Check for updates

Original Research Article

BIG DATA & SOCIETY

Big Data & Society
January–June 2018: 1–12
© The Author(s) 2018
DOI: 10.1177/2053951718757253
journals.sagepub.com/home/bds
SAGE

Isomorphism through algorithms: Institutional dependencies in the case of Facebook

Robyn Caplan[1] and danah boyd[2]

Abstract

Algorithms and data-driven technologies are increasingly being embraced by a variety of different sectors and institutions. This paper examines how algorithms and data-driven technologies, enacted by an organization like Facebook, can induce similarity across an industry. Using theories from organizational sociology and neoinstitutionalism, this paper traces the bureaucratic roots of Big Data and algorithms to examine the institutional dependencies that emerge and are mediated through data-driven and algorithmic logics. This type of analysis sheds light on how organizational contexts are embedded into algorithms, which can then become embedded within other organizational and individual practices. By investigating technical practices as organizational and bureaucratic, discussions about accountability and decision-making can be reframed.

Keywords

Algorithms, accountability, Facebook, institutional theory, isomorphism, bureaucracy

This article is a part of special theme on Algorithms in Culture. To see a full list of all articles in this special theme, please click here: http://journals.sagepub.com/page/bds/collections/algorithms-in-culture.

Concerns about the impact of data-driven intermediaries on the news media industry have been growing steadily over the last several years (Saurwein et al., 2015). Major social media and information companies like Facebook, Google, and Twitter play a central role in what news and information people consume (Gottfried and Shearer, 2016). The popularity of these systems—and the scale with which they impact both viewership and finances—has forced many news media producers to alter how they produce and disseminate content for their audiences. In short, long-standing news outlets must construct their content with algorithmic and data-centric intermediaries in mind. Furthermore, a whole host of new digital-first outlets such as *BuzzFeed* and *Breitbart* have emerged to capitalize on the way in which this ecosystem is architected.

The news industry has long been interwoven with other industries and institutions—most notably, advertising and government. At various points in history, news media has been reconfigured by shifts in those ecosystems. From the rise of "penny press" to the dynamics of government-driven propaganda, journalism has had to change depending on the broader landscape (Schudson, 1987: 14). More recently, the internet, social media, and algorithmic and data-driven systems have altered many aspects of the news and information landscape. Through the use of algorithms that rely on signals from both the content and interactions of consumers, these technologies help curate what news and information is presented to whom. Furthermore, by providing services that allow everyday people to actively serve as content distributors, their systems

[1]Data & Society Research Institute, USA; School of Communication and Information, Rutgers University, USA
[2]Microsoft Research, USA; Data & Society Research Institute, USA

Corresponding author:
Robyn Caplan, Data & Society Research Institute, 36 West 20th Street, Floor 11, New York, NY 10011, USA.
Email: Robyn@datasociety.net

Creative Commons CC-BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (http://www.creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

help distribute some content more widely than others. Traditional news enterprises, dependent on attention and clicks over digital advertising (often delivered by programmatic advertising networks owned by Facebook and Google), are forced to respond to these shifts. As a result, these platforms have upended the organizational practices of news-producing platforms, altering how both the newsroom and individual journalists operate (Christin, 2014; Petre, 2015).

This paper underscores how efforts to increase accountability within algorithmically-mediated fields need to consider the organizational values and institutionalized mechanisms embedded within algorithms that have been driving organizational change across the news media industry. Part of the challenge for algorithmic accountability work is to understand how algorithms and data-driven technologies are both situated within larger macro-social trends, such as the increased privatization of public services in the current era of capitalism as well as changes in ownership structures of industries, and also influence a wide-range of actors and organizations that have become dependent on algorithmic and data-driven intermediaries. We draw on concepts from institutional theory, such as *isomorphism*, to understand how algorithms structure disparate businesses and aims into an organizational field, leading them to change their goals and adopt new practices (DiMaggio and Powell, 1983: 148). This paper provides an analysis of how the media industry has shifted through its dependence on powerful algorithmic intermediaries, such as Facebook. In doing so, we examine both how technology has shaped media industries as well as how these systems have conferred value and legitimacy to specific individuals and organizations. Using DiMaggio and Powell's (1983) theory of isomorphic change, we highlight how Facebook and its algorithmic and data-driven practices have become an institutionalized organization within this domain, structuring the media system as an organizational field. In this sense, algorithms and data-centric technologies, like bureaucracy, act as a mechanism of *legitimation* in the process of institutionalization, reflecting broader macro-structural social processes, inducing a process of isomorphism, or homogenization, among dependent organizations (DiMaggio and Powell, 1983; Meyer and Rowan, 1977). We argue that algorithmic and data-driven technologies can be ''de-mythified'' and viewed more akin to bureaucratic or administrative mechanisms than intelligent systems.

Viewing algorithmic systems as akin to bureaucratic or centralized administrative instruments can help reorient technology companies that have eluded classifications, back into their regulatory domains. The fact that social media and information technologies are not automatically labeled as part of the news business has led to confusion about their role and responsibilities within the news media ecosystem. Ambiguity has stemmed from the emergence of algorithmic and data-driven platforms, built by people who typically lack domain-specific expertise, entering into wide spectrum of sectors and industries, such as transportation, public health, criminal justice, or media. Often, technology companies position themselves as ''platforms,'' which both serves to highlight their intermediary role and allow them to position themselves as ''neutral'' in ways that would make them more immune from more top-down regulation or from complaints by users within the United States (Gillespie, 2010). Julie Cohen (2016) argues that this has created problems for a regulatory environment developed during industrialism that is dependent on ''well-defined industries'' with specifications for what would ''trigger regulatory oversight'' (p. 4). Information technology companies have come to mediate more and more of everyday life, without a clear understanding of how the incentives or goals of the organizations developing technologies can affect diverse sectors or industries. Despite the ubiquity of digital and information technologies now, the language of technology—in this case *data* and *algorithms*—is often used to make a company's activities distinct from previously regulated institutions. At the same time, the presence of these technologies also serves to homogenize the sector's practices and incentives.

