5-2013

The Algorithm as Institution: Toward a Theoretical Framework for Automated Media Production and Consumption

Philip M. Napoli

Fordham University, pnapoli@fordham.edu

The Algorithm as Institution:

Toward a Theoretical Framework for Automated Media Production and Consumption

Philip M. Napoli

*Professor, Communication and Media Managemen*t

Schools of Business

Director, Donald McGannon Communication Research Center

Fordham University

Media Policy Fellow

New America Foundation

Abstract

Media scholars have only recently begun to recognize and investigate the importance of algorithms to a wide range of processes related to the production and consumption of media content. There have been few efforts thus far, though, to connect these developments to potentially relevant bodies of existing theory and research. This paper seeks to address this gap by exploring the utility of institutional theory as a potentially useful analytical framework for continued inquiry into the role of algorithms in the operation of media systems; and by offering some suggestions for ways in which an institutional analytical frame can be extended into algorithmic contexts.

The Algorithm as Institution:
Toward a Theoretical Framework for Automated Media Production and Consumption

Introduction

Algorithms are widely recognized as playing an increasingly influential role in the political, economic, and cultural spheres (Mayer-Schonberger, & Cukier, 2013; Pariser, 2011; Steiner, 2012). Algorithms are serving particularly prominent roles in the media sector, where the processes of media production, consumption, and even advertising placement, are increasingly automated and algorithmically dictated (see, e.g., Danaher, Lee, & Laoucine, 2010; Mager, 2012; Steiner, 2012). Clearly, then, the *algorithmic turn* (to borrow Uricchio's [2011] phrase) that is taking place in the media sector should be a focal point for communication and media studies scholarship. Researchers have begun to examine this transition in a number of contexts and from a variety of analytical perspectives (see, e.g., Beer, 2001; Gillespie, 2011; Webster, 2011).

But, as is to be expected in these early stages of an emergent area of inquiry, there has been relatively little discussion of useful theoretical frameworks (for exceptions, see Anderson, in press; Webster, 2011). This paper attempts to address this gap via an exploration of how institutional theory can inform and guide future research on the algorithmic turn in media production and consumption, as well as provide a lens through which to interpret extant research in this area.

Institutional theory has frequently and fruitfully been employed in the study of media organizations and practices (see, e.g., Cook, 2005; Napoli, 2011; Schudson, 2002; Sparrow, 1999) and has, in a few instances, been applied to new media organizations and platforms in which algorithms play a prominent role (see, e.g., Anderson, in press; Webster, 2010, 2011).

The goal here, however, is to more explicitly link institutions (and institutional theory) and algorithms. Toward this end, the contentions of this paper are: 1) that algorithms can be conceptualized as a distinct form of institution; and 2) that institutional theory can be useful for understanding the functioning and impact of algorithms in media production and consumption.

This paper begins with overview of institutional theory and how it has been applied to the study of media institutions. Drawing upon this body of work, the next section argues that algorithms fit comfortably within established definitional frameworks for institutions and the functions that they serve. This paper then considers the implications and opportunities that arise from an institutional approach to the study of algorithms, through an exploration of how specific lines of institutional theory can inform what we know at this point about the role that algorithms are playing in the dynamics of media production and consumption. The concluding section considers how this theoretical perspective on algorithms illuminates directions for future research.

Institutional Theory and the Study of Media Institutions

It is important to note at the outset that institutional theory constitutes a very broad tent, encompassing everything from economics-driven agency theory (see, e.g., Alchian & Demsetz, 1972) to political science-grounded rational choice theory (see, e.g., Moe, 1990), to more sociologically-oriented theoretical approaches such as social constructivism (see, e.g., Berger & Luckmann, 1966) and rationalization (for a more detailed overview, see Scott, 2008).

Given the theoretical and disciplinary breadth that characterizes the field, it is perhaps not surprising that institutional theory and research has long been characterized as possessing a high degree of definitional ambiguity and interpretive inconsistency (DiMaggio & Powell, 1991a; Scott, 2008). The goal here is not to delineate the many contours of the field, or to engage in the debates about the strengths and weaknesses of various theoretical perspectives. Rather, the goal

here is to provide a brief grounding in the definitional parameters and core objectives of institutional theory and research. Subsequent sections of this paper will delve into specific branches of institutional theory more deeply, with a focus on identifying and exploring those that seem to provide the greatest potential for enhancing our understanding of the institutional role and function of algorithms in media production and consumption.

