletter from the time period. Faced with the complex challenge of rapidly isolating all the "abnormal people" in Denmark, the company congratulated itself on developing a particularly efficient solution in the form of a "new sorting machine brush holder" able to filter the population in 3 different ways simultaneously (IBM Germany, Reprinted in 1983). In this context, mental illness, like race, becomes a variable to be sorted on. Information assisting in the creation of a racially pure society is obtained from a process which is bureaucratic, even banal. The transformation of race into information, then, was not just productive managerially but ideologically, dehumanizing those it targeted. As Rainer Zons once stated (2001, p. 37, translation mine): "It is precisely the neutrality of the archival medium and its procedures of storage, transmitting and calculating that ensure its inevitability and mercilessness." Presented as an informational archive, the horrific objective of filtering out "inferior races" and "abnormal people" from a population became a technical problem with an optimum solution.

4 Master Code, Master Race

Yet the apex of the Reich's informational imaginary was still to come. As noted, the polycentric nature of state agencies, particularly early on, meant that data collection methodologies diverged, producing heterogeneous datasets. This issue wasn't insoluble. Indeed, it appears that it was primarily the cross-

---

3 To avoid any misconceptions, that race is "constructed" in no way implies that it is illusory or insignificant. If a distinct Jewish race was fabricated and formalized into a procedural logic by the Nazi regime, this doesn't lessen its impact. Indeed, one of the "advantages" of investigating the tragic history of the Holocaust is that it absolutely crystallizes the stakes of race as a category and the power a hegemonic regime wields in constructing it. If too, in our modern era, the biological foundations of race have been thoroughly debunked, race and racism is a daily reality. As Teresa Guess notes, drawing on Durkheim, race is a social fact: "every way of acting, fixed or not, capable of exercising on the individual an external constraint; or again, every way of acting which is general throughout a given society" (quoted in Guess, 2010, p. 655). Thus, if race is constructed, it nevertheless generates enormous forces, embedded within sociomaterial configurations, that shape practices, structure opportunity, and impinge on life itself.

indexing of three datasets that supported the profiling, location, and deportation of many Jews. As Luebke and Milton explain, (1994, p. 33) total numbers of Jews in areas were produced from a subset of the 1939 census called the *Volkstumkartei* (Ethnic Register); in order to pinpoint victims this was matched against two other sources: local police registers and Jewish congregational lists (*Judenkartei*); "in short, Berlin dictated aggregate numbers, while local police and Jewish communal officials drew up lists of the names and addresses of individual deportees."

But as power was consolidated under the SS and datasets formalised, a new dream arose—a single number for every subject in the Reich, a master code that could be cross-matched to any and all information on an individual. If this number, "were to be integrated into a central filing system based on the punch card principle, then it would be possible to create links to other existing card files," explain Aly and Roth (2004, p. 134), then "the era of a 'final accounting of humans' would be ushered in, at which point individuals would not only be inventoried at certain intervals, but on an individual and permanent basis." Rather than the speculative correspondence of name or location, information from disparate sources and times could be cross-indexed authoritatively—a unique identifier would provide the common token.

The responsibility for its implementation fell to Albert Speer's *Maschinelles Berichtswesen* (MB) or Mechanical Reporting System. In some respects, the MB was well placed to carry out this programme, having recently incorporated the Hollerith Department and its technocratic leader Kurt Passow. "In 1941 Passow proposed a standardized, national numbering system that would make the entire economy 'machine-readable'" (Tooze, 2001, p. 256). But up until this point, the agency's main duty was to maintain up to date information about the use and location of weapons and munitions. From 1943, the MB also commanded a medical reporting division, using punch card technologies to document the medical status of millions of soldiers and maintain a running tally of the wounded and dead. Suddenly, Speer's System "found itself as the focal point of the ultimate utopian vision of population registration" (Aly and Roth, 2004, p. 124). This required new cards, new procedures, and an expansion of responsibilities. And yet the machinic apparatus was already in place—a shift from weaponry to citizenry, from patients to persons was evolutionary rather than revolutionary. From a British intelligence document we learn that "at the beginning of 1944, the MB developed a modern method of registration of persons" and that already "towards the end of 1944 the method of individual registration was applied experimentally for the evaluation of an investigation, made by the *Rüstungsamt*, on toolmakers and jig-makers, employed in armament factories" (Hoeffding et al., 1945).

