

The Political Theory of Data: Institutions, Algorithms, & Formats in Racial Redlining

Political Theory
2022, Vol. 50(2) 337–361
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00905917211027835
journals.sagepub.com/home/ptx



Colin Koopman¹ 

Abstract

Despite widespread recognition of an emergent politics of data in our midst, we strikingly lack a political theory of data. We readily acknowledge the presence of data across our political lives, but we often do not know how to conceptualize the politics of all those data points—the forms of power they constitute and the kinds of political subjects they implicate. Recent work in numerous academic disciplines is evidence of the first steps toward a political theory of data. This article maps some limits of this emergent literature with an eye to enriching its theoretical range. The literature on data politics, both within political theory and elsewhere, has thus far focused almost exclusively on the *algorithm*. This article locates a further dimension of data politics in the work of *formatting* technology or, more simply, *formats*. Formats are simultaneously conceptual and technical in the ways they define what can even count as data, and by extension who can count as data and how they can count. A focus on formats is of theoretical value because it provides a bridge between work on the conceptual contours of categories and the technology-centric literature on algorithms that tends to ignore the more conceptual dimensions of data technology. The political insight enabled by format theory is shown in the context of an extended interrogation of the politics of racialized redlining.

¹University of Oregon, Eugene, OR, USA

Corresponding Author:

Colin Koopman, University of Oregon, 1295 UO, Susan Campbell Hall, Eugene, OR 97403, USA.

Email: koopman@uoregon.edu

Keywords

Data, formats, algorithms, institutions, power, redlining

Data's Politics

It is increasingly undeniable that data technology is significant for contemporary politics. Although few would deny this, there are perhaps fewer who feel convinced that they have a thorough understanding of the terms on which data has become a political force. We acknowledge the presence of data across so much of our political lives—from mass surveillance to discriminating algorithms to automated disinformation. Yet we often do not know how to conceptualize the politics of all those data points—the forms of power they constitute and the kinds of political subjects they implicate. We lack a political theory of data despite our widespread affirmation of the politics of data.

Consider two contemporary exemplars of data politics. First, a multinational quarantine in response to data-driven forecasts predicting a pandemic's spread. Governments issued restrictive public health measures not on the basis of direct clinical experience, but rather because of predictive modelling developed by small armies of epidemiological statisticians. It is striking how the response to SARS-CoV-2 was driven by forecasting data models that could not have been conceived of, let alone implemented, as real-time models one hundred years ago in the midst of the influenza pandemic of 1918. A second case is that of the infiltration of a democratic election in an advanced liberal nation by a foreign government through the medium of a relatively decentralized communication platform. Where our imaginary of espionage is still trained by Cold War spy thrillers, the Russian-based Internet Research Agency influenced at least one (and quite possibly a second) U.S. presidential election by exploiting the inattention to, or better yet widespread lack of understanding about all the attention to, the informational ecology of social media. As event after event like these two attest, data increasingly drive politics.

Too often what we can clearly see we also find opaque to the understanding. How do we make sense of the politics of data we so readily perceive? What conceptual repertoire would be adequate to data such that we could understand and evaluate their political effects? How, in short, do we theorize the politics of all those data?

The inherited conceptual tools of political theory are often inapplicable to contemporary problems of data politics. Consider, for example, how traditional conceptions of political power, arguably the fulcrum of modern political theory, are increasingly insufficient for a clear articulation and cogent

assessment of data's politics. Modern political theory stretching from Thomas Hobbes (1968 [1651], II.XVII/227, II.XVIII/238) to Max Weber (2004 [1919], 33) has modelled power as an exercise of coercion, exploitation, or domination. This arguably remains the most influential model of power in contemporary political scholarship. Yet the model stands in contrast to contemporary data technologies that produce political relations that are irreducible to the coercion model. Today's social media empires, for example, often appear as sites for a form of power that cannot be adequately comprehended in terms of a model of power as coercion. How a social media platform channels the ways its users can interact with one another is not clearly coercive, and yet it concerns basic issues of who we can be, what we can do, and how we are situated in relation to one another.

One way to think of the kind of shift I here urge is in terms of the work of Michel Foucault, who showed that political power does not always exhibit the form of coercive sovereign power, but sometimes assumes more subtle forms such as disciplinary training (1995 [1975]) and biopolitical management (1990 [1976]). Foucault showed that even institutions as central to modern politics as the prison and the public health agency are sites for the exercise of forms of power that are not simply coercive.¹ Yet for all that, Foucault's own analyses of power do not necessarily help us understand the specificity of the modes of power operative through data, and even where they do help in this task, certainly they do not do so automatically.² Rather, their value for a political theory of data may be in how they point to the need to continually retheorize power as it assumes changed shapes in transforming sociotechnical milieu.

Our conceptual repertoire needs to be enriched where our existing theoretical resources do not help us perceive what is going on right before our eyes, or right beneath our typing fingertips. Many of our inherited conceptual tools remain broadly relevant to contemporary politics. Yet most are increasingly inapplicable to the political orders emerging through our data. We no longer need to be on guard *only* against coercive power, and we ourselves are no longer *only* subjects of dignity whose liberties would be guaranteed by rights, for we are *also* now subjects of data over whom new forms of power are exercised.

These theoretical gaps have been explored in recent work across a spate of academic disciplines, marking the first steps toward a theory of data politics.³ This emergent literature on data politics can be seen as contributing, even if not always consciously, to now-canonized scholarship in the political theory of technology.⁴ What this recent work uniquely contributes is theoretical grip on the specificities of data technologies, which might not be wholly comprehensible through the categories that helped make sense of older

technologies from hammers to railroads to televisions. I here track some of the specific gains of this emergent scholarship on the politics of data technology in order to also map some of its current limits with an eye toward further theoretical expansion.

I proceed as follows. The first section describes a familiar approach in political inquiry focused on the study of *institutions*. I begin at a place outside of technology as commonly understood because my argument benefits from comparison to what functions as a disciplinary baseline for political theory—and if anything is an assumed baseline for a theory of politics it is political institutions. It is undeniable that political institutions are important for understanding data politics. But they are not, I shall argue, sufficient. I turn in the following sections to forms of political inquiry in which the field of focus is expanded to include a survey of the *technological* dimensions of data alongside its role within political institutions. Here it is notable that almost all of the work now attending to technologies of data politics is focused on *algorithmic technology*, or the processing operations internal to computational uses of data. The second section explains how a political theory of the algorithm brings into view the inner technical workings of some of today's most powerful technologies. But the recent spate of attention to algorithmic politics is, I argue, incomplete. Indeed, a focus on the politics of algorithms already suggests the need for inquiries into the politics of other technical elements essential to information systems. I thus argue in the third section for locating crucial operations of the politics of data in the work of *formatting technology*, or what I more simply refer to as *formats* or *data structures*.