One of the core definitional inconsistencies that has characterized institutional research involves whether institutions are conceptualized in very concrete terms as formal, complex organizations, or more abstractly as formal or informal routines, norms, rules, or behavioral guidelines (Jepperson, 1991). Of course, these conceptualizations need not be mutually exclusive, and can become intertwined in a variety of interesting ways, particular within the context of the study of media institutions. For instance, the institution of journalism in many ways resides at the intersection of complex and evolving formal organizations (i.e., news outlets) and equally complex and evolving norms and procedures related to the professional practice of journalism.

Institutions can be broken down into three broad components: regulative, normative, and cultural-cognitive (Scott, 2008). The regulative dimension refers to the ways in which institutions "constrain and regularize behavior" (Scott, 2008, p. 52). This dimension entails a focus on elements such as regulatory processes, rule-setting, and sanctioning activities. The normative dimension refers to the role of social values and norms and how they contribute to the definition of goals and objectives, as well as the appropriate means of pursuing them (Scott, 2008). This dimension entails a focus on elements such as common beliefs and values within organizations and communities. The cultural-cognitive dimension refers to shared interpretive frames and conceptions of reality. This dimension entails a focus on the mechanisms via which

shared meaning and knowledge are created and disseminated (Scott, 2008). As will become clear (see below), all three of these dimensions, to varying degrees, resonate within the role and function that algorithms serve in media production and consumption

Through the examination of these various elements, institutional theory seeks to explain phenomena such as commonalities in the structure and behavior of organizations; the role of conventions, routines, and habits in individual and organizational behavior, and how those reflect or deviate from the pursuit of rational interests; and the construction and evolution of laws, rules, interests, and environmental cognitions (see Scott, 2008).

Institutional Theory in the Media Sector

Institutional approaches to the media have a long history (for reviews, see Moe & Syvertsen, 2007; Schudson, 2002), though it is worth noting that a large proportion of the scholarship that has examined media institutions (i.e., media organizations and/or established norms, procedures, and practices in the media sector) has not employed institutional theory per se. The overwhelming majority of media institutions scholarship that has employed institutional theory has focused on the news media (see, e.g., Benson, 2006; Cook, 2005, 2006; Lowrey, 2011; Lowrey & Wu, 2010; Napoli, 1997; Schudson, 2002). However, reflecting the broad scope of institutional research outlined above, media and communication-focused scholarship has also examined realms such as regulation and policymaking (Galperin, 2004), cultural production (Ahlkvist, 2001; Guzman, 2005; Kim, 2012), technology development (Flanagin, Flanagin, & Glanagin, 2010; Hrynyshyn, 2008), and the construction of audiences and publics (Herbst, 1993; Napoli, 2003, 2011).

Much of this work has focused on processes of "gatekeeping," via which decisions about which content to disseminate to the public are reached (Moe & Syvertsen, 2007). More broadly,

underlying much of this work is a concern with how organizational and supra-organizational forces affect media organization and industry structures, behavioral patterns, environmental cognitions, and (perhaps most significantly), content. In many ways, it is this need to understand the institutional forces that affect content outputs and flows that is the driving force behind this body of research, given the social, political, and cultural impacts of various forms of media content. From this standpoint, a key dimension of media institutions research to date is the extent to which it has compellingly illustrated that the media function as a political, cultural, and social institution (Sparrow, 2006) – an institutional breadth that, as will become clear, is equally applicable to the role and function of algorithms.

<div align="center">The Algorithm as Media Institution</div>

The argument that this section seeks to develop is that, when we consider algorithms through the lens of institutional theory, the functionalities and effects of algorithms map quite closely with those of institutions in general, and media institutions in particular. In developing this argument, it is first necessary to provide a framework for conceptualizing algorithms, which represent a distinctive integration of human and technological agency, within an institutionalist theoretical context.

In many ways, this argument can be thought of as a companion to Katzenbach's (2011) argument that media technologies should be thought of as an institution. As he illustrates, media technologies have a regulatory dimension (constraining and facilitating communicative behaviors and preferences) that is a key characteristic of institutional structures. Media technologies are able, through the characteristics of their design, to both constrain and facilitative communicative practices and preferences, and thus essentially provide base structures and parameters that regulate the production, distribution, and consumption of content. But of course,

the development of these technologies emerges from – and is shaped by social processes, thereby reflecting a duality that is in many ways a defining characteristic of institutions (see, e.g., Giddens, 1984). Algorithms can be characterized similarly, in terms of the extent to which they have the capacity to structure user behaviors, impact preference formation, and impact content production decisions, through mechanisms that are technological in nature but are developed and frequently refined and recalibrated within complex social processes that are impacted by organizational and supra-organizational environmental conditions (see below; also see, e.g., Goldman, 2006; Grimmelman, 2008/2009; Jiang, in press).