The process of homogenization is not unique to algorithms and data. Twentieth century organizational and neo-institutional scholarship focused on a different set of mechanisms—namely, bureaucracy—that also induced similar changes across sectors and industries. Scholars have previously highlighted this automation of bureaucratic processes through software. For example, James Beninger's (1986) *The Control Revolution* focuses on these dynamics long before the totalizing effects of Big Data and algorithmic intermediation had even begun to take shape. More recently, popular articles in publications like *Slate* and *Real Life Mag* make connections between Big Data and algorithms and their bureaucratic administrative predecessors (Clair, 2017; Elkus, 2015). Given long-standing efforts to introduce accountability into bureaucratic systems, there is good reason to examine sociotechnical organization practices through the lens of bureaucracy to open up ways of rethinking accountability in this environment.

Some who still consider social media platforms as focused on personal experiences rather than news may think an analysis of Facebook's effect on the news media an odd choice for an analogy of algorithmic systems as centralized bureaucratic institutions. Yet, conversations around algorithmic accountability often center on Facebook and Google in ways that reveal

the entanglement of social media, news media, and algorithms (Napoli, 2015). Rooted in sociology and political science, DiMaggio and Powell's (1983) theory of isomorphism provides a novel perspective for debates about algorithms and culture, providing a different vantage point for understanding the relationships between organizations embedded within algorithmic logic.

Data-driven algorithms homogenize

Although algorithmic processes are shaping sociotechnical systems more than ever before, the notion of an ''algorithm'' still lacks analytic stability and coherence (Seaver, 2017). Many scholars and computer scientists take as their starting point the definition provided by Donald Knuth (1968), who argued that the word ''algorithm'' refers to a ''finite set of rules which gives a sequence of operations for solving a specific type of problem,'' (p. 27). Meanwhile, there has been a growing understanding within the burgeoning field of ''algorithm studies'' that algorithms are a powerful ''rationalizing force'' within the network society (Pasquale, 2015: 15). Thus, a number of researchers have been looking to other past methods of ''rationalization'' of societies to understand how to both frame an analysis of algorithms, as well as decrease its importance as *the* object of analysis in algorithm studies. Algorithms that serve to rationalize industries, also work to homogenize or make an industry more uniform and similar. Institutional and neoinstitutional definitions of algorithms examine at the role algorithms have begun to play as mediators of macropolitical processes. Philip Napoli (2014) and Mike Ananny (2016) have both noted the usefulness of neoinstitutional theory for relocating algorithms within these complex social, political, and economic relationships. Both Napoli's argument of ''algorithms-as-institutions'' and Ananny's concept of ''algorithmic assemblages'' seek to understand how algorithms come to mediate supra-organizational processes, and automate them to directly ''structure user behaviors.'' What is needed to expand on both Napoli and Ananny's theories is a way to understand how algorithms, as mechanisms of institutionalization, lead to broader system-wide changes among organizations and individuals structured through algorithms.

In the 1980s, James Beninger warned that the mechanisms of algorithms, which define individuals and actions into discrete categories (inputted as variable types) are underpinned by a belief in the value of such processes of rationalization to organize societies, that ''control can be increased not only by increasing the capability to process information, but also by decreasing the amount of information to be processed'' (1986: 16). In this sense, algorithms that serve to preprocess, categorize, and classify individuals and organizations should be viewed *as extensions* of bureaucratic tools such as forms, that have been associated with the state in the past. This comparison, however, has fallen by the wayside in contemporary studies of algorithms. It is difficult to understand why, as much of the contemporary information technology industry has focused on the digitization of records and practices that were previously done offline. For instance, early software development focused on business applications like Microsoft Office and Lotus 1-2-3, which were designed to enable bureaucratic aims like the collection and storage of records about actors (and relationships between actors), events, and processes. Algorithmic and information systems have in many ways served to re-mediate the record-keeping function and standardization of bureaucratic mechanisms. In the process, however, privately-owned software companies that have undertaken this work have fundamentally transformed business and government, as well as consumer technologies, reconfiguring most sectors into data-driven bureaucracies where algorithms promise efficiency and optimization.

Despite the obvious link between bureaucratic modes of information production and online modes of data production using algorithmic models, there are some key differences in the current information environment which may have obscured this relationship. Firstly, bureaucratic modes of information production and management have often been associated with the state and more centralized and hierarchical information organizational contexts. Bureaucracy, for Weber, is a mechanism used by the state to induce *rationality* within complex political and economic environments. The critiques we make of the characteristics of bureaucracy now in the twenty-first century—that it is too hierarchical, too rigid, and requires workers to be specialized—was seen by Weber as precisely the instruments that could be used by the state to create fairer, more just, and more equitable treatment of citizens within societies (Du Gay, 2000: 2; Green, 2008: 201). The ethos of bureaucracy, according to Du Gay's reading of Weber, was to be ''impersonal, expert, and procedural'' through a commitment and subordination to the bureaucratic hierarchy (Du Gay, 2000: 4; Green, 2008: 201; Weber, 1978: 958ff). Like bureaucracies, algorithms are also often deployed with an expressed interest in limiting the subjectivity of decision-making systems (Beninger, 1986: 15; Du Gay, 2000: 2; Green, 2008: 201). Often this is done as a way to make algorithmic systems appear more objective than their human counterparts, even when humans directly play a role in the creation, training, and deployment of algorithmic systems.

Viewing algorithmic systems as an extension of bureaucratic mechanisms can both serve to temper anxieties about the role algorithms are playing in re-structuring industries, and highlight potential avenues for critique. Algorithms, like bureaucracies in the past, have been positioned as the necessary antidote to subjective decision-making processes within large-scale and complex systems that are coordinating between many individuals, industries, and organizations, simultaneously. Companies that have emerged to digitize records and automate the delivery of information and services to users, using proprietary algorithms, have been able to enter into new industries and spaces under the guise of internet exceptionalism (Wu, 2010). In the process, the bureaucratic mechanisms that came to frustrate so many individuals have been closed down and hidden behind not only the "black box" of algorithms, but a mythology that suggests that the work done by algorithms is fundamentally different from that done by offline administrative mechanisms in the past (Pasquale, 2015). Viewing algorithms in this way also highlights how they can work to organize, homogenize, and synthesize industries, such as the news media industry, through processes of "isomorphism" studied by DiMaggio and Powell (1983), within the context of bureaucracy.