The term “format,” as I use it here, refers to the technical-conceptual apparatus that structure data such that they can be recorded, stored, processed, and retrieved. Formats become ineluctably political when they are involved in structuring data about us. Formats are widely visible in our everyday interactions but are also almost always looked over. Formats for gender, race, health, and credit specify the shapes of our data, be it via high-performance, machine-learning systems or legacy paper machines like punch-card indexes. These formats are political not only in the way that they function as political prostheses for classical political dynamics of coercion, but more significantly they are political in the way they serve to perform the work of what I call “fastening” subjects to all manner of databases and systems.

My argument concerns a generalizable political theory that would be relevant for all kinds of data formats. Despite having such generality in view, I shall thread my analysis through a single, and singularly burdensome, case: the role of data in the history of racial redlining in the United States.⁵ If

formats can be shown to have been a central political technology for a project that produced more race-based socioeconomic inequality than any other in the twentieth-century United States, and in the midst of the greatest period of household wealth accumulation in the country's history, then data formats clearly command the attention of political theory. The racial dimensions of redlining make painfully plain why the normative stakes of a political theory of data could not be higher in the face of a growing suite of contemporary data technologies, including predictive policing software, facial recognition surveillance systems, and health-care algorithms.

Focusing on the politics of data through the lens of equality, and more specifically racial equality, helps bring into view a central theoretical contribution of my proposal for a political theory of formats or, more simply, format theory. Format theory can serve as a meeting ground, or focal space, for multiple trajectories in (and beyond) recent political theory that otherwise remain disconnected. Consider my claim above that formats are simultaneously conceptual and technical. If this is right, then format theory could provide a bridge between scholarship on the politics of racial categorization and on technological racialization.

On the one hand, work on racial categories, concepts, and representations has proven crucial for attending to pressing normative questions about identity schemes (Crenshaw 1991; Hall 1981). This literature can be enriched by attention to technical dimensions of categorization that illuminate what otherwise remain purely representational accounts of racial categories. On the other hand, more recent scholarship on racializing technology cultivates a much-needed sensibility for the algorithmic reproduction of racial discrimination (Benjamin 2019; Noble 2018). This literature too can be enriched, here by a fuller attention to how racial conceptualizations are implemented not by algorithms alone but more so by categorization schemas and data structures.

It is not my claim that these two literatures are incomplete on their own. My argument is rather that format theory provides a bridge between conceptual and technical analyses of racial politics (and other topics for political inquiry) that enables us to simultaneously mobilize and enrich the insights of each. Format theory creates analytic capacities for interrogating the conceptual and the technical in their connections with one another. In doing so, the attention to a politics of formats offers a way of building out connections already being made in recent political scholarship—for instance, in work on the statisticalization of racialized criminality (Muhammad 2010), the racialization of censuses (Thompson 2016), and the surveillance of blackness (Browne 2015).

From Institutions to Technologies

Institutional forms like legislatures and elections, along with their standard features of policy and law, remain the preeminent referent for politics as a field of inquiry. This is because institutions clearly matter to politics. They contribute to the structuring and restructuring of our political relations (Berk and Galvan 2009; Jabko and Sheingate 2018). The very idea of the politics of data would thus seem to refer in the first place to data's impacts on formal political institutions, as explored by Cass Sunstein (2007) and others, extending therefrom outward to more nonformal institutional forms, such as social movements, interest groups, and political culture.

Despite the clear value of scholarship on the effects of data on political institutions, it also has its limits. Much of it tends to treat data technology as only derivatively political—that is, as political only insofar as data figures as a prosthetic tool of the exercise of political power in traditional institutions. But there are cases for which we need a theoretical perspective that interrogates technology itself as a political terrain or a political medium. Deepening opinion polarization on social media indeed has profound effects on democratic elections. But social media are also political in the way that they shape how we see ourselves and one another irrespective of participation in political institutions. Facebook might help turn a voting district to extremist paranoia. But at the same time it invites hundreds of millions of users globally to understand their own lives (and those of others) in terms of brief and episodic status updates. We need an expanded theoretical repertoire capable of comprehending the politics that inhere in both kinds of dynamics. Consider why this kind of expansion is called for with respect to my focal case of the politics of redlining.

Within normative political theory, discussions of redlining have thus far attended focused almost exclusively to institutional dynamics. That focus is crucial. But that it has been the exclusive site of attention would be problematically limiting if it were the case that redlining was a data-driven project whose technology reverberated politically both within institutional contexts but also far beyond them. Consider how widespread an exclusive focus on institutions is across normative theorists arguing for quite different conceptions of racial justice. Elizabeth Anderson's influential *The Imperative of Integration* (2010) develops an argument for racial integration as a requirement of justice. This argument explicitly depends on first establishing that "[t]he state's role in constructing segregation has been large" (2010, 68). In describing the "state's role" in redlining, Anderson focuses on how the actions of the Federal Housing Administration (FHA) "promoted racial redlining" by "denying mortgage guarantees in black and integrated neighborhoods" (2010,

68). In contrast to Anderson, Jonathan Kaplan and Andrew Valls argue for the very different demand of reparations for redlining, and yet on the same basis. They note that “the role of the Federal Government in lending was by far the most important factor in creating and solidifying racially segregated housing,” explicitly citing both FHA and the Home Owners’ Loan Corporation (HOLC) (Kaplan and Valls 2007, 260). This institutionalist perspective also frames recent work that takes up the politics of homeownership from the perspective of the civil society organizations and citizens’ groups who resisted it, as detailed in Chloe Thurston’s (2018) work on the role of hybrid “boundary groups” in the history of redlining. And indeed the most influential recent argument about redlining also exhibits the same institutionalist focus. Ta-Nehisi Coates argues that “the federal government—through housing policies—engineered the wealth gap” (Coates 2014, §IX, online). The normative deficiency of unjust inequality is, for all of these commentators, above all a function of political institutions.

It is no surprise that political theorists who have analyzed the politics of redlining have focused exclusively on institutional matters of administrative agencies, statutes, rulings, and policy activism.⁶ For this approach provides an argumentative route to unambiguous normative judgment. Ascribing moral and political culpability to the state yields a powerful argument, because the state, even if only indirectly, impugns us all in the injustice of redlining.

Notwithstanding its clear argumentative force, such an approach neglects the complexity of political dynamics that involve more than just, and play out elsewhere than within, political institutions. Was redlining politically more extensive in this way? To begin to see how it was, consider that it could not have been purely accidental that the data-intensive technologies that implemented redlining were bootstrapped into successful operation just at that moment when the coercive brutality of the lynch mob became an embarrassment. A data-driven informatics of race helped launch a more understated racism that dispensed with the sporadic and overt violence of the noose in order to take up the seemingly benign tools of the clipboard, the derivative equation, and the racializing format. Hence could redlining be perpetrated in full daylight by professionalized administrators at the FHA and HOLC. As part of racism’s long migration from the plantation whip to the manicured database, technologies were developed that enabled the starkly casual production of what Ruha Benjamin calls “default discrimination” (2019, 49, 77).