However it's also worth complicating this imaginary of totalizing information. Developing a workable system of *Personaleinzelerfassung* (Individual Personal Registration) was always going to be extremely complicated. As Tooze asserts (2001, p. 257), this would have required "coordinating the registers of the personnel offices of local business, the local population register, the registers of local hospitals and the local police." Given the informational challenges, it is unsurprising that no functional system was ever implemented. Albert Speer himself painted a more haphazard picture of the Mechanical Reporting System. Interrogated by the Allied Forces in 1945, he admitted that there would often be time lags, "(up to two months) between completion of questionnaires and production of results" (Speer, p. 2). Though the System provided a central hub for reporting and analysis, the collection of production statistics itself from each of the sectors was a manual affair. Speer's assistant would "spend a couple of nights on the telephone at the end of every month calling the *Sonderausschuss* leaders and getting from them figures for their respective products" (p. 2). As bombing took its toll on infrastructure and communications began unraveling, even this became difficult. Already by the autumn of 1944, as the Allied forces noted in the same report (p. 2), "this kind of telephoning became impossible." The chaos of the war also began to impinge on the statistical predictions that aided the Nazi government in its military planning. In 1942, the System "was still able to furnish such figures for 1-1 ½ years in advance," but this predictive ability was gradually shortened, until by mid-1944, "they were only able to commit themselves for one month ahead" (p. 2). Opposing military forces were rapidly closing in. Time was not on their side, and the dream of perfect administration through perfect information would never be realised.

5 Big Data Solution?

In fact even before the hard-stop of military surrender, the fantasy had been bittersweet—Dehomag's technology seemed to be a double-edged sword that often swung profits and power in its own favour. Firstly, Dehomag established a big-data condition. This relentless push towards more information often overwhelmed government agencies. Take, for example, the Race and Settlement Office (RASO). Responsible for verifying the purity of ancestry and approving marriages, the RASO quickly became inundated with paperwork. As the conditions became tighter and the categories more articulated, the labor involved in verification spun out of control. Their statistical chief protested that, "the [manual] way in which the files are [currently] stored, makes any quick and efficient survey impossible. It would require months of work looking through individual files to answer even one [racial] question" (Black, 2001, p. 294). Painstakingly combing through records by hand would simply not suffice. Automation was therefore critical, and the solution evident. Indeed, the Hollerith system had already been effectively deployed at other agencies, agencies vying for attention and funding within a highly competitive government arena. "For every single one of the additional future tasks," the chief argued (p. 294), "months of tedious clerical work would be necessary just to determine how many and which [racial] petitions are involved. The punch card system would be able to determine this easily, quickly to the desired date...therefore, card indexing is indispensable...the exact instrument for complete surveillance both on a large scale and down to the smallest detail." Agencies like the RASO were faced with a choice—adopt the Dehomag system, lease its machines and pay its licenses—or drown in a deluge of data.

Secondly, Dehomag required transparency. The Hollerith system was far from being a plug-and-play solution. Granted, the Hollerith machines comprised a common set of core routines: counting, tabulating, sorting. But each client has its own domain and its own particular problems. This meant a long-term partnership rather than a one-time sales pitch. As Black stresses (2001, p. 273), the collaboration "began with a protracted investigation of the precise data needs of the project, as well as the people, items, or services being tabulated. This required IBM subsidiary 'field engineers' to undertake invasive studies of the subject being measured, often on-site." To design a punch card that could be used for capturing information and producing insights meant understanding how that information was captured and what insights were desired. This, in turn, meant understanding the economies of each sector, the objectives of each organisation, the operations of each client. Over time more and more national agencies adopted the platform, working closely with the firm. IBM's lawyer in Berlin stated that the degree of collaboration and customisation necessary for the integration of the Hollerith system "secures to the Dehomag a contact and insight into the big business of the nation superior to any other company" (p. 296).