This perspective also calls to mind some of the central propositions of actor-network theory (see, e.g., Latour, 2005). For instance, a central proposition of actor-network theory is that agency need not be restricted to humans. Non-human actors – such as, for example, media technologies (see, e.g., Couldry, 2008; Plesner, 2009), or, relatedly (as is the contention here), algorithms, interweave on equal footing with human actors to affect social conditions. This perspective further helps us to form a conceptual approach to institutions that accommodates a proposed institutional structure such as an algorithm, which in many ways epitomizes the complex intermingling of human and non-human actors that is central to an actor-network theory perspective on institutions.

From a more concrete standpoint, perhaps the most basic starting point for illustrating the intersection of algorithm and institution is Lessig's (2006) widely embraced notion that "code is law." Lessig (2006) employed this analogy to illustrate the ways in which the programming that controls the operation of communications networks and platforms is a powerful tool for regulating the behavior of users in ways that are not always obvious. If we extrapolate from Lessig's (2006) metaphor, systems of laws, are, of course, widely understood as institutions (see,

e.g., La Torre, 2010); and algorithms can be thought of as a form of code. And so, if code is law, and law is an institution, then forms of code such as algorithms should reasonably be thought of as institutions as well, particularly given, as will be illustrated below, algorithms often serve a similar function in terms of regulating individuals' behaviors (in addition to serving other core institutional functions)

This is, of course, a somewhat superficial, and semantic starting point. As a next step, we might consider Jepperson's (1991) detailed definitional assessment of the concepts of institutions and institutionalization, in which he notes the following: "Within any system having multiple levels or orders of organization . . . primary levels of organization can operate as institutions relative to secondary levels of organization. A microcomputer's basic operating system appears as an institution relative to its word-processing program (especially to a software engineer)" (p. 147). Here, of course, the notion of code (and thus, by extension, algorithm) as institution becomes a bit more explicit. Extending this analogy, we can think of algorithms functioning at a primary level of organization, providing the parameters of the framework in which subsequent organizational functions are carried out.

Jepperson's (1991) notion of relativity in the structure and functionality of institutions is particularly important to understanding the algorithm as a media institution. That is, there may be some algorithms (such as those employed by a dominant platform such as Google or Facebook), that are essentially *absolute* institutions in their own right, due to the wide-ranging scope of their significance and influence. Others may be better thought of as *relative* institutions, due to the way in which they may play an institutional role in more limited social or organizational contexts. Consider, for instance, the rise of "algorithmic journalism," in which algorithms are increasingly being embedded in the practice of journalism (Anderson, 2011a). As

will be discussed in greater detail below, algorithms are increasingly being used to establish the parameters in which journalists conduct their work. To a large extent, this is happening through a remediation of the relationship between journalists and their audiences, resulting in an ongoing redefinition of what have been termed "institutionally effective" audiences (Ettema & Whitney, 1994, p. 5).

The Algorithmic Turn and its Institutional Connections

The next step is to delve with greater specificity into the ways in which algorithms are functioning in the dynamics of media production and consumption, and to see how these functions connect with foundational articulations of what institutions are and what they do. For the purposes of this brief overview (which to some extent can only capture the tip of the iceberg in terms of the full extent of the algorithmic turn that is occurring in the media space), this discussion will be divided into the role of algorithms in: a) media consumption; and b) media production.

The Algorithmic Turn in Media Consumption

One of the key functions that algorithms perform in contemporary media consumption is to assist audiences in the process of navigating an increasingly complex and fragmented media environment. Central to this navigation process are the typically algorithmically-driven search and recommendation systems that facilitate searching for and selecting content in an environment of such extreme content abundance that technologically unaided forms of search and navigation are no longer practical or effective (Anderson, 2006). These algorithmically-driven systems are, of course, central to new media platforms ranging from search engines to online content providers such as Amazon, iTunes, YouTube, and Netflix.