Such an insidious politics of racial inequality will remain obscured from view so long as we interrogate projects like redlining only with respect to their institutional manifestations. This point can be put in terms of the

familiar contrast between structural and attitudinal racism. If racism is always at least partly structural in that it operates beyond the conscious intentions of some of its perpetrators, then it is incumbent upon us to investigate those structures. As structural racism migrated across the twentieth century, one of its primary sites became technology, specifically information technology. As such, resisting and dismantling racism today would require inquiry into the specifically technological dimensions of structural racism, including the study of what I call “the informatics of race.” To interrogate twentieth- and twenty-first-century structural racism as purely institutional, and not at all technological and informational, is to fail to interrogate a significant swath of how racism has operated for at least a century.

Algorithmic Technology

Although political scholarship in our so-called “information age” remains largely focused on institutions, movements, and interests, there is a growing body of political research that attends to the specific technological operations internal to information systems in order to excavate the politics internal to technologies. This work is welcome, and indeed necessary, for a deeply technological society such as ours.

Interestingly, however, almost all of our contemporary technology-focused contributions to the political theory of data are solely concentrated on a single technology: the algorithm. There is now an impressive roster of work on the politics of algorithms from perspectives in science and technology studies (Gillespie 2014), communications (Bucher 2018), literary studies (Finn 2017), geography (Amoore 2020), law (Rouvroy 2016), and political science (Eubanks 2018). There has also been a small surge of critical analyses of algorithmic politics within political theory, including work by DuBrin and Gorham (2021), Panagia (2021), Sheehey (2019), and my own earlier contributions in Koopman (2019, 66–107 [chap. 2]).

To take one recent instance of the political theory of algorithms, Davide Panagia’s work brings into focus the way in which algorithms exercise political power by disposing us as subjects of data.⁷ The algorithm is politically salient, argues Panagia, precisely because of its technical capacity “to predict future outcomes and to coordinate action” such that we become disposed to be predictable and coordinated (Panagia 2017). Algorithms, on Panagia’s analysis, are effective technologies for producing a political order of things oriented toward predictability (2021, 126). As Bernard Harcourt (2007, 16) and Bonnie Sheehey (2019) have similarly shown in work on predictive policing algorithms, the temporality of prediction is one in which the present is rushed toward its own anticipated future.

The crucial point here is that we who equip ourselves with algorithms dispose ourselves to become what the algorithms say about us: “an algorithm is a dispositif not because it constrains freedom through various forms of domination, but because it proliferates controls on variability and, in this way, governs the movement of bodies and energies” (Panagia 2021, 128). Such algorithmic disposition can be understood as, to borrow a straightforward idea from Ian Hacking (1995), a “looping effect.”⁸ Just as I cultivate a special relationship to my anxiety when a psychologist labels me anxious, I (and others) cultivate a closeness to those labels algorithms ascribe to me: financially trustworthy (credit reporting), potential perpetrator (predictive policing), or unlikely to succeed in graduate school (standardized testing).

Algorithms are portable technologies that travel across institutional domains, as shown by Louise Amoore’s work on algorithms that bridge between finance and the military: “At the level of the algorithm, there is a profound indifference to the context of whether these norms and anomalies pertain to financial trades or the movement of insurgent forces” (2020, 43; cf. Amoore, 2013, 39–45). The mobility of algorithms across social fields suggests that their political power is not only a function of their institutions. The sorts of dispositions, or forms of subjectivation, instilled in us by the algorithm involve political effects extending well beyond institutions.

Turning again to my exemplary case of racialized redlining, a political theory of the algorithm enables us to scrutinize two crucial political features of redlining. First, it offers a view of how race was factored into home real estate appraisal algorithms. Consider FHA’s 1938 edition of its *Underwriting Manual*, the appendix to which contains a reprint of Form 2015, “Report of Valuator.” This form provides a simple (arithmetical) algorithm for calculating home values. One of the variables factored into the equation falls under the heading of “Rating of Location” and concerns “Protection from Adverse Influences.” One subcategory of this is “Quality of Neighborhood Development,” by which is meant: “Areas surrounding a location are investigated to determine whether incompatible racial and social groups are present, for the purposes of making a prediction regarding the probability of the location being invaded by such groups” (FHA 1938, §937; cf. FHA 1936, II. §233). An arithmetical weighting of such a factor within an appraisal algorithm shows not only that inequality was generated by valuation techniques but also precisely how this occurs. These algorithms were imported into 1930s FHA procedures from private-industry real estate appraisal protocols developed in the 1920s. Exemplary is F. M. Babcock’s 1924 *The Appraisal of Real Estate*, proposing a valuation model for which “residential values are affected by racial and religious factors” in ways that can be quantitatively

computed (1924, 70). Just over ten years later, Babcock was chief underwriter for the FHA.⁹

A second crucial feature of redlining illuminated by a political theory of the algorithm concerns the effects of racialized algorithmic apparatus. Algorithms frequently feed into predictive systems such as appraisal reports in a way that unleashes what sociologist Rose Helper, in one of the first critical surveys of redlining, detailed as a self-fulfilling prophecy (1969, 94). Not only do race-focused appraisal algorithms create the possibility for unintentionally race-unequal mortgage underwriting, but they also predictively create an impression that lender decisions track the housing market itself, an impression that appears bolstered by a market that tends to conform to what professionals predict of it. The outputs of valuation algorithms loop back into the housing market in the form of buyers, sellers, and real estate agents who are increasingly disposed to factor algorithmic appraisals into their own subjective valuations.

By illuminating such two-step processes of algorithmic prediction, the political theory of the algorithm offers a way of attending to the politics of data technology itself. This work thus enriches political analyses focused on the actions taken at traditional political fora such as institutions. But more important is that the political theory of the algorithm enables us to see how the informatics of race can at times function as an exercise of political power outside of the explicit purview of institutions, and therefore as political in its own right. Racialized algorithms are political not only in functioning as prosthetic tools for gargantuan institutions like FHA, but they can also be understood as enacting series of disposing looping effects that constitute the terms of our political relations to one another and even to ourselves.

These and other insights afforded by a political theory of the algorithm should be seen as a specific instance of a more general insight we can gain from the political theory of data technology. For, as I shall now argue, the algorithm is just one technological species within the wider genus of data technology (which, in turn, is of course itself a species within the wider genus of technology). The kinds of insights gained by the analysis of algorithms may also be brought into view by considering other specific aspects of data technologies. I turn now to the case for a political theory of the format.