Finally, Dehomag locked in the client. The broader Dehomag ecosystem was intentionally designed *not* to be interoperable with other systems. The core information medium—the cards—were composed of specific paper stock with particular dimensions and holes punched at certain intervals. They couldn't simply be transferred into another system. On top of this, Dehomag kept a tight rein on its supply chain and material ecosystem, a business strategy long ingrained within IBM (Pugh, 1995, p. 247). The paper necessary for the punch cards was controlled by Dehomag, who could draw on IBM's US reserves rather than Europe's scarcity; the printing of the cards themselves was controlled by Dehomag, who only approved a handful of factories; and finally the sorting and counting machines themselves were controlled by Dehomag, who always leased them rather than selling them (Black, 2001, p. 293). If any one of these components was missing, the entire system would collapse. Of course, alternative systems did exist, such as the Powers Tabulating Machine Company in Britain and France. But the switching costs established heavy friction—locating, integrating, and scaling up an alternative solution would be exorbitantly costly both in time and money. Platform lock-in was complete.

Dehomag thus offered an automated solution for information processing, the only viable solution in a rapidly expanding data arms race. This solution was highly customised, entailing an intimate knowledge of each client's internal operations. And this solution was tightly controlled—the company's vice-like grip on its supply chain and machines meant it always kept the upper hand. Thus, when IBM in New York finally began to cut ties, the National Socialist government was anxious indeed. As Black summarises (p. 292), "elaborate

data operations were in full swing everywhere in Germany and its conquered lands. The country suddenly discovered its own vulnerable over-dependence on IBM machinery." The NS government scrambled, seizing rival machines and attempting to develop a bespoke solution, to somehow sever their reliance on machinery that Dehomag had orchestrated. But, as Black outlines (p. 296): "Berlin really didn't know what to do. They stole some IBM machines in France, purchased control of a Powers subsidiary, and brought in Bull machines, all envisioning a new cartel. None of it was coordinated, but something had to be done to counteract Germany's dependence on IBM." The flip side of the informational imaginary was beginning to emerge. A private company had profited greatly by providing governments on both sides with power underpinned by their engineering expertise, their processing labor, and their informational architecture. But this power—like Dehomag's leases and licenses—was temporary. It could be retracted at any time.

6 Deconstructing Race

Dehomag provides an example of a historical moment in which technically amplified regimes of classification were deployed on populations. Punch card technologies were certainly slower and cruder than their contemporary big-data counterparts. But if their mechanisms were clunkier, their imaginaries of power through knowledge are not dissimilar: by capturing all the information about a population, integrating it into a single system, and rendering it machine-readable, individuals can be made knowable and therefore sortable. "In enabling politically effective social engineering through sociological insights and especially in its general phenomenon of 'process-generated data' the Nazi practice of general data collection "connects with present day attempts at data analysis and their attendant methodological limits" observes Andreas Kranebitter (2012, p. 97, translation mine). New technologies renew age-old dreams.

Of course, in contemporary informational architectures, race is generally treated as a "protected class," a group of people who share a common trait that cannot be discriminated against based on that trait. Yet as numerous scholars have pointed out, race continues to persist, albeit in less overt ways. Safiya Noble's *Algorithms of Oppression* (2018) explores how search engine results for terms like "black girls" perpetuate racial stereotypes, reproducing historically racist tropes in a modern informational environment. Similarly Cathy O'Neil's *Weapons of Math Destruction* (2018) investigates a litany of cases, such as parole recommendations, where algorithmic decisions based on ostensibly neutral data are anything but unbiased. Northpointe's widely adopted software, for example, provides courts with recidivism risk scores based on a defendant's information. As information volumes increase and public agencies come under performance pressures, technologies like these seem to offer a lifeline, combing through case files and rapidly delivering a decision with certitude. But a closer look reveals highly uneven results. A *Propublica* report compared an individual's risk score against their actual criminal charges over the next two years. In their statistical analysis, the authors (Angwin et al., 2016) found that even when race was isolated: "Black defendants were still 77 percent more likely to be pegged as at higher risk of committing a future violent crime and 45 percent more likely to be predicted to commit a future crime of any kind." None of these systems explicitly leverage race in their calculations, yet once scrutinized, their systemic bias toward whiteness often becomes apparent. What these and other scholars have begun to reveal in the last few years is a form of informational discrimination that persists despite protected classes—a kind of racism without race.