Algorithms as administrative mechanisms

As technology companies have come to support state-based processes like predictive policing and algorithmic sentencing (Brayne et al., 2015; Christin et al., 2015), administrative power has transferred from the state to private enterprise, particularly as the technology industry has advanced (Owen, 2015). Over the last two decades, the buzzwords *disrupt* and *disruption* have, according to Taylor Owen, come to stand in "for a form of libertarianism deeply rooted in the technology sector, a sweeping ideology that goes beyond the precept that technology can engage social problems to the belief that free market technology—entrepreneurialism—should be left unhindered by the state" (Owen, 2015). During waves of hype surrounding automation and "Big Data," the mythology surrounding these technologies implies it is more legitimate than existing institutions, with more accurate claims to objectivity (boyd and Crawford, 2012). Yet, the adoption of particular technological practices and vocabulary often serves more as a signal of "legitimacy" than an attempt to improve productivity and performance of an industry (DiMaggio and Powell, 1983; Meyer and Rowan, 1977). This has been seen in the adoption of algorithmic and data-driven processes across a wide spectrum of sectors and institutions, despite evidence that demonstrates that these processes are subject to similar biases and concerns as previous bureaucratically driven institutions. Though this process of institutional delegitimization can find its roots in many histories, the narrative of technology as that which could *disrupt* existing institutional structures can be traced to the ideologies embraced by many of early proponents of the internet.

This is where the theories of neoinstitutional scholars, such as DiMaggio and Powell (1983) and Meyer and Rowan (1977), are especially useful for analyzing how these industry-wide changes can occur. For Weber, bureaucracy rationalized society by trapping people into a structured order, or an "iron cage." DiMaggio and Powell (1983) revisit the concept of his "iron cage" of bureaucracy, seeking to make sense of the existence of a system that persists and continues to structure social life, despite its removal from the context in which that iron cage emerged (and despite it no longer being an efficient way to structure society). The concept of an "iron cage" is fruitful for considering the impact of algorithms, whose formalized logics often contradict the rhetoric of personalization, choice, and freedom.

Algorithms and Big Data function in a similar way—in a world where surveillance is the norm, merely existing in the world means you are structured into the technologies and systems of data collection, production, and analysis that structure most of social life today. For DiMaggio and Powell, the iron cage of bureaucracy persisted 80 years after Weber was writing, not because bureaucracy increased competition or made the state more efficient or just or equal, but *because* it served as a mechanism of rationalization and structuration. In effect, bureaucracy demanded legibility of every actor or organization that interacts with the state within the terms the state defined.

DiMaggio and Powell (1983) look to these mechanisms of structuration to explain how organizations and individuals become more similar or "homogenous" (p. 147). They use a concept, "isomorphism," to provide a way to understand system-wide changes in an industry, sector, or "organizational field" that forces "one unit in a population to resemble other units that face the same set of environmental conditions" (DiMaggio and Powell, 1983: 149). This concept is helpful for understanding how both the mechanisms and the rhetoric of *algorithms* have become embedded within the structure of social life, despite rampant critiques about their capacity to accurately represent reality, increase efficiency, or remain free from bias.

These theories provide an orientation to examine how organizations become more similar through dependence of one organization on another organization, using predictors such as "the greater the

dependence of an organization on another organization, the more similar it will become to that organization in structure, climate, and behavioral focus'' (DiMaggio and Powell, 1983: 154). DiMaggio and Powell offer three frames or forces with which one can analyze these changes—*coercive, mimetic,* and *normative*—though these forces can hardly be considered as separate. In analyzing news media industry in response to algorithms and Big Data, what becomes relevant is how these frames of analysis can help identify new forms of legitimation within organizations, professionals, and audiences that can work to tease out *what* or *whose* values are being prioritized through the current media ecosystem, that are then structured through algorithms. The increasing dominance of the algorithmic form is thus analogous to the role played by bureaucracy in neo-institutionalist work in sociology, where bureaucracy is as much a carrier of *legitimation*, as it is the outcome of more macro-structural social processes.

Algorithmic isomorphism: Re-orienting the organizational field of news media

The news media industry provides an interesting case study to study the increasing dominance of the algorithmic form at competing levels of organizational contexts (the technology company versus the news media company) and individual practices (engineering versus journalism). Technology companies that produce search engines, social media, aggregators, and recommendation engines are currently operating as intermediaries in spaces where news media content is both produced and distributed. Though platforms have repeatedly tried to distance themselves from traditional classifications of sectors (Napoli and Caplan, 2017), theories of isomorphism can work to re-situate these technologies back into their domains. DiMaggio and Powell's theory of institutional isomorphism provides one mechanism to re-orient these companies into the *organizational field* of media. This lens reveals the disconnect between (1) the values and assumptions being embedded into the technology that shapes the media industry and (2) the (often problematic) values that have dominated journalism historically. This has been made more complicated as platform companies engaged in the prioritization and filtering of news media content, like Facebook and Google, have repeatedly sought to differentiate themselves from traditional media companies (Napoli and Caplan, 2017).

The power and centrality of Facebook's News Feed algorithm, in particular, is important. According to DiMaggio and Powell's predictors of isomorphic change (or more homogeneous and similar), the more dependent an industry becomes on one organization who is exerting a dominant administrative function (in this case, Facebook's News Feed algorithm), the more that organization will be able to exert change on other organizations that rely upon them (DiMaggio and Powell, 1983: 154). By defining and re-defining the concept of *relevance* or ''value'' of information and news media, Facebook increasingly writes the rules, or code, that defines which content succeeds or fails in no small part because Facebook is now playing an outsized role in how people access news content (Gottfried and Shearer, 2016). Furthermore, by altering the economics of journalism through the reconfiguration of attention and advertising, Facebook drives news media organizations to incorporate metrics such as click-rate, likes, and shares.