Formatting Technology

A widely circulated computer science formulation from a 1976 textbook instructs readers that *Algorithms + Data Structures = Programs* (Wirth 1976). Insofar as this conceptualization is adequate to contemporary data systems, formats (or data structures) are the missing piece of the equation in

recent scholarship on the politics of data, focused as this scholarship is entirely on algorithms (or processing).¹⁰ For example, communications theorist Taina Bucher cites this canonized equation in *If. . . Then: Algorithmic Power and Politics* (2018, 22). But she then immediately notes that “the focus in this book will almost exclusively be on algorithms” (2018, 23).¹¹ Why adopt a narrowing strategy for a complicated technology? An analytics of algorithmic power is necessary, but attending exclusively to the algorithm as the only operator of data power is not. In fact, such exclusivity may even be conceptually and computationally inoperable. As Niklaus Wirth, author of the 1976 textbook, states, “the structure and choice of algorithms often strongly depend on the structure of the underlying data,” such that the two are “inseparably intertwined” (xiii).¹² A political theory of data requires an interrogation of both the algorithms and the formats intertwined within the politics of data computation. Formats, just like algorithms, contribute something unique to the exercise of power and politics through data.

As I use the term, a “format” refers not so much (or more precisely not only) to a file type (e.g., .HTML or .PDF), but much more broadly to those technical-conceptual specifications that organize forms, registers, records, dossiers, databases, and files of all kind. Formats define, in a manner simultaneously conceptual and technical, how information is itself constituted. In specifying what data must be, formats delimit and determine the specific shapes that are allowable for any data point. The delimitations specified by formats can be more or less restrictive, but formats in every instance require some shaping such that there can be data rather than an incomputable nothing. The very idea of “raw data” (according to which there could be information without formatting) is nonsensical—as if chaos could be computable or disorder could be data. As Lisa Gitelman (2013, 1) deftly puts it, following Geoff Bowker (2005, 184), “‘Raw data’ is an oxymoron.” In more technical terms: all data are already-structured data.

This insight is crucially important for a political theory of data, but to get there we need to understand its technical underpinnings. Computer science (especially in recent theorizations of machine learning) often misleadingly relies on a distinction between unstructured data and structured information. Yet information theory (on which all digital computation today rests) teaches that data are necessarily structured in at least some way. The more appropriate distinction for computer science is one between relatively less-structured and relatively more-structured data (i.e., between data functionally structured for a purpose or not). This point can be traced to the founding statement of information theory: Claude Shannon’s 1948 paper “A Mathematical Theory of Communication.” Shannon defined the concept of “information” quantitatively as the measure of choice in a communication system (1949

[1948], 3, 18; cf. Weaver 1949, 100). My concept of the format speaks precisely to this technical point in information theory. Consider Shannon's point as neatly exemplified by the drop-down-list format. The possible amount of information in a drop-down is determined by the number of options in the predefined list. If the drop-down is formatted with two options, the amount of information provided by a user of that drop-down is (by Shannon's measure) less than that given by a user of a drop-down with three or thirty options. Quantity of information equals measure of choice. My point about formats, then, is only this: formats specify the minimal structuring whereby choice—that is, information itself—is made possible.

Why do such technical details of information theory matter to the politics of data? Consider a specific instance of a drop-down list that is ubiquitous across the online forms we regularly fill out: the one labelled "Gender" on a user-profile settings form. Clicking on that label populates a list of options (often two, sometimes more). Such a drop-down enacts a very precise formatting of a user's information and thereby of the user as a subject of data. Certain options are possible. Others simply do not matter (regardless of whether they matter to the user) because they cannot be formatted by that system as storable data. The form formats what can count as gender information. A user's "Gender" on the form may or may not be their gender, but it is crucial to recognize that for their gender to be any kind of data at all, it must be formatted in some way. Add to this the fact that for most such forms, "Gender" is a required field such that one's gender must be made as data.¹³

Formats of all kinds categorize us in all manner of ways. Some of these categories appear, and in fact are, innocent. Others seem innocent but actually serve as sites for the reproduction of political dangers. Still others cannot even begin to appear as neutral.

If formats at least sometimes bear political weight, then this weight is already internal to the algorithms that depend upon formats, because formats are functional preconditions for algorithms. Formats set the terms according to which any algorithm might operate. Only after information already exists, constituted by a format, can algorithms then do their work. This point too was observed by Wirth in 1976: "the choice of structure for the underlying data profoundly influences the algorithms that perform a given task" (1976, 56). Wirth offers an almost purely-technical example of how the choice of a numeric sorting algorithm depends on whether the data to be sorted are formatted for storage in sequential files on disks or in a computing machine's random access memory (1976, 57). The politics of the algorithm is of enormous importance, but it is also necessarily continuous with the politics of the format, and for specifiable technical reasons.

Why, then, are data formats so visibly neglected in a moment when the politics of algorithms has begun to so brightly emanate? To explain this critical gap, consider that throughout the literature on algorithmic politics the very idea of the algorithm is ambiguous between two possible meanings. In most instances, the term functions as a synecdoche such that “algorithm” stands in for the entirety of “data technology” (or “data science” or “big data”).¹⁴ Insofar as the attention to algorithms is in part an attempt to attend to the technical specificity of data systems, we should be more precise by carefully separating out techniques by which data is formatted from those by which it is processed. Work that respects this distinction moves toward a second meaning of “algorithm.” Here the term refers rather precisely to information “processing” operations—those stepwise procedures coded into programmable machines that calculate over informational inputs. Given this second sense, there is indeed a politics of processing in need of scrutiny. But there is also a politics of data structuring equally commanding critical attention. And since algorithms and formats are designed interoperatively within any data system, the political theory of the algorithm (in its precise sense) already solicits the political theory of the format.

My discussion thus far yields what might be referred to as a dependent argument for the politics of the format. If algorithms are political, and if algorithms depend upon formats, then any politics internal to the algorithm is a politics already reliant upon the formats that constitute data. While this argument motivates an attention to the politics of formats from within the sphere of attention already accorded to the politics of algorithms, there is also a second argument that there is a politics internal to formatting itself independent of downstream uses of data by algorithms, institutions, or anything else.

The independent argument concludes that the very formats that constitute data themselves are political in the same way that guns or nuclear reactors are political: not because they themselves kill or pollute but rather because they are such effective instrumentalities for killing or polluting and so massively useless for doing so many other things (you would not use a pistol as a doorstop, or a nuclear plant as a schoolhouse).¹⁵ A binary drop-down gender list on an account registration form does not automatically affiliate a user with a predefined gender identity, but it is pretty much useless for doing anything other than that. This conclusion now in view, I turn now to sketching the argument for it.