Arguably the closest informational system to Dehomag's punch-card processing is Gotham, a big-data analytics product developed by Silicon Valley firm Palantir.[4] Founded by Peter Thiel in 2003, the company

---

4 Of course, the different understandings of race in the U.S.A. and Germany should be acknowledged. While *rasse* may equate to race in English, the concept of race in each country emerges from a distinct history and a particular set of ideologies and philosophies. One of the lenses to understand these distinctions is through the treatment of children of color in post-wartime, a comparison undertaken in *Race After Hitler: Black Occupation Children in Postwar Germany and America* by Heide Fehrenbach. As George Fredrickson writes: "If Germans endowed themselves with a 'racial' identity and excluded others from it, Americans tended to racialize others and consider themselves simply human" (quoted in Fehrenbach 2005, p. 1). These distinctions acknowledged, there are also strong overlaps and connections. Scales-Trent (2001) notes similarities in the way that race was sociolegally defined by both governments, while studies by Kühl (2002) and Whitman (2017) trace how Germany used the United States as a model for racial laws.

sought to synergistically bring algorithmic processing and human analysis together in one platform. In the backend, Gotham provides the technological stack necessary to store massive repositories of information and render it queryable in real-time; in the frontend, Gotham provides a highly visual interface and a suite of intuitive tools allowing analysts to discover patterns, follow leads and locate individuals (Munn, 2017). Palantir's technology has now been adopted by dozens of blue chip corporations as well as government agencies; Airbus, Walmart, and GlaxoSmithKline sit alongside clients like the NSA, the US Army, and Immigration and Custom Enforcement (Alden, 2016). Like Dehomag, Palantir bills itself as the solution to the problem of overwhelming information, providing clients with the ability to extract signals from all the big-data noise. Like Dehomag, Palantir operates on a licensing business model, providing servers, software and its "forward-deployed" engineers on a contractual basis for millions of dollars per year. And like Dehomag, Palantir promises to integrate all of a client's disparate informational sources into one compatible whole, providing one platform to rule them all.

One of Palantir's long standing clients is the Los Angeles Police Department or LAPD, who credit the platform with "helping them make sense of it all" (Palantir, 2013). Yet the use of this information architecture by the LAPD is far from racially neutral. In 2018 the grassroots activist organisation Stop LAPD Spying released a report on predictive policing in Los Angeles. The patterns and insights provided by Gotham are based on the data available, yet the source of that data is often overlooked—racially inequitable policing practices produce a racially inflected archive. "Compared to whites" notes the report, "Black and Latinx individuals are roughly 75% more likely to be stopped by police, up to 127% more likely to be frisked, and up to 76% more likely to have their vehicle searched" (Stop LAPD Spying, 2018, p. 17). Every traffic stop creates a report; every stop-and-frisk contributes another data point. As the report notes: "Historical crime data is thus a reflection of law enforcement's responses to particular kinds of behaviors committed by certain subsets of the population, rather than a reflection of actual criminality" (p. 18). All of these reports are entered into various data systems, which are consolidated and cross-indexed through the Gotham system. Biased policing creates an informational asymmetry where data on people of color predominates. The data itself becomes racialized: with vastly more references and more data points, people of color appear more often in queries. Officers follow up on these leads, feeding even more data into the system. The result is a "racist feedback loop" in which "an increasingly disproportionate amount of police resources are allocated to historically hyper-policed communities" (p. 18). The racist feedback loop suggests a powerful but largely invisible form of inequality, a data-driven discrimination that derives much of its force from its apparent neutrality.