To some extent, one could argue that the functionality of these search and recommendation algorithms is of equal or greater value than the actual content that these platforms provide. By some accounts, for instance, Netflix was able to withstand a challenge in the online/by mail video distribution business from better-resourced competitor Blockbuster due in part to its superior recommendation system (Keating, 2012). The value of the navigation algorithm is well-illustrated by the Netflix Prize competition, in which the company made a large amount of user data public and offered a one million dollar prize to the individual or team who could improve upon the accuracy of its Cinematch movie recommendation system (Bennett & Lanning, 2007). The effectiveness of Google's PageRank algorithm is widely credited with the search engine's continued dominance of the search market, despite being a late arrival in what was, at that point, already a highly competitive landscape (Battelle, 2006). The design of a more effective recommendation algorithm was seen as fundamental to enhancing the revenue prospects of YouTube, by increasing the average time users spent with the site (Helft, 2009).

From an institutional standpoint, it is important to emphasize, as Webster (2011) does, the extent to which these systems focus attention in particular ways and "structure decision making within certain bounds" (p. 50). In this regard, these systems exhibit characteristics that are inherently institutional in nature. Research has shown, for instance, that not only do Netflix users rely quite heavily on the service's recommendation system in the process of selecting video programming (Keating, 2012), and also that the design of recommendation systems can influence the content feedback that users provide -- altering the opinions that serve as inputs to; and thus, ultimately, the outputs of, the recommendation system (Cosley, et al., 2003).

It is also worth considering a case such as Twitter's Trends list, which provides users with a list of the most popular topics currently being discussed on the platform. This list is an

algorithmically-generated output from over 250 million tweets sent daily that serves in part to guide the media consumption behavior of Twitter users (pointing them toward popular topics). The broader significance of a case such as this was illustrated via the recent controversy over the seemingly premature disappearance of the Occupy Wall Street movement from the Twitter Trends list (see Gillespie, 2011). In response to charges of politically-motivated censorship, Twitter released details regarding the operation of its Trends algorithm. The company noted that the Trends algorithm is not based on a simple calculation of the most used terms, but rather seeks to also take into account factors such as whether the term is recently surging in popularity, the clustering patterns of the users of the term, the ratio of unique tweets to retweets, etc. (Gillespie, 2011).

This case is notable for a number of reasons. First, the very fact that the presence or absence of the Occupy Wall Street movement in the Twitter Trends list was a widely discussed and controversial topic in its own right illustrates the extent to which the list and the algorithm that generates it possess institutional characteristics. The magnitude of the controversy illustrates the extent to which the list represents a fundamental mechanism by which individuals and organizations form their cognitions of the online public sphere and the broader political dynamics reflected there. This brings us to the other notable dimension of this example – the explicitly political ramifications of the output of this particular algorithm, which is in keeping with the well-established understanding of the media as a political institution.

This political dimension of the functionality of algorithms has been demonstrated in other contexts as well. Recent research, for instance, has demonstrated that the manipulation of search engine rankings can potentially affect election outcomes, and that such manipulations can easily be conducted surreptitiously (Epstein & Robertson, 2013). Findings such as these reflect

increasing concerns about the influence of such algorithmically-driven platforms, due in large part to the already ingrained perceptions amongst users that such search returns represent objective and reliable representations of relevant online content. In a call for the imposition of public interest obligations on search engines (essentially, treating search engines comparably to another prominent media institution – commercial broadcasting), Laidlaw (2008) emphasizes that search engines are "authoritative and reliable, and shape public opinion and meaning" (p. 124). In these ways, an algorithmically driven platform such as a search engine genuinely is functioning in a similar political capacity to established media institutions.

In some media consumption contexts, algorithms go beyond providing users with a list of content options to choose from (i.e., search returns or recommendations). In some instance, the user even abdicates the final content decision to the algorithm in what we might term the *algorithm as programmer.* Consider, for instance, an online music provider such as Pandora, which, based on the user's expressed preferences, delivers a stream of music that its selection algorithm has determined is reflective of the user's expressed preferences (for a discussion of Pandora's algorithm, and those of other online music providers, see Magno & Sable, 2008).

In many ways, Facebook's well-known EdgeRank and GraphRank algorithms operate similarly, drawing upon the behaviors and the expressed preferences of users to deliver a specific bundle of content to the user's News Feed (Bucher, 2012a, 2012b). These algorithmically-driven platforms go one step beyond providing a set of recommendations for the user to choose from, instead making the final decision about which content will be consumed and which content will be ignored.