From a technical perspective, formats are what functionally specify how data are defined. This work of specification is politically salient insofar as the data points we accumulate around us are formatted in specific ways such that different formats dispose different subjects of data toward different effects. Consider as an example that some social media platforms allow users to

choose any username they like, whereas others require that users represent themselves through the fixed identifier that is their official legal name. Social media platforms thereby make visible how even our names, which feel so natural and obvious to us, are in fact highly formatted pieces of data technology. You need only think of the remarkable variety of possible names people could have but which fail to fit the formatting requirements of most social media platforms (as well as most legal registration databases): a name containing multiple last names (not the familiar technology of hyphenation but actually having two different last names whose orders can be switched), a name without a last name (be it familial or any other kind of second datum), a name that frequently changes depending on social context, an extremely long name, or a name written with certain special characters. My point is not that the formatting of names is always a coercive harm, as if there is some political good to be had in people being able to have any kind of name at all. Rather, the point is that names are precisely formatted, and their formatting cannot but carry consequences that in some cases help dole out unequal burdens and benefits. You can run a simple algorithm to validate my claim: sum together the hours of work that women put into changing their legal names across dozens of bureaucracies upon marriage (or divorce) and subtract from that number the hours that men put into doing the same upon marriage (or divorce). Recall that my argument locates the politics of data primarily in how formats function to generate and maintain inequalities.¹⁶

The independent argument for countenancing the politics of formats takes on increasing gravity when we observe that formats only rarely operate in isolation, but in almost all instances function within networked ensembles of formats.¹⁷ Name formats are again a good example. The formatting of a name such that it can be stored in a database intersects with innumerable other formats that further database us. In most cases, the name format functions as a technical condition for the storage, collection, processing, and distribution of other individualizing data points. If we grant that there may be an ethics buried within the format of the name itself, then much more palpable is the politics spread across entire ensembles of formats in which the name is a central node.

The ensembles of ground-level formats through which data become constituted are, alongside the algorithms and institutions in which they operate, sociotechnical instrumentalities for the exercise of political power. It is true that formats and algorithms are often implemented in the context of institutions to produce political effects. In many such cases formats are leveraged by bad actors in ways that produce political harms such as unjustified inequalities. But both kinds of data technology can also be used to operate these effects on their own outside of institutional dynamics. That is, they can also

be operationalized in ways that produce political harms even where nobody intended those harms, and even in cases where those involved explicitly aimed to avoid harm. A format can have a profound impact when burned into the operation of a gadget, a website, or a printed blank. And that impact can exceed its being merely a conduit for the intentions of actors (be those actors institutionally affiliated or not).

Some would say that this means that formats themselves exhibit agency.¹⁸ Such formulations are enticing, but I sidestep the irresolvable debates over the metaphysics of agency these views engender. All that political theory really needs to grasp of the situation is that formats and algorithms produce real politics effects in the world. For example, data technologies can be crucial ingredients in the generation and reproduction of inequalities. If institutions can unequally dole out the benefits and burdens of social resources, then so too can algorithms do exactly that, and so too can the formats that make one illegible to the algorithm or the agency, or perhaps legible only in specific ways such that one cannot but be burdened by unequally distributed effects.

The inestimable media archaeologist Cornelia Vismann speaks to format politics in her book *Files: Law and Media Technology* when she claims that, “File plans give birth to a transcendental order of files prior to all content” (2008 [2000], 142).¹⁹ This is a crucial statement of the role played by formats (a term Vismann indeed used, albeit sparingly) prior to the informational contents they already order by making them possible (2008 [2000], 7). Formats, which are the technical operationalization of file plans within data systems, organize in advance all that a data system can be used for. That cannot but be political in any context where information technology is widely operated on people from the databases into which they are born (birth certificates) to those that mark their passing (death registries).

The mode of political power operated by formats is therefore not just the classical Hobbesian or Weberian power of coercion. As such, the power of formats might be taken to be continuous with Foucauldian models of disciplinary power and regulatory biopower.²⁰ Or it might even be related to conceptualizations of the politics of prediction inherent in modern science.²¹ Such continuities are important, but the politics of formats is not merely one more instance of what has already been theorized. For there is a specificity, in the way that formats imprint us, and dispose us, specifically as the information (or data) through which we have come to live and to act.

The grip that formats exercise involves us in a unique operation of power that I call “fastening.” The work of fastening is the achievement of the double operation of the format: the subject of data is pinned down to the format’s delimitations and definitions at the same time that it hustles us up in how

efficiently we are handled by whatever agency or network we have been formatted to fit. Canalizing and accelerating, entrenching and quickening, the conceptual-technical hybrid that is the format bores deeply into us with remarkable celerity. We are pinned down to the prefab formats of our favored social media platforms, and then, in virtue of this pinning, our communicative interaction is quickened. We are tied down to the racial categorization presented on the census form, which in turn fastens us further when the form is loaded into database upon database and is used to rapidly generate a volley of community analyses.

Though the grip of fastening is tight, and is wound tighter over time, I am not asserting that the work of fastening is either inevitable or incontestable. As is abundantly evident in the case of the gender drop-down noted above, formats are today increasingly sites of contestation and transformability. My point is only that we need an analytical category to understand how such contestations take place against a background of our lives and actions being pinned down and sped up. For only with such a background in place can contestation be taken to be truly political (rather than, say, merely idiosyncratic).

The operation I am calling “fastening” was almost pinpointed by Marshall McLuhan in his observation that “all patterns of personal and political interdependence change with any acceleration of information” (1994 [1964], 199). What McLuhan should have said is that patterns of personal and political interaction accelerate with any change of information. Information, since it is choice, makes precise in a way that always affords an acceleration that tends toward, but will never fully achieve, automation.²²

To more clearly focus the politics of fastening enacted by data formats, I turn now to the crux of the exemplar case I have been discussing: the formatting of racialized data at the heart of redlining. As recounted above, racialized algorithms helped produce the infamous redlining projects at the governmental institutions that steered the residential mortgage market in the middle of the twentieth century. Those racialized algorithms and their institutions were therefore deeply reliant in the first place on a surfeit of racial data. They were, for that very reason, deeply reliant on the formats in virtue of which racial data could even come to be constituted. Formats, in other words, constituted the technical starting points for redlining’s political informatics of race.

What were the specific mechanisms through which the informatics of race were operationalized within the history of redlining? This question can again be answered by surveying the technical manuals describing the protocols through which the racial data of redlining were collected. The valuers producing the appraisal reports that factored in race did not necessarily

themselves survey the racial composition of neighborhood locations, as required by their algorithms. But somebody did. And thus could those valuers consult a map, table, or some other quickening technology for racial data lookup. One technical manual from the era describing the process by which such a ground-level informatics of race was produced is the 1935 *Technique for a Real Property Survey*. This manual was a joint product of the Works Progress Administration and the Federal Housing Administration and described in exquisite detail the process for conducting a property survey. In a section titled “General Procedure for Real Property Survey” is included a reproduction of “Form B” on which all data on individual dwellings were to be collected. Form B collects up to 29 pieces of data about each property, including, in Box N, the “Race of Household,” which can be “1. White; 2. Negro; 3. Other” (FHA/WPA 1935, §1, 14). In the instructions to the enumerators who collected these data, the format specifications for Box N are clearly stated. In their entirety, they are: “Mark in the appropriate box the race of the household, whether, white, negro, or other. If any member of the household, other than servant, is negro or of a race other than white, consider the whole household as belonging to that race” (FHA/WPA, §7, 38). Earlier it is stated that throughout the form “one entry and one only is generally to be made” with the clear implication that all dwelling data will be single-race rather than multi-race (FHA/WPA, §7, 21).