These informational systems continue a racist legacy, but in a strangely flipped fashion. Race is excluded as a classifier: it can no longer be explicitly sorted on or integrated into decision-based procedures. Yet what Gotham, Northpointe, and the other examples show is that, in their intersections with historical stereotypes and social practices, these information architectures retain systemic racism in more subtle ways. Distributed throughout an ostensibly neutral dataset and only enacted when a query is run, this diffused racism means that the systemic advantage accorded to whiteness is retained, while "race" as an obviously problematic classifier is never invoked. Indeed these systems show that race can be reconstituted by a *functional* rather than *ideological* relation: the cluster of people who receive a disproportionately bad decision can be aligned post facto with non-whiteness. Race is imperceptibly performed rather than overtly proclaimed. Race is what is done to someone. For Ta-Nehisi Coates (2018, p. 369), this is the real material difference that underpins the productive fiction of race: "The line dividing black and white America was neither phenotypical, nor cultural, nor even genetic," he writes, "in fact, there was no line at all, no necessary division of any kind…To be black in America was to be plundered. To be white was to benefit from, and at times directly execute, this plunder." The platforms sketched here suggest that the same concept might be applied to the inner world of big-data analytics and algorithmic governance: to be white is to benefit from informational enhancement; to be non-white is to suffer from the same informational mediation.

Drawing a line from Dehomag to modern examples like Palantir thus points to the transformations necessary to preserve existing power relations. While punch-card technologies provided the necessary foundations for indexing and sorting massive populations, contemporary systems maintain existing hegemonies in more subtle ways. Exploring both systems demonstrates the ways in which race is not a

predefined or projected concept, but rather emerges from specific conditions and technical operations. Rather than a static category, race is a flexible device that derives much of its power from its ability to be embedded and enacted in various ways. Both underpinned and amplified by the logics of information technologies, raciology constantly reinvents itself.

References

Agre, P. (1994). Surveillance and Capture: Two Models of Privacy. *The Information Society*, *10*(2), 101–127.

Alden, W. (2016, May 7). *Inside Palantir, Silicon Valley's Most Secretive Company*. BuzzFeed. https://www.buzzfeed.com/williamalden/inside-palantir-silicon-valleys-most-secretive-company

Aly, G., & Roth, K. H. (2004). *The Nazi Census: Identification and Control in the Third Reich*. Temple University Press.

Amoore, L. (2014). Security and the Claim to Privacy. *International Political Sociology*, *8*(1), 108–112. https://doi.org/10.1111/ips.12044

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). *Machine Bias*. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Arendt, H. (2012). *Antisemitism: Part one of the origins of totalitarianism*. Harcourt Brace & Company. http://public.eblib.com/choice/publicfullrecord.aspx?p=3304619

Aul, W. (1972, November). Herman Hollerith: Data Processing Pioneer. *Think*, 22–24.

Austrian, G. D. (1982). *Herman Hollerith: Forgotten Giant of Information Processing*. Columbia University Press.

Bauman, Z. (2001). *Modernity and the Holocaust*. Cornell University Press.

Beller, J. (2018). *The message is murder: Substrates of computational capital*. Pluto Press.

Black, E. (2001). *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation*. Crown Publishers.

Bowker, G., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. The MIT Press.

Braun, C. von. (2000). Und der Feind ist Fleisch geworden: Der rassistische Antisemitismus. In C. von Braun & L. Heid (Eds.), *Der ewige Judenhass* (pp. 149–213). Philo.

Coates, T.-Nehisi. (2018). *We were eight years in power: An American tragedy*.

Deutsche Hollerith Maschinen Gesellschaft m.b.H. (1934). *Denkschrift zur Einweihung der neuen Arbeitsstätte der Deutschen Hollerith Maschinen Gesellschaft m.b.H. in Berlin-Lichterfelde am 8. Januar 1934*.

Domscheit-Berg, D. (2018, July 12). Wehret den Anfängen. *Frankfurter Rundschau*. https://www.fr.de/fr7/wehret-anfaengen-10966885.html

Fehrenbach, H. (2005). *Race After Hitler: Black Occupation Children in Postwar Germany and America*. Princeton University Press.

Gilroy, P. (2001). *Against race: Imagining political culture beyond the color line*. Harvard University Press.

Guess, T. J. (2006). The Social Construction of Whiteness: Racism by Intent, Racism by Consequence. *Critical Sociology*, *32*(4), 649–673. https://doi.org/10.1163/156916306779155199

Harney, S., & Moten, F. (2013). *The Undercommons: Fugitive Planning and Black Study*. Minor Compositions.

Heide, L. (2002). IBM and the Holocaust (Review). *IEEE Annals of the History of Computing*, *24*(2), 94–94.