These examples are all reflective of the ways in which algorithms have been identified as having a "governmental power" and a "gatekeeping function" (Bucher, 2012b, pp. 8-9) and thus

operate as an extension of the institutional functionalities that long have been associated with traditional media. Consider the parallels with DiMaggio and Powell's (1991a) statement that "Institutions do not just constrain options; they establish the very criteria by which people discover their preferences" (p. 11). This statement perfectly encapsulates the functionalities of search and recommendation systems in users' media consumption behaviors (see generally Pariser, 2011).

In the realm of social media, algorithms are mediating not just the consumption of traditional media content, but also the dynamics of individual social relations and interactions. Ultimately, as Beer (2009) has illustrated, "algorithms have the capacity to shape social and cultural formations and impact directly on individual lives" (p. 994). Research has, for instance, demonstrated the political dimension of algorithmically-driven platforms such as social media platforms and search engines. In their examination of the role of Facebook in the formation of political groupings, Langlois, et al. (2009) emphasize the ways in which software and protocols (i.e., algorithms) function as "actors that intervene directly in cultural and communicational processes," and thus need to be thought of as "a new type of actor . . . [that] fundamentally changes the dynamics of the constitution of issues and their publics" (p. 429).

Here again, we see strong parallels between the functioning of algorithms and established understandings of the functioning of institutions. According to North (1981), institutions "provide the framework through which human beings interact. They establish the cooperative and competitive relationships which constitute a society" (p. 20). Clearly, this is increasingly what algorithms are doing, particularly via their centrality to media platforms such as social networking sites and applications

The Algorithmic Turn in Media Production

Algorithms are also playing an increasingly prominent role on the production side of the media equation. As the media environment grows more complex, with audiences increasingly fragmented and empowered, and with a growing array of technologies and platforms at their disposal, media organizations are increasingly turning to data and algorithms to help them effectively navigate this environment (Davenport & Harris, 2009). The two primary functions that algorithms are performing in the media production realm at this point are: a) serving as a demand predictor; and b) serving as content creator.

Looking first at the realm of demand prediction, in this "big data" era, media organizations have an ever-expanding supply of data on audiences' media consumption patterns and preferences to draw upon (Napoli, 2011), and algorithms play a central role in producing actionable insights from these stores of data. The motion picture industry, for instance, has begun to rely on predictive software packages such as Epagogix, which employ algorithms to predict the success of prospective film projects based upon the plot elements contained within the individual film scripts (Gladwell, 2006; Davenport & Harris, 2009). In this way, the algorithm's predictive outcome can supplement – or even replace – the judgment of the studio executives who have traditionally made such decisions on the basis of experience, intuition, and fairly superficial analyses of current and past box office trends.

Similarly, Netflix has been developing its new slate of original programming (such as *House of Cards* and *Hemlock Grove*) via feeding its enormous trove of audience behavior and preference data into a predictive algorithm that then identifies the type of original programming most likely to succeed (Carr, 2013; Leonard, 2013). The inputs in this case are obviously very different than the inputs being utilized by a system such as Epagogix, but the outcome is

essentially the same – algorithmically-derived performance estimates that are increasingly dictating production decisions.

Perhaps the most controversial application of such algorithmically-driven demand predictors has been in the realm of journalism, where news organizations are increasingly relying on analyses of various forms of user behavior and feedback data to more precisely calibrate their news gathering and reporting activities. Many news rooms now operate with comprehensive and immediate feedback related to various aspects of online news consumption, ranging from page views to time spent on a site/story, to ratings, to volume and valence of comments (see, e.g., Anderson, 2011b; Napoli, 2011).

But the issue here goes beyond the availability and analysis of new forms of audience data. Rather, the specific concern is the role that algorithms play in making sense of these data, and how these algorithmic analyses then affect content decision-making. Consider, for instance, the case of "content farms." Content farms mine search engine data to estimate demand for content and various topics, and then produce that content rapidly and cheaply in order to meet that demand. Once again, the process is algorithmically-driven. Leading content farm Demand Media, for instance, feeds its algorithm three types of data: a) popular search terms from search engines; b) the ad market for keywords (i.e., which keywords are currently being sought and for how much); and c) the competitive environment (in terms of content that's already available online) (Roth, 2009). The output then represents essentially a prediction of the type of content for which there is the highest audience demand, and Demand Media produces that content accordingly (see Anderson, 2011a).