An entire politics is contained in such a format. Every dwelling must be accounted for; every dwelling must be racialized; every dwelling must be single-race. The algorithmic processors and institutional policies built atop such a database of race clearly can do some things but not others—a whole series of efficiencies is already coded into any such processing and policy prior to its creation. The subjects whose lives are shaped by these algorithms and institutions can henceforth be shaped only in some ways but not others. And prior to that shaping, the subjects who are recorded on the form are disposed by the forms in certain ways but not others. They are constituted as data by specific formats, and not others. On Form B, they are accounted for by their race (but not occupation or education); their race is accounted for singularly (and without mixed-race categorization); their household’s race is accounted for by a real-estate equivalent of the one-drop rule (and in a way that overcodes nonwhite races). In all of this is an entire apparatus of fastening, one result of which was the enactment of mass-scale racially inegalitarian subsidies with a rapidity (fastening’s acceleration) and demobilizing effect (fastening’s pinning) of which no racist ideologue of the preceding decades could have even dreamt.

The surveys that composed the racialized data underlying redlining are visible exemplars of the formats that were contemporary with the algorithms that

soon turned survey data into valuation criteria. Such algorithms and formats together were then-new techniques deployed by a whole raft of pioneer data scientists of the 1920s and 1930s. What went on in real estate accounting was occurring at increasing scale in domain after domain: educational assessment (the SAT), psychological evaluation (personality profiling), medical care and insurance (health records), biology (the constitution of genetic code as the information transfer process that is the key to life), and more.²³ An entire data epistemology was taking shape with its concomitant data politics.

The first-generation data science of the roaring 1920s and declining 1930s, as big and systemic as it has since become for us, holds true to a penetrating observation of Bruno Latour's: "I have never followed a science, rich or poor, hard or soft, hot or cold, whose moment of truth was not found on a one- or two-meter-square flat surface that a researcher with pen in hand could carefully inspect" (1999, 53). Latour's late-twentieth-century claim for the indispensability of the desk may have shifted to the computer desktop for many twenty-first-century sciences, but we can also follow his point back in the other direction to consider the early-twentieth-century technologies that helped produce the data science of redlining. What we will find there is only a slightly smaller flat surface: an intake form carted on a clipboard by an enumerator, who at the doorstep filled out the checkboxes, including Box N, to produce that moment's own moment of truth with all of the disastrous consequences that followed from this racial accounting.

The politics of redlining was, to be sure, enacted by the institutions of the state that made decisions about people's lives and operated by way of the algorithms that disposed how they would go about living. But it was always shot through with the formats that established the very data through which, in a prior moment, people were fastened to specific variables that defined them. Once paper data had pinned people down to race, that paper could travel in two ways. It could spiral outward centrifugally in ways that would quicken all kinds of processes concerning those new subjects of data (e.g., mortgage decisions and credit evaluations at far-flung banks). And it could also loop back to them with the deep speed of that centripetal feedback that quickly alters what we take ourselves to be capable of. Fastening was crucial to institutionalized redlining, yet more insidious was how it conveyed an exercise of power that travelled much more widely and deeply, in all kinds of ways that are not all reducible to institutional politics.

Categories, Technologies, & Politics

In a hypertechnological world, technologies are political in manifold ways. Our technologies have not only been instrumentally deployed by political

institutions, but they have also been directly involved in disposing us in ways that contribute to the reordering of political relations. These general observations about the politics of technology apply with equal force to the particular politics enacted by the contemporary data technologies in which we are increasingly becoming ensconced.

In their political work of fastening, data technologies are relatively unique among technologies in that they operate not only technically but also conceptually. Unlike technologies as tiny as hammers or as gargantuan as hydroelectric dams, data technologies are more than mere machines. They are machines that operate through conceptual categories. There is thus a powerful combination of the conceptual and the technical that inheres in whatever is data-driven. That power, in its political sense, comes into view only if we interrogate the politics of data in light of the formats that, alongside algorithms, enable data technology to operate at all. Otherwise we are prone to mistakenly understand data technology as inexorable, inevitable, and automatic.

Absent a more fulsome attention to the politics of formats, data will continue to dispose and divide us in ways with which we will fail to come to terms. It is amply clear that this has taken place with respect to the massive racial inequalities of homeownership in the United States generated across the twentieth century. It is ongoing not only in that domain but also in numerous other high-profile political affairs such as online election interference, electronic pandemic surveillance networks, computational climate change modelling, facial recognition systems in public spaces, and the suddenly rampant propagation of disinformation. Continuing inattention to the specific technological dynamics operating in these and other political contexts cannot but contribute to the perpetuation of existing political realities and all their visible scars. In politics, if not everywhere else, our refusal to look more deeply into what is going on right in front of us either visibly makes things worse or faintly lets them go that way.

Acknowledgments

For comments on this article, I thank Verena Erlenbusch-Anderson, Bonnie Sheehy, editors and reviewers for this journal, and audiences at New York University, the New School for Social Research, Eastern Michigan University, and the University of Oregon.


Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Colin Koopman  <https://orcid.org/0000-0003-0354-7522>

Notes

1. Many have followed Foucault's approach here; for a recent exemplar of work further building out a nuanced theory of power in a Foucauldian vein, see Allen (2008).
2. See footnote 20 below explicating why Foucauldian concepts of biopolitics and anatomopolitics (i.e., discipline) are inadequate for contemporary configurations of data politics in contrast to my elaboration of a theory of infopolitics.
3. In addition to the work across multiple disciplines cited below in the section on algorithms, important contributions toward a political theory of data include Galloway and Thacker (2007), Dean (2010), and Harcourt (2015).
4. See for two canonical contributions Winner (1980) and Feenberg (1991). A fuller mapping of scholarship on the politics of technology, as well as related work in the philosophy of technology, to the emergent literature on the politics of data technology is beyond my scope here. I briefly consider Marshall McLuhan, Bruno Latour, and Cornelia Vismann below, but a fuller survey would certainly also include at least Martin Heidegger, Jacques Ellul, Hannah Arendt, and of course Friedrich Kittler.
5. For general background on redlining see Coates (2014); my analysis of redlining here builds on my prior work in Koopman (2019, 108–150).
6. This is also due at least in part to the fact that all of the literature on first-wave pre-1948 redlining relies on the historical research in Jackson (1985); on the nuances of later post-1968 discriminatory lending by hybrid private-public organizations see Taylor (2019).
7. For Panagia's broader theory of political disposition via Foucault's analytic of the *dispositif*, see Panagia (2019).
8. In drawing on Hacking's epistemic account of looping I depart from Panagia's (2021) bid for an ontological (121), rather than epistemic (112), approach to technological disposition. Panagia's argument against epistemic accounts argues that representational epistemologies cannot account for the effects of algorithms (118). My turn to Hacking invites an antirepresentationalist (because pragmatist) epistemology. This prompts the question of whether Panagia's ontology of algorithms is consistent with, or perhaps even requires (as I would argue), a nonrepresentationalist epistemology of algorithms.
9. On the migration of appraisal algorithms from private-industry textbooks to public-agency guidance, see Thurston (2018, 78–82), Koopman (2019, 138–147), and Light (2011, 486–492).