News media and the news feed

In their analysis of the bureaucratic state, DiMaggio and Powell describe coercive forces that stem from political influence and what they refer to as ''the problem of legitimacy'' (DiMaggio and Powell, 1983: 150). In the context of information intermediaries, Lawrence Lessig (1999) takes a similar approach through his adage ''code is law,'' which emphasizes the degree to which code and hardware serve to structure and regulate certain activities and outcomes within computational systems.

Facebook itself has used its News Feed algorithm, and changes being made to it, to exert powerful coercive pressures on organizations operating within its walls. Evidence that news media organizations are subject to the informal and formal pressures Facebook's platform places upon them can be seen in their relative success following changes to Facebook's News Feed algorithm. Publishers that had early success in News Feed effectively subsumed their own organizational practices to the logic of Facebook's algorithms. Outlets like *BuzzFeed*, *EliteDaily*, and *UpWorthy* were early winners in the Facebook ecosystem, using tactics like engineering headlines to include emotional directives for readers (to click or like and share) and creating digestible and relatable content that could be easily shared among users (Meyer, 2013; Oremus, 2016). In August 2013, Facebook engineer Lars Backstrom authored a ''News Feed FYI'' for the corporate blog, explaining why sites engaging in ''clickbait'' practices were prioritized by Facebook's algorithms. At that time, Facebook's algorithms prioritized stories with many likes and comments, and re-prioritized content to the top of a user's feed that had significant engagement (Backstrom, 2013). The impact on the organizations willing to play by these rules was significant. In October 2013, Facebook announced that referral traffic to media sites from Facebook had grown by over 170%

from the previous year. Publishers that engaged with this system saw even more significant growth, with BuzzFeed reporting referral traffic increases of 855% and Bleacher Report at 1081% (Wong, 2015). It was during this period that many of the dependencies between Facebook and news publishers were strengthened. SimpleReach, a content measurement and distribution company, announced in 2013 that Facebook was driving "more traffic than any other social network," surpassing other social media sites popular at the time, such as Twitter and StumbleUpon (Scottberg, 2013).

Facebook's central position within this emerging organizational field led to repeated changes within the media industry as organizations adapted to Facebook's algorithms, and as Facebook changed its algorithms to adapt to these organizations. In their effort to combat the dominance that news media organizations engaging in 'click-bait' were having over their network, Facebook released another change to the News Feed in late 2013 to identify "high-quality" news content (Meyer, 2013). "High-quality" was defined by Facebook as whether users continued to interact with an article after-the-fact, which meant that some publishers saw older articles begin to re-emerge on the network, with traffic driven to this older content. In August 2014, Facebook released another change to the News Feed to address "Click-Baiting Headlines," further defining their concept of quality news sources. In this version, Facebook used variables like "how long people spend reading an article away from Facebook" as a way to calculate how users determine content that is valuable to them (El-Arini and Tang, 2014). Facebook warned publishers relying on click-baiting headlines, that their referral traffic may decrease. Outlets like Eli Pariser's *UpWorthy* were particularly affected by this change, with a decrease of 46% of referral traffic over two months (McArdle, 2014). Pariser responded to the change by re-evaluating the metrics by which *UpWorthy* calculated its own success, to be more in line with Facebook's own goals. Pariser is quoted as saying he was shifting his organization towards "Facebook's focus on engaged time" (Kafka, 2014). This shifting of goals in response to Facebook is perhaps indicative of the less explicit forms of coercive isomorphism described by DiMaggio and Powell (1983), in which organizations are driven to conform to gain support from the organizations upon which they are now dependent (p. 151).

Over the course of 2015 and 2016, several other changes Facebook made to its News Feed had effects on news media organizations. As status updates and personal sharing among users began to decline over 2015, Facebook began to invest more of their resources in products geared towards news media distribution (Efrati, 2016). This included a new emphasis on "native videos" embedded directly in the News Feed, which was communicated out to news publishers directly by Facebook (Oremus, 2016). It also included the launch of Instant Articles in May 2015, a platform developed exclusively for the hosting of content from recognized news media publishers to reduce load time for users clicking on news media stories (Reckhow, 2015). This new platform was also the first step in a revenue-sharing model between Facebook and news publishers, albeit limited—they offered that publishers could sell ads in their articles and "keep the revenue" or use Facebook's Audience Network, the site's own targeted advertising product, already used by many brands and publishers (e.g. *The Huffington Post)* for audience measurement and targeted ad delivery (D'Onfro, 2016). One report, by *Digiday*, said publishers using Instant Articles saw a drop in referral traffic over the last quarter of 2015, though the author notes that few publishers were willing to speak about declines in traffic on the record (Moses, 2016). During another tweaking of its algorithm in June 2016, Facebook's algorithm re-prioritized friends and family over publishers, and news media organizations again saw significant declines in referral traffic (Mosseri, 2016). This corresponded (though there is no documented causal link) to an uptick in a spread of misinformation (commonly referred to as "fake news") over the Facebook network over the same period (Silverman, 2016).

Though Facebook has in many cases claimed that its algorithms merely neutrally reflect the aggregate activities of users (Zuckerberg, 2016, 2017), the framing and re-framing of the News Feed's prioritization of content challenges this claim. This pattern of changing the algorithm to meet their own organizational incentives also highlights how accountability proposals that focus primarily on gaining access to algorithms or data will fall short, given that changes can be made to the News Feed algorithm quickly and with widespread effects on industry practices. Of course, not all news organizations were as responsive to the changes made to the News Feed algorithm. Understanding the broader contexts through which some companies adapted to changes made to the algorithm, when others did not, would be a worthwhile area of future investigation.