In cases such as these, content production decisions are increasingly being algorithmically dictated in a context in which traditional institutional values emphasized a

decision-making process based less on audiences' expressions of their interests and wants and more on professionally-established norms and criteria that facilitated more public interest-oriented judgments regarding audiences' informational needs in order to be better-informed citizens (Anderson, 2011a). This transition essentially has as its core a tension involving a transition from one guiding institutional infrastructure to another in terms of the practice of journalism and the formulation and application of news values.

It is also worth noting the ways in which the content farm example represents a tension between different, but intersecting, algorithmic institutional structures. Specifically, in an effort to combat the extent to which "low quality" content farm articles appeared prominently in its search returns, Google adjusted its algorithm to reduce the relevance of "low quality" sites (defined as sites that copy content or provide low value-add to users) (Tartakoff, 2011). Here, the mechanisms for algorithmic media consumption are being adjusted in response to a system of algorithmic media production that relies, in part, on data on users' algorithmically-driven media consumption behaviors.

The algorithmic turn in media production is, in some instances, being enhanced in ways that go beyond demand prediction and extend into the realm of content creation. Essentially, the direct human element in the process of content creation is, in some contexts, being eliminated. Algorithms have been developed and employed to perform comparably to human content creators in areas such as poetry and music composition (see, e.g., Steiner, 2012). This model is also at the core of Narrative Science, a start-up based around a software package that can generate complete news stories once it is fed the core data around which the stories will be based (e.g., sporting event scores/stats, company financial reports, housing data, survey data, etc.) (see

Lohr, 2011). In this way, the direct human element is, to a large extent, removed from the process of generating news stories.

In this realm of algorithmic media production, we once again see strong intersections between the functionality of algorithms and the functionality of institutions. Consider, for instance, Douglas' (1986) statement in her book *How Institutions Think* (a title that compels a consideration of the intersection of algorithms and institutions): "Institutions encode information. They are credited with making routine decisions, solving routine problems, and doing a lot of regular thinking on behalf of individuals" (p. 47).

As this section has illustrated, in the realm of media production, common decisions relating to what content to produce and how to produce it are being delegated to algorithms. Given the extent to which media institutions research has focused over the course of its history on understanding the processes via which such culturally and politically significant decisions are reached within media organizations, incorporating the algorithmic turn into such analytical frameworks is now clearly vital.

The key point here is that, when we examine the ways in which algorithms are functioning in the processes of media production and consumption, we see strong parallels between these functionalities and those that have traditionally been ascribed to institutions. Indeed, algorithms are encroaching upon many of the functionalities (gatekeeping, production decision-making) that have traditionally been at the core of media institutions research.

Linking Institutional Theory and Algorithms

Now that some basic (but hopefully convincing) connections between the nature and function of algorithms and the nature and function of institutions have been drawn, the next step is to delve more deeply into institutional theory in an effort to identify the ways it can serve as an

interpretive lens for what we know at this point about the functionality and impact of algorithms in the media sector. This is not intended as a comprehensive overview of relevant institutional theory and its potential applicability to the algorithmic turn in media, but rather as an effort to identify some of the more obvious and/or potentially useful theoretical and/or subject matter connections.

Structuration

In Giddens' (1984) well-known theory of structuration, a key element of structuration is the notion of duality, which refers in this context to the extent to which agents and structures (translated in the context being examined here to individuals and institutions) mutually reproduce the social world. In a useful application of structuration theory to the new media environment, Webster (2011) illustrates how individuals and institutions mutually construct the media environment. Central to this process are what Webster (2011) terms "user information regimes" – the typically algorithmically-driven search and recommendation systems (see above) that facilitate searching for and selecting content. The notion of duality is of particular relevance here, as these user information regimes construct and constrain the how users perceive and engage with their media environment. At the same time, the activities of these users feed into (in the form of user data), and to some extent construct (in terms of how these user data are employed by algorithms in the functioning of the search and recommendation systems) these user information regimes and the roadmap to the media environment that they provide.

Specific patterns of mutual influence (i.e., duality) in the intersection of users and user information regimes are already being identified. One important pattern, for instance, is a certain amount of reflexivity that is inherent in much algorithmically-driven media consumption. For instance, as Bucher (2012) has illustrated in her analysis of Facebook's GraphRank algorithm,

the algorithm monitors users' behavior to find the most interesting patterns. Once these patterns are found, they are fed back to the users via Facebook's News Feed. "Consequently, even more users will apparently act in the way that the algorithm predicts" (Bucher, 2012, p. 14). In this regard, GraphRank acts as a mechanism "for predicting the future and [preventing] another future from happening" (Bucher, 2012, p. 14).