Hilberg, R. (1985). *The destruction of the European Jews*. Holmes & Meier.

Hoeffding, O., Selwyn, J., Stella, A. J., & Noton, C. H. (1945). *Symposium of Interrogations and Reports on German Methods of Statistical Reporting* (No. 273; 28). British Intelligence Objectives Sub-Committee. http://www.cdvandt.org/BIOS-273.pdf

Hollerith, H. (1889). *Art of Compiling Statistics* (Patent No. US395782 A). http://www.google.com/patents/US395782

IBM Germany. (1983). News From Yesterday. *IBM Nachrichten*, *265*(33).

Kranebitter, A. (2012). *Zahlen als Zeugen: Quantitative Analysen zur „Häftlingsgesellschaft" des KZ Mauthausen-Gusen* [Magister der Sozial- und Wirtschaftswissenschaften]. Universität Wien.

Kühl, S. (2002). *The Nazi Connection: Eugenics, American Racism, and German National Socialism*. Oxford University Press.

Lindemann, A., & Levy, R. (Eds.). (2006). *Antisemitism: A History*. Oxford University Press.

Luebke, D. M., & Milton, S. (1994). Locating the Victim: An Overview of Census-Taking, Tabulation Technology, and Persecution in Nazi Germany. *IEEE Annals of the History of Computing*, *16*(3), 25–39.

Luhmann, N. (2006). System as Difference. *Organization*, *13*(1), 37–57.

Martin, R. (2003). *The organizational complex: Architecture, media, and corporate space*. MIT Press.

Munn, L. (2017). Seeing With Software: Palantir and the regulation of life. *Studies In Control Societies*, *2*(1). https://studiesin-controlsocieties.org/seeing-with-software/

Noble, S. (2018). *Algorithms of oppression how search engines reinforce racism*. New York University Press.

O'Neil, C. (2018). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin Books.

Palantir. (2013, January 25). *Palantir at the Los Angeles Police Department*. https://www.youtube.com/watch?v=aJ-u7yDwC6g

Pugh, E. (1995). *Building IBM: Shaping an Industry and its Technology*. MIT Press.

Law for the Restoration of the Professional Civil Service, § 175 (1933). https://pages.uoregon.edu/dluebke/Holocaust444-544/Arierparagraph.html

Law for the Prevention of Offspring with Hereditary Diseases, (1933). http://germanhistorydocs.ghi-dc.org/pdf/eng/English30.pdf

Erste Verordnung zum Reichsbürgergesetz, §5 (2) d (1935). http://www.verfassungen.de/de/de33-45/reichsbuerger35-v1.htm

Scales-Trent, J. (2001). Racial Purity Laws in the United States and Nazi Germany: The Targeting Process. *Human Rights Quarterly*, *23*(1), 259–307.

Schor, P. (2017). *Counting Americans: How the US Census Classified the Nation*. Oxford University Press.

Speer, A. (1945, May 29). *Interrogation of Albert Speer* (Allied Expeditionary Force, Interviewer) [Interview]. http://hydrastg.library.cornell.edu/fedora/objects/nur:01462/datastreams/pdf/content

Stop LAPD Spying. (2018). *Before the Bullet Hits the Body: Dismantling Predictive Policing in Los Angeles*. Stop LAPD Spying. https://stoplapdspying.org/wp-content/uploads/2018/05/Before-the-Bullet-Hits-the-Body-May-8-2018.pdf

Tooze, J. Adam. (2007). *Statistics and the German State, 1900-1945: The making of modern economic knowledge*. Cambridge Univ. Press.

Weindling, P. (2010). Germany Eugenics and the Wider World: Beyond the Racial State. In A. Bashford & P. Levine (Eds.), *The Oxford Handbook of the History of Eugenics* (pp. 315–331). Oxford University Press.

Whitman, J. Q. (2017). *Hitler's American Model: The United States and the Making of Nazi Race Law*. Princeton University Press.

Zons, R. (2001). Domesday, Buchenwald, Weimar. In G. Theile (Ed.), *Das Archiv der Goethezeit* (pp. 31–43). Ordnung Macht Matrix.