'Innovation' through imitation

Changes stemming from coercive forces, especially when frequent, lead to an environment of uncertainty that prompts dependent organizations to learn from other dependent organizations that have successfully conformed to the structuring mechanisms. This process of "mimesis," or imitating models for success, is

another process DiMaggio and Powell (1983: 151) argue will induce similarity across an organizational field. In this sense, the dominant organization's incentives or goals become embedded across an industry through the borrowing of practices that lead to success over the network. In the case of Facebook, this was seen in the adoption of data-driven metrics and analytics into newsrooms, as well as the growth of a new set of intermediaries that were fed directly by the Facebook API, whose role it was to analyze and communicate Facebook metrics back to publishers.

During the early era of the News Feed, relationships between Facebook and media organizations were far from static or one-directional. Rather, an ecosystem of social media analytics businesses, using the Facebook API, acted as intermediaries between Facebook and the news media industry which was growing dependent on the social media platform to reach audiences. Throughout 2013, a number of tools and products were rolled out for media organizations in order to bring in data or content from Facebook users directly into their newsroom. In September 2013, Facebook rolled out tools for publishers and "media partners" including BuzzFeed, CNN, NBC's Today Show, BSkyB, and Slate to integrate "public posts of real-time activity about any given topic," in the form of Keyword Insights API, and the Public Feed API (Osofsky, 2016). The development of these products, as well as other partnerships, quickly lead to an additional ecosystem of businesses who used Facebook's public API and served as an intermediary between Facebook and media organizations. Facebook lists these "media solutions" partnerships on their site, which includes CrowdTangle, a social media analytics company that was bought by Facebook in November of 2016 (Newton, 2017).

As Facebook and other online intermediaries began to take on a larger role in the distribution of journalism and other news media content, the media industry continued to shift in respond to this algorithmic and data-driven environment. While some content providers, such as *BuzzFeed* and *The Huffington Post*, emerged out of these new algorithmic markets producing content directly for Facebook and other social media networks (Herrman, 2016), other news agencies, including *The New York Times*, had to quickly grapple with how to incorporate metrics and analytics into their newsroom cultures (Sobel Fitts, 2015). These pressures were greater for some organizations who saw falling readership on both their website, and on mobile, as they competed with other digital content producers more able to quickly adapt to the algorithmic and data-driven ecosystem.

The news industry was responding to the impact that new digital technologies were having on their industry, incorporating the organizational incentives of these technologies (how Facebook was structuring value of an article) into the structure of their organizations, as well as into the system of incentives that were being used to drive coverage among journalists. *The New York Times* addressed this issue of "disruption" of the existing media industry by new technology players directly within their report, in a section titled "What is Disruption?" (*The New York Times*, 2014: 16). Featured in the report were media businesses that had adopted metrics and analytics into their coverage (such as BuzzFeed, ESPN, and Quartz) and who were expanding their digital offerings rapidly by using "social search and community-building tools and strategies" (p. 24). Other media organizations sought to similarly adapt their business models, highlighting the role that data-driven and algorithmic processes can take in compelling news media organizations to take on the characteristics of social media platforms—an example being Tribune Publishing rebranding itself as the tech company "tronc" which purports to use machine learning to better serve audience interests (Napoli and Caplan, 2017). As DiMaggio and Powell argue, modeling of one organization's practices by another is "a response to uncertainty" through the borrowing of practices that may enhance legitimacy or possibility for success, or to demonstrate to others that they are working to change their practices to be in-line with those of the dominant organization (DiMaggio and Powell, 1983: 151). To that end, many of the practices that we now associate with "innovation," such as the adoption of Big Data methods by the Tribune, is actually due to an uncertain environment which induces one organization to copy the practices of another.

Influence (breadth) versus reputation (depth)

The third source of isomorphic change described by DiMaggio and Powell occurs at the individual level during processes of professionalization of a workforce, which they refer to as "normative pressures" (p. 152). There are many reports that the work of journalism has changed significantly in response to digital media, however, to what extent this is due to Facebook in particular is not known. At the same time, broader trends—the rise of citizen journalism, journalists adapting data-driven metrics into communicating the value of their work, and the incorporation of computational skills into journalistic work—need to be considered alongside the emergence of Facebook and other data-driven intermediaries to assess how algorithms have changed the journalist-audience relationship (Anderson, 2011). Additional factors for consideration include algorithmic methods of surfacing news content (through the

Trending Topics module on Facebook) and the production of news content through automated methods (Podolny, 2015).

As Anderson (2011) and other scholars (Tandoc, 2014; Tandoc and Ferrucci, 2017; Vu, 2013) have noted, as online metrics and data-driven processes gained an increased status in the news media ecosystem, news media organizations began to adapt, using these metrics to learn more about what their audiences search for online and what topics can drive revenue. As a result, they began to choose their ''subjects solely on these computer-generated metrics'' (Anderson, 2011: 536). Anderson stressed that media companies began to embrace an ''algorithmic understanding of democratic processes,'' viewing the process of news delivery as one in which data and algorithms are used to assess individual user wants and needs, relying on a model of ''individual consumer choice'' that becomes equated with democratic values (Anderson, 2011: 541).

Journalists themselves were reporting pressures within newsrooms to adapt to the new digital ecosystem, as systems of incentives that valued views and clicks began to dominate. Though media companies have always used audience measurement to guide reporting, Petre (2015) argues that the ''tracking capabilities of the internet, as well as the ability to store and parse massive amounts of data, mean that audience metrics have grown far more sophisticated in recent years.'' Analytics companies contribute to the ecosystem of data-driven journalism, impacting the professional practices of journalists within newsrooms. Reports by *Nieman Lab* (2014) have detailed the manner by which journalistic and editorial practices began to shift in response to the emergence of web metrics, leading to the rise of a ''culture of the click,'' with web metrics often used as a management tool, particularly when websites rely on ''traffic-based financial incentives'' (Christin, 2014). Metrics influence journalistic practices even when only editors—and not journalists—have access to the data (Sobel, 2015). Even when boundaries between journalists and metrics exist, journalists still seek out ways to numerically compare their work to others (Petre, 2015).