Further, the dynamics of many search, recommendation, and navigation algorithms emphasize popularity as a key criterion in generating results (see Jones, 2013; Webster, 2011), which again leads to a certain reflexivity in their operation. Popular content is what is most frequently and prominently recommended, thus further enhancing its popularity relative to other available content, and inhibiting less popular content from gaining popularity (see Cho & Roy, 2004). Given this structural characteristic of many algorithms, it is perhaps not surprising that predictions that the new media environment would facilitate unprecedented levels of audience consumption of "long tail" content (see Anderson, 2006) have, to some extent, failed to materialize (Napoli, 2011; Webster, 2011).

Institutional Isomorphism

Institutional isomorphism refers to the tendency for organizations in a particular field to resemble one another across a variety of dimensions. Explanations for this tendency include: a) coercion, which involves the "formal and informal pressures exerted on organizations by other organizations upon which they are dependent and by cultural expectations in the society within which the organization functions" (DiMaggio & Powell, 1991b, p. 67); b) mimetic processes, which involve organizations responding to uncertainty in their environment or objectives by modeling themselves on similar, or more successful organizations in their field; and c) normative

pressures, which involve the processes of professionalization (education, training, acculturation) that result in increased similarity across organizations (DiMaggio & Powell, 1991b).

This theoretical framework may be useful for understanding some patterns we see in the algorithmic turn in media production and consumption. Research has suggested for instance, that despite employing different Web crawling procedures and ranking algorithms, major search engines often exhibit high levels of similarity in their search returns (Hindman, 2009). Such findings raise the question of whether processes of institutional isomorphism may be at work. Institutional isomorphism in the realm of algorithmic media may even occur across algorithmic and non-algorithmic media platforms. For instance, as Gillespie (2011) points out, the emphasis on novelty that has characterized news reporting in the traditional media sector seems to be reproducing itself in the online space in contexts such as the design of the Twitter Trends algorithm (discussed above). Similarly, Rogers (2004) has demonstrated how Google search returns tend to map quite closely with the issues and sources characteristic of mainstream media. Obviously, going beyond such observations, and untangling the likely causes, is an important next step; one in which the institutional isomorphism framework may provide a useful analytical perspective.

It also seems reasonable to consider algorithms as a potential driver of institutional isomorphism. To the extent that more and more content producers are, for instance, relying upon algorithmic demand prediction models that are derived from historical data on audiences' exhibited consumption patterns, there would seem to be the possibility that different organizations are increasingly likely to produce similar outputs, to the extent that they are all essentially data mining the same history and producing content on the basis of those results. For instance, the more motion picture studios that rely on Epagogix to determine their production

slate, the more likely it would seem that these studios will produce similar films, as the same algorithm and underlying data are driving their decision-making. In this regard, algorithmically-driven institutional isomorphism essentially results in diminished diversity of content output.

The mimetic processes identified by DiMaggio and Powell (1991) as a way that organizations cope with uncertainty in their environment would seem, on the surface, as the most likely driver of institutional isomorphism within the context of algorithmically-driven media enterprises. The reliance on big data and algorithmic outputs in organizational decision-making can be seen as the latest manifestation of a long-running trend in the media sector to more fully embrace data-driven decision-making as a means of counteracting environmental uncertainty (Napoli, 2011).

Social Constructivism

Social constructivism addresses the ways in which groups created shared meanings and cognitions, which then serve as important mechanisms for guiding behaviors. This theoretical perspective emphasizes that social reality is a product of social processes directed at the establishment of shared knowledge and belief systems (for a review, see Scott, 2008). This approach is particularly relevant in terms of its emphasis on cognitions, as well as in terms of its applications to the construction of technological systems (see, e.g.,). Both (inter-related) contexts – the social construction of knowledge and the social construction of technological systems, would seem to bear directly on the construction, operation, and impact of algorithms.

Those who focus on the cultural-cognitive dimension of institutions emphasize "the extent to which behavior is informed and constrained by the ways in which knowledge is constructed and codified. Underlying all these decisions are socially constructed models, assumptions, and schemas" (Scott, 2008, p. 68). Algorithms obviously serve as prime examples

of constructors and codifiers of knowledge, particularly in contexts such as search engines, which play a central role in aggregating, categorizing, organizing, and presenting information, to the point that both Google and Googlization function as verbs in contemporary language (see, e.g., Halavais, 2009; Vaidhyanathan, 2011).