10. Two books that are exceptional in attending to formats are Tenen (2017) and Vismann (2008 [2000]). Another apparent exception is Sterne (2012), but the focus there is really on formats as standards, as theorized for instance by Kittler (2010 [2002], 37).
11. Consider also Amoore's assertion that "the architecture of the cloud is defined spatially by the relations between algorithms and data" in light of her general neglect of data structures (2020, 33). Contrasting somewhat is Dourish (2016), who cites Wirth (2) and explicitly laments that discussions of "data structures . . . have been less prominent" (8) than work on algorithms within the literature on data and computation, and yet ultimately also focuses almost exclusively on the algorithm.
12. Wirth was also explicit that computer science was overly focused on algorithms to the exclusion of data structures (1976, xii).
13. For one example of an empirical analysis of gender drop-downs that is attentive to the theoretical concerns discussed here, see Bivens and Haimson (2016).
14. Gillespie similarly observes that the term "is typically used as an abbreviation for everything described above, combined: algorithm, model, target goal, data, training data, application, hardware" (2016, 22).
15. This argument builds on a view commonly associated with Winner (1980).
16. Implicit in my view is that formatting is not itself a normative category—formatting in itself is not good or bad, but rather it is various formats that can engender good or bad effects.
17. Thanks to Lisa Gitelman for a provocation leading to this point.
18. See for instance a recent special issue of *Big Data & Society* edited by Kennedy, Poell, and van Dijck (2015).
19. I would however decline Vismann's characterization of this process as transcendental (meaning that it must be assumed but can never be shown)—it is precisely my argument that we can empirically interrogate the operations of formats.
20. I have elsewhere elaborated the contrasts between my theory of infopolitical fastening and Foucault's attention to data-related elements of disciplinary registration, biopolitical statistics, and the risk calculations internal to security apparatus. Central to the contrasts I discern between my concept and Foucault's are the presumptive targets of these various modalities of power. Foucault's disciplinary anatomopolitics described a mode of power directed at individual bodies; see Foucault (1995 [1975]). Though it involved documentation, dossiers, and other such paperwork, these data were always in the service of bodily training. Foucault's regulatory biopolitics and the security apparatus it is continuous with proposes a description of power targeted on populations of living beings; see Foucault (1990 [1976] and 2007 [2004], 56ff., 104ff., and 315), as well as later work by Hacking (1990). Though the statistical elements in securitarian biopolitics clearly involves data, it is not a mode of power that grips (or what I call "fastens") the person as data (or as what I call "the informational person"). As Foucault himself shows, a population of living organisms is neither epistemologically nor politically equivalent to a disciplined body (2007 [2004], 55–67).

And both, I argue, are epistemologically and politically distinct from the informational person who is fastened by countless formats. I further develop these contrasts from Foucault in Koopman (2019, 161–168) and in an at-a-glance table in Koopman (2018, 115).

21. On prediction in my period of focus here, see Ross (1991, 390, 472).
22. On automation, see McLuhan (1994 [1964], 346–359).
23. I have explored some of these other domains in Koopman (2019, 2020) and Critical Genealogies Collaboratory (forthcoming).

Bibliography

- Allen, Amy. 2008. *The Politics of Our Selves*. New York: Columbia University Press.
- Amoore, Louise. 2013. *The Politics of Possibility*. Durham, NC: Duke University Press.
- Amoore, Louise. 2020. *Cloud Ethics*. Durham, NC: Duke University Press.
- Anderson, Elizabeth. 2010. *The Imperative of Integration*. Princeton: Princeton University Press.
- Babcock, Frederick Morrison. 1924. *The Appraisal of Real Estate*. New York: Macmillan, 1924.
- Benjamin, Ruha. 2019. *Race After Technology*. Cambridge: Polity Press.
- Berk, Gerald, and Dennis Galvan. 2009. "How People Experience and Change Institutions: A Field Guide to Creative Syncretism." *Theory and Society* 38, no. 6 (2009): 543–580.
- Bivens, Rena, and Oliver Haimson. 2016. "Baking Gender Into Social Media Design: How Platforms Shape Categories for Users and Advertisers." *Social Media and Society* 2, no. 4 (Oct.): 1–12.
- Bowker, Geoffrey. 2005. *Memory Practices in the Sciences*. Cambridge: MIT Press, 2005.
- Browne, Simone. 2015. *Dark Matters: On the Surveillance of Blackness*. Durham, NC: Duke University Press.
- Bucher, Taina. 2018. *If . . . Then: Algorithmic Power and Politics*. Oxford: Oxford University Press.
- Coates, Ta-Nehisi. 2014. "The Case for Reparations." *The Atlantic*. June 2014. <https://www.theatlantic.com/magazine/archive/2014/06/the-case-for-reparations/361631/>.
- Crenshaw, Kimberlé. 1991. "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review* 43, no. 6 (July): 1241–1299.
- Critical Genealogies Collaboratory (coauthored by C. Koopman, P. Jones, V. Simon, P. Showler, and M. McLevey). Forthcoming. "When Data Drive Health: An Archaeology of Medical Records Technology," *BioSocieties*.
- Dean, Jodi. 2010. *Blog Theory*. Cambridge: Polity Press.
- Dourish, Paul. 2016. "Algorithms and their Others: Algorithmic Culture in Context." *Big Data & Society* 3, no. 2 (Dec.): 1–11.