A focus on numeracy and digital data shapes every aspect of how journalists are expected to do their work. For example, journalists are increasingly encouraged to develop technical skills, become skilled at using social media, and learn to code (Broussard, 2015). The pressure to learn digital skills appears to have affected journalism training as early as the 1990s, with roots stemming even earlier to the digitization of some newspapers in the 1970s (Fahmy, 2008: 24). What is unique to the current ecosystem in which intermediaries play a more central role are calls for journalists to engage in ''journalism-as-a-process'' which stresses to journalists that newsmaking is an ongoing process and a collaborative venture between journalists and audiences. Meyer (2013) found that, because of social media, audiences expected journalists to write stories they could ''relate'' to, mirroring expectations from the personalization of content that occurs over algorithmic and data-driven networks, such as Facebook.

Coercive, mimetic, and normative pressures, influenced through algorithmic prioritization of content determined through metrics, can thus have important implications for the production of content. As social media and search engines have centralized the production and distribution of content, news media content in particular, they have also shifted how publishers and journalists determine not only what is important to cover, but how the value of this coverage is communicated to a wide range of actors communicating value through metrics. These actors include publishers communicating the reach of their stories and site, journalists communicating their value to potential employers, and the communication of value (and cost) to advertisers, using these same services to place and distribute advertisements to consumers. Due to their centrality within this network of interrelated actors, however, power to drastically alter these relationships has become centralized within opaque proprietary companies like Facebook, which can work define, and redefine the rules structuring these relationships with no accountability nor oversight.

Algorithms as administration—Implications for oversight

Arguments both in favor of and against incorporating algorithmic decision-making tend to over-emphasize the role algorithms specifically play in the construction of reality. Instead, algorithms should be viewed more as administrative mechanisms that organize relationships between organizations and individuals. As part of this organizational structure, not only are algorithmic systems working to automate the administrative mechanisms of a dominant organization, but they are also providing a common language or structure that serves a legitimizing function that affects other organizations and individuals within that field. This is both positive and negative for increasing oversight into this process of structuration. Using neoinstitutional theory, it is possible to trace a network of organizational dependencies and relationships that we cannot see through the code alone.

Through the administrative function of algorithms, organizational incentives become deeply embedded within many layers of an organizational field, making the encoded algorithmic models almost invisible and

less amenable to change. Within the area of media policy, tracing this map of interdependencies, in terms of both sociotechnical and economic dependencies, begs us to question how one changes system-wide incentives and organizational structures that become embedded at multiple levels. It also begs the question of *how*, or according to what principles, we begin to assess the constraints, limitations, and goals of the dominant organization as their commitments become embedded into algorithmic or computational mechanisms. As part of this re-definition of values and criteria, neoinstitutional scholars Meyer and Rowan (1977) stress the need to include mechanisms that would not necessarily be subsumed within this dominant formal structure (p. 356). In other words, one approach Facebook could take would be to develop a mechanism by which news content is excluded from the same metric-ization of clicks and likes as other content on Facebook. In this way, Facebook could "decouple" its valuation of news from its organizational logics.

What makes algorithms so seemingly necessary and powerful is the sheer *amount* of data they parse and sort through—far beyond what individuals are capable of processing on their own. And yet, evidence increasingly suggest that algorithms are *no less immune* to the biases and inefficiencies of the humans that created them, within the domains or sectors within which they are situated (Kroll et al., 2017). Consistently, technology remains embedded within the specific historical social, political, and economic contexts—and existing systemic social injustices—in the domains and sectors in which they are deployed, despite their owner's hopes for their *disruptive power* (Owen, 2015).

Understanding algorithms as an extension of the concept of bureaucracy, in terms of both their organizing and legitimizing functions, is one step towards understanding how organizations have become more homogenous in the era of algorithms and data-driven processes. In their analysis about the similar function bureaucracy played centralizing the power of the state, DiMaggio and Powell provide one set of mechanisms—predictors of isomorphic change—that is useful for conceptualizing how organizational contexts become embedded within other dependent organizations, through the administrative processes of data and algorithms. This work is useful for re-orienting discussions about how to assess the values driving the shaping of an organizational field, such as the news media industry.

Where the analogy between algorithms and bureaucracy falls short can also illuminate broader concerns about introducing accountability mechanisms within algorithmic systems. Bureaucracies were implemented by human beings and tend to have humans at most ends of the bureaucratic process. Within these more algorithmically-driven systems, humans are part and parcel of how individuals, behaviors, and content become classified and embedded within algorithmic-systems (e.g. the use of journalists to train the Facebook Trending Topics algorithm (Nunez, 2016)). However, automation and algorithms may take over many processes, leading to fewer points of access for individuals to question or critique how they have been classified into the system (Eubanks, 2018). Additionally, the human element of bureaucratic systems may in some way reduce the complexity of the bureaucratic systems. If humans are responsible for carrying out mechanisms of bureaucracy, to some degree, they must understand how the bureaucracy works. Within algorithmic bureaucracies, code often serves to mediate these relationships, as well as the relationships between data collected through the platform, reducing the number of people capable of understanding the complexities of the whole system, as well as opportunities for critiquing process or whistleblowing.

The news media industry has been irrevocably shifted in the era of data and algorithms. In many ways, these mechanisms brought about positive changes to an industry that has long been hegemonic, top-down, and controlled by powerful corporate interests that were allowed to consolidate in the late 1990s after changes to the Telecommunications Act (Croteau and Hoynes, 2006). While algorithms are only a small part of what is inducing change across the news industry, the Facebook example shows that power rests not in an algorithm's capacity to induce a new *logic* into an industry, but in its function as an administrative mechanism that is embedded with the values and cultural and economic environment of their creators. In this process, algorithms shape all other organizations and individuals operating within a given landscape.