- DuBrin, Rosie, and Ashley Gorham. 2021. "Algorithmic Interpellation." *Constellations*. <https://doi.org/10.1111/1467-8675.12568>.
- Eubanks, Virginia. 2018. *Automating Inequality*. New York: St. Martin's Press.
- Feenberg, Andrew. 1991. *Critical Theory of Technology*. New York: Oxford University Press (later republished in 2002 under a different title).
- FHA (U.S. Federal Housing Administration). 1936. *Underwriting Manual*. Washington, D.C.: Gov't Printing Office.
- FHA (U.S. Federal Housing Administration). 1938. *Underwriting Manual: Underwriting and Valuation Procedure Under Title II of the National Housing Act*. Washington, D.C.: Gov't Printing Office.
- FHA/WPA (U.S. Federal Housing Administration and U.S. Works Progress Administration). 1935. *Technique for a Real Property Survey: Part I: Survey Procedure*. Washington, D.C.: Gov't Printing Office.
- Finn, Ed. 2017. *What Algorithms Want*. Cambridge: MIT Press.
- Foucault, Michel. (1976) 1990. *The History of Sexuality, Volume 1*. Translated by Robert Hurley. New York: Vintage Books.
- Foucault, Michel. (1975) 1995. *Discipline and Punish*. Translated by Alan Sheridan. New York: Vintage Books.
- Foucault, Michel. (2004) 2007 (lectures originally delivered 1978). *Security, Territory Population: Lectures at the Collège de France, 1977-78*. Translated by Graham Burchell. Chicago: University of Chicago Press.
- Galloway, Alexander, and Eugene Thacker. 2007. *The Exploit*. Minneapolis: University of Minnesota Press.
- Gillespie, Tarleton. 2014. "The Relevance of Algorithms." In *Media Technologies*, edited by Tarleton Gillespie, Pablo Boczkowski, and Kirsten Foot, 167–193. Cambridge: MIT Press.
- Gillespie, Tarleton. 2016. "Algorithm." In *Digital Keywords*, edited by Benjamin Peters, 18–30. Princeton: Princeton University Press.
- Gitelman, Lisa, ed. 2013. *"Raw Data" Is an Oxymoron*. Cambridge: MIT Press.
- Hacking, Ian. 1990. *The Taming of Chance*. Cambridge: Cambridge University Press.
- Hacking, Ian. 1995. "The Looping Effects of Human Kinds." In *Causal Cognition*, edited by Dan Sperber, David Premack, and Ann James Premack, 351–383. Oxford: Oxford University Press.
- Hall, Stuart. 1981. "The Whites of Their Eyes: Racist Ideologies and the Media." In *Silver Linings*, edited by George Bridges, and Rosalind Brunt, 28–52. London: Lawrence & Wishart.
- Harcourt, Bernard. 2007. *Against Prediction*. Chicago: University of Chicago Press.
- Harcourt, Bernard. 2015. *Exposed: Desire and Disobedience in the Digital Age*. Cambridge: Harvard University Press.
- Helper, Rose. 1969. *Racial Policies and Practices of Real Estate Brokers*. Minneapolis: University of Minnesota Press.
- Hobbes, Thomas. (1651) 1968. *Leviathan*. New York: Penguin.
- Jabko, Nicolas, and Adam Sheingate. 2018. "Practices of Dynamic Order." *Perspectives on Politics* 16, no. 2 (2018): 312–327.

- Jackson, Kenneth. 1985. *Crabgrass Frontier*. New York: Oxford University Press.
- Kaplan, Jonathan, and Andrew Valls. 2007. "Housing Discrimination as a Basis for Black Reparations." *Public Affairs Quarterly* 21, no. 3 (2007): 255–273.
- Kennedy, Helen, Thomas Poell, and Jose van Dijck. 2015. Special issue on "Data & Agency." *Big Data & Society*.
- Kittler, Friedrich. (2002) 2010. *Optical Media*. Translated by Anthony Enns. Cambridge: Polity Press.
- Koopman, Colin. 2018. "Infopolitics, Biopolitics, Anatomopolitics: Toward a Genealogy of the Power of Data." *Graduate Faculty Philosophy Journal* 39, no. 1: 103–128.
- Koopman, Colin. 2019. *How We Became Our Data: A Genealogy of the Informational Person*. Chicago: University of Chicago Press.
- Koopman, Colin. 2020. "Coding the Self: The Infopolitics and Biopolitics of Genetic Sciences." *Hastings Center Report* 50, no. 3: S6–S14.
- Latour, Bruno. 1999. *Pandora's Hope*. Cambridge: Harvard University Press.
- Light, Jennifer. 2011. "Discriminating Appraisals: Cartography, Computation, and Access to Federal Mortgage Insurance in the 1930s." *Technology and Culture* 52, no. 3 (July): 485–522.
- McLuhan, Marshall. (1964) 1994. *Understanding Media*. Cambridge: MIT Press.
- Muhammad, Kahlil Gibran. 2010. *The Condemnation of Blackness*. Cambridge: Harvard University Press.
- Noble, Safiya. 2018. *Algorithms of Oppression*. New York: NYU Press.
- Panagia, Davide, with Çağlar Köseoğlu (interviewer). 2017. "#datapolitik: An Interview with Davide Panagia." *Contrivers' Review* (Nov.). <http://www.contrivers.org/articles/40/Davide-Panagia-Caglar-Koseoglu-Datapolitik-Interview-Political-Theory/>
- Panagia, Davide. 2019. "On the Political Ontology of the Dispositif." *Critical Inquiry* 45, no. 3 (Spring): 714–746.
- Panagia, Davide. 2021. "On the Possibilities of a Political Theory of Algorithms." *Political Theory* 49, no. 1: 109–133.
- Ross, Dorothy. 1991. *The Origins of American Social Science*. Cambridge: Cambridge University Press.
- Rouvroy, Antoinette (in conversation with Bernard Stiegler). 2016. "The Digital Regime of Truth: From the Algorithmic Governmentality to a New Rule of Law." *La Deleuziana*, no. 3. http://www.ladeleuziana.org/wp-content/uploads/2016/12/Rouvroy-Stiegler_eng.pdf
- Shannon, Claude. (1948). 1949. "The Mathematical Theory of Communication." As reprinted in Shannon and Warren Weaver, *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
- Sheehy, Bonnie. 2019. "Algorithmic Paranoia: The Temporal Governmentality of Predictive Policing." *Ethics and Information Technology* 21, no. 1 (March): 49–58.
- Sterne, Jonathan. 2012. *MP3: The Meaning of a Format*. Durham, NC: Duke University Press.

- Sunstein, Cass. 2007. *Republic.com 2.0*. Princeton: Princeton University Press.
- Taylor, Keeanga-Yamahtta. 2019. *Race for Profit*. Charlotte, NC: UNC Press.
- Tenen, Dennis. 2017. *Plain Text: The Poetics of Computation*. Stanford: Stanford University Press.
- Thompson, Debra. 2016. *The Schematic State*. Cambridge: Cambridge University Press.
- Thurston, Chloe. 2018. *At the Boundaries of Homeownership*. Cambridge: Cambridge University Press.
- Vismann, Cornelia. (2000) 2008. *Files: Law and Media Technology*. Translated by Geoffrey Winthrop Young. Stanford: Stanford University Press.
- Weaver, Warren. 1949. "Recent Contributions to the Mathematical Theory of Communication." In Claude Shannon and Warren Weaver, *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
- Weber, Max. (1919) 2004. "Politics as a Vocation." In *The Vocation Lectures*, edited by David Owen and Tracy Strong, translated by Rodney Livingstone. Indianapolis: Hackett.
- Winner, Langdon. 1980. "Do Artifacts Have Politics?" *Daedalus* 109, no. 1 (Winter): 121–136.
- Wirth, Niklaus. 1976. *Algorithms + Data Structures = Programs*. Englewood Cliffs, NJ: Prentice-Hall.

Author Biography

Colin Koopman is a professor and the head of philosophy, as well as director of new media & culture, at the University of Oregon. His work has appeared in *Critical Inquiry*, *Constellations*, *Contemporary Political Theory* and elsewhere, including public venues such as *The New York Times*, *Aeon*, and *Public Books*. His most recent book, *How We Became Our Data: A Genealogy of the Informational Person*, tracks the history of the politics of information over the past century of data-driven technological deployments within the American scene. His first two books dealt with methodological questions arising out of pragmatist and genealogical philosophy.