In studying the role data and algorithms have in reproducing the structure of social life, there has been a significant emphasis placed on gaining access to the specific technologies themselves. Efforts to increase the transparency or auditability of algorithms and data rest on the assumption that gaining access to code, or even the data used to train machine learning models, can provide us insight into *what goes wrong* when there are biases or inefficiencies within systems (Diakopoulos, 2014; Sandvig et al., 2014). However, institutional theory shows that, to the extent that one organization is able to control the behaviors, actions, or incentives of other organizations through changing the structure of their system unilaterally, the ripple effects go far beyond the code itself. More research is needed to understand how the interplay between organizations in the news ecosystem are influencing each other and being implemented through code. Increasing oversight or accountability into how an industry has been shaped

by a dominant organization is incredibly complex, and entails a system-wide analysis of how organizational incentives, built into algorithms, both operate as a constraint across a field, and are themselves constrained by larger macro-sociological trends.

Acknowledgments

We thank Philip M Napoli, CW Anderson, Emmanuel Moss, and our anonymous reviewers for their generous feedback on this paper. Participants at three workshops also helped us hone early versions of this document: the Digital Technologies and Democratic Theory workshop at Stanford University in May 2017, the Algorithms in Culture workshop held at UC Berkeley in November 2016, and the Eclectic Workshop held at Data & Society in December 2016. We are deeply grateful for all of their feedback, critique, and recommendations.

Declaration of conflicting interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: in part by a grant from the Open Society Foundations.

References

Ananny M (2016) Toward an ethics of algorithms: Convening, observation, probability, and timelines. *Science, Technology, & Human Values* 4(1): 93–117.

Anderson CW (2011) Deliberative, Agonistic, and Algorithmic Audiences: Journalism's Vision of its Public in an Age of Audience Transparency. International Journal of Communication 5: 529–547.

Backstrom L (2013) News Feed FYI: A window into news feed. Available at: https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed.

Beninger J (1986) *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, MA: Harvard University Press.

boyd D and Crawford K (2012) Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication, & Society* 15(5): 662–679.

Brayne S, Rosenblat S and boyd d (2015) *Predictive Policing*. New York, NY: Data & Society. Available at: http://www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf (accessed 10 October 2016).

Broussard M (2015) Teaching Coding in Journalism Schools: Considerations for a Secure Technological Infrastructure. *Computation + Journalism*, October 2–3, 2015, New York, NY.

Christin A (2014) *When it Comes to Chasing Clicks, Journalists Say One Thing but Feel Pressure to do Another*. Cambridge, MA: Nieman Lab.

Christin A, Rosenblat A and boyd d (2015) *Courts and predictive algorithms*. Data & Civil Rights: A New Era of Policing and Justice. New York, NY: Data & Society.

Clair A (2017) Rule by nobody: Algorithms update bureaucracy's long-standing strategy for evasion. In: Real Life Mag. Available at: http://reallifemag.com/rule-by-nobody/.

Cohen J (2016) The regulatory state in the information age. *Theoretical Inquiries in Law* 2: 370–414.

Crain M (2009) The rise of private equity media ownership in the United States: A public interest perspective. *International Journal of Communication* 3: 208–239.

Croteau D and Hoynes W (2006) *The Business of Media: Corporate Media and the Public Interest*. Thousand Oaks, CA: Pine Forge Press.

Diakopoulos N (2014) *Algorithmic Accountability Reporting: On the Investigation of Black Boxes*. New York, NY: Columbia Journalism School, Tow Center for Digital Journalism.

DiMaggio J and Powell W (1983) The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review* 48: 147–160.

D'Onfro J (2016) How an 'oddball' team created one of Facebook's biggest threats to Google. Business Insider. Available at: http://www.businessinsider.com/what-is-facebook-audience-network-and-why-does-it-matter-2016-2 (accessed 15 October 2016).

Du Gay P (2000) *In Praise of Bureaucracy*. London: SAGE.

Efrati A (2016) Facebook struggles to stop decline in 'original' sharing. In: The Information. Available at: https://www.theinformation.com/facebook-struggles-to-stop-decline-in-original-sharing (accessed 14 November 2016).

El-Arini K and Tang J (2014) News Feed FYI: Click-baiting. *Facebook Newsroom*. Available at: https://newsroom.fb.com/news/2014/08/news-feed-fyi-click-baiting/ (accessed on 16 October 2016).

Elkus A (2015) You can't handle the (algorithmic) truth. In: Slate / Future Tense. Available at: http://www.slate.com/articles/technology/future_tense/2015/05/algorithms_aren_t_responsible_for_the_cruelties_of_bureaucracy.html (accessed 25 January 2017).

Eubanks V (2018) *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. New York, NY: St. Martin's Press.

Fahmy S (2008) How Online Journalists Rank Importance of News Skills. *Newspaper Research Journal* 29(2): 23–39.

Gillespie T (2010) The politics of 'platforms'. *New Media & Society* 12(3): 347–364.

Gottfried J and Shearer E (2016) *News Use Across Social Media Platforms 2016*. Washington, DC: Pew Research Center. Available at: http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/ (accessed 10 October 2016).

Green JE (2008) Max Weber and the Reinvention of Popular Power. *Max Weber Studies* 8(2): 187–224.

Herrman J (2016) Inside Facebook's (Totally insane, Unintentionally gigantic, Hyperpartisan) Political-Media Machine. The New York Times Magazine. Available at: https://www.nytimes.com/2016/08/28/magazine/inside-

facebooks-totally-insane-unintentionally-gigantic-hyper-partisan-political-media-machine.html (accessed 17 October 2016).

Kafka P (2014) Upworthy's traffic is still headed down. Blame us, not Facebook, says Upworthy. In: Recode. Available at: http://www.recode.net/2014/5/14/11626848/upworthys-traffic-is-still-headed-down-blame-us-not-facebook-says.

Knuth D (1968) *The Art of Computer Programming*. Boston, MA: Addison-Wesley.

Kroll JA, Huey J, Barocas S, et al. (2017) Accountable algorithms. *University of Pennsylvania Law Review* 165: 633–704.

Lessig L (1999) *Code and Other Laws of Cyberspace*. Basic Books.

McArdle M (2014) Facebook puts a downer on upworthy. In: Bloomberg. Available at: https://www.bloomberg.com/view/articles/2014-02-11/facebook-puts-a-downer-on-upworthy (accessed 10 October 2016).

McQuail D (1992) *Media Performance: Mass Communication and the Public Interest*. Thousand Oaks, CA: Sage.

Meyer J and Rowan B (1977) Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology* 83(2): 340–363.

Meyer R (2013) Why are upworthy headlines suddenly everywhere? *The Atlantic*. Available at: http://www.theatlantic.com/technology/archive/2013/12/why-are-upworthy-headlines-suddenly-everywhere/282048/ (accessed 12 October 2016).

Meyer R (2016) Facebook purges journalists, immediately promotes a fake story for 8 hours. *The Atlantic*. Available at: http://www.theatlantic.com/technology/archive/2016/08/facebook-steps-in-it/497915/ (accessed 15 October 2016).

Moses L (2016) Uh-oh, some publishers see a drop in Facebook traffic. In: Digiday. Available at: http://www.digiday.com/publishers/publishers-just-saw-decline-facebook-traffic/.

Mosseri A (2016) News feed FYI: Addressing hoaxes and fake news. In: Facebook Newsroom. Available at: https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/ (accessed 15 January 2017).

Napoli PM (2014) An institutional theory perspective on algorithmic media production and consumption. *Communication Theory* 24(3): 340–360.

Napoli PM (2015) Social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekeeprs. *Telecommunications Policy* 39: 751–760.

Napoli P and Caplan R (2017) When Media Companies Insist They're not Media Companies, Why They're Wrong and Why That Matters. *First Monday* 22(5).

Newton C (2017) Instant recall. In: The Verge. Available at: https://www.theverge.com/2017/4/16/15314210/instant-articles-facebook-future-ads-video (accessed 15 August 2017).

Nunez M (2016) Want to Know What Facebook Really Thinks of Journalists? Here's What Happened When It Hired Some. Gizmodo.com. Available at: https://gizmodo.com/want-to-know-what-facebook-really-thinks-of-journalists-1773916117 (accessed 15 October 2016).

Oremus W (2016) Who controls your Facebook feed. In: Slate. Available at: http://www.slate.com/articles/technology/cover_story/2016_01_how_facebook_s_news_feed_algorithm_works.html (accessed 15 October 2016).

Osofsky J (2016) Information about trending topics. In: Facebook Newsroom. Available at: http://www.newsroom.fb.com/news/2016/05/information-about-trending-topics/ (accessed 17 October 2016).

Owen T (2015) *Disruptive Power: The Crisis of the State in the Digital Age*. Oxford: Oxford University Press.

Pasquale F (2015) *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge: Harvard University Press.

Petre C (2015) The traffic factories: Metrics at Chartbeat, Gawker Media, and The New York Times. In: Tow Center for Digital Journalism. Available at: http://www.towcenter.org/research/traffic-factories (accessed 17 October 2016).

Podolny S (2015) If an algorithm wrote this, how would you even know?. *The New York Times*. Available at: https://www.nytimes.com/2015/03/08/opinion/sunday/if-an-algorithm-wrote-this-how-would-you-even-know.html (accessed 15 November 2016).

Reckhow M (2015) Introducing Instant Articles. *Facebook Media*. Available at: https://media.fb.com/2015/05/12/instantarticles/ (accessed 15 October 2016).

Tandoc EC (2014) Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society* 16(4): 559–575.

Tandoc EC Jr and Ferrucci PR (2017) Giving in or giving up: What makes journalists use audience feedback in their news work? *Computers in Human Behavior* 68: 149–156.

Sandvig C, Hamilton K, Karahalios K, et al. (2014) Auditing algorithms: Research methods for detecting discrimination on internet platforms. In: *Data and discrimination: converting critical concerns into productive inquiry (preconference at international communications association '14)*, 22 May 2014, Seattle.

Saurwein F, Just N and Latzer M (2015) Governance of algorithms: Options and limitations. *info* 17(6): 35–49.

Schudson M (1987) *Discovering the News: A Social History of American Newspapers*. New York, NY: Basic Books.

Scottberg E (2013, October 21) Facebook continues to be the biggest driver of social traffic. In: SimpleReach Blog. Available at: http://blog.simplereach.com/facebook-continues-to-be-the-biggest-driver-of-social-traffic/ (accessed 15 October 2016).

Seaver N (2017) Some tactics for the ethnography of algorithmic systems. *Big Data & Society*. Epub ahead of print 2017. Available at: https://doi.org/10.1177/2053951717738104.

Silverman C (2016) This Analysis shows How Viral Fake Election News Stories Outperformed Real News. BuzzFeed News. Available at: https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm_term=.drE8WWBDW1#.aa17xxwKxa (accessed 15 January 2017).

Sobel Fitts A (2015) When metrics drive newsroom culture. *Columbia Journalism Review*. New York. Available at: https://www.cjr.org/analysis/how_should_metrics_drive_newsroom_culture.php (accessed 17 October 2016).

The New York Times (2014) 2015 Annual report. *The New York Times*.

Vu HT (2013) The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism* 15(8): 1094–1110.

Weber M (1978) *Economy and Society*. Berkeley: University of California Press.

Wong D (2015) In Q4, social media drove 31.24% of overall traffic to sites. Report. In: Shareholic. Available at: https://blog.shareaholic.com/social-media-traffic-trends-01-2015/ (accessed 15 October 2016).

Wu T (2010) Is internet exceptionalism dead? In: Szoka B and Marcus A (eds) *The Next Digital Decade – Essays on the Future of the Internet*. Washington, DC: TechFreedom.

Zuckerberg M (2016) I want to share some thoughts about trending topics. Available at: Facebook/Mark Zuckerberg Public Posts: https://www.facebook.com/zuck/posts/10102830259184701?pnref=story (accessed 10 October 2016).

Zuckerberg M (2017) Building global community. In: Facebook.com. Available at: https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634 (accessed 14 July 2017).