This theoretical perspective points us to the importance of understanding the *social construction of algorithms.* As Berger and Luckmann (1966) have emphasized, "To understand the state of the socially constructed universe at any given time, or its change over time, one must understand the social organization that permits the definers to do the defining" (p. 116).

Here we start to see an important point of connection between the social construction of knowledge and the more specific context of the social construction of technological systems. Technological systems have been described as having messy, complex problem-solving components and are both socially constructed and society shaping (here we see a recurrence of the duality theme that is central to our understanding of institutions and algorithms) (Hughes, 2012). Their components include a physical component, organizations, and regulations and laws (Hughes, 2012). Importantly for this context, they are also defined to include both hardware and software (Constant, 2012). Algorithms would seem to fit quite well within this definitional framework.

Continuing with this line of reasoning, it seems reasonable to think of algorithms as "the definers" that "do the defining," in many contexts, thereby playing an increasingly influential role in the social construction of knowledge. Understanding the social organization underlying these definers means understanding the social processes underlying the construction of the algorithms (i.e., technological system) that play an increasingly influential role in the social construction of knowledge

Approaching algorithms from this perspective is in many ways an extension of recent calls for research on, for example, search engines, to shift from an emphasis on the impact of search engines to an emphasis on the "social practices and power relations involved in the construction of search engines" (Mager, 2012, p. 782). Recent work that has taken initial steps toward "political economy of the algorithm" (Pasquinelli, 2009) fits into this analytical framework, as does emerging qualitative work examining the thought processes and perspectives are a reflection of the specific cultural schemas possessed by those engaged in algorithm development and maintenance (Van Couvering, 2007). Van Couvering (2007), for instance, illustrates how market (oriented around competitive and revenue-generating concerns) and science/technology (oriented around norms of objectivity, innovation, and experimentation) schemas dominate the conceptualizations of "quality" in the design, assessment, and modification of search algorithms. In this way, we see how specific social values and norms get reflected in the design of algorithms – important steps to understanding the social construction of these technological systems.

Conclusion

This paper has attempted to serve a number of functions. Hopefully, it has made a compelling case for approaching algorithms as a distinctive form of media institution; and in so doing has illustrated a number of ways in which institutional theory can meaningfully inform research on algorithms as (or as part of) media institutions. In developing these lines of reasoning, this paper has also hopefully illustrated a number of specific contexts in which the political and cultural implications of the operation of algorithms in the media sector are quite pronounced – essentially rising to a level of institutional significance.

This concluding section puts forward some lines of inquiry common to institutional theory that seem particularly relevant for developing a deeper understanding of algorithms' institutional functionality and impacts. Although there is currently a welcome surge in research attention to the role of algorithms in the media sector (Barocas, Hood, & Ziewitz, 2013), up until very recently scholars noted that it is "surprising how little attention such software infrastructures have received from the academic community" (Bucher, 2012a, p. 1176).

Greater attention to this subject may be able to revive an area of media scholarship that has, by some accounts, grown dormant. Specifically, a fairly recent assessment of the field of media sociology has described the field has having "largely stalled" after the many pathbreaking institutional analyses of news organizations in the 1970s and 1980s (Kaplan, 2006, p. 173). While recent examinations of how news organizations are adapting to the online era to some extent belie this characterization (see, e.g., Paterson & Domingo, 2008, 2011), it does seem reasonable to contend that an emphasis going forward on what Anderson (2011c) has called the "sociology of the algorithm" (p. 6) could represent a vital new direction for media sociology. And so, the goal here is to identify some areas of institutional theory that seem particularly well-suited to informing future work.

First, it is important to more fully examine algorithms from the perspective of institutional effects. Institutional effects refer to those effects that "feature institutions as causes" (Jepperson, 1991, p. 153). These effects may be upon other institutions, or upon other categories such as individuals or organizations. If we think, then of algorithms as institutions, the question then becomes, what kinds of effects are algorithmically-driven systems of media production and consumption having not only on other institutions, but on other aspects of civic

and cultural life? In what specific ways is cultural production (that is, the actual content) being

transformed by increased reliance on algorithmic systems for decision-making?

Another important line of inquiry is the process of institutionalization – that is, the

process via which specific practices, norms, structures, or environmental cognitions become

widely accepted and adopted. Some scholars have emphasized that future institutional research

should shift from the current focus on the outcomes or products of institutional influence and

more on the processes via which institutions are established (Suddaby, 2010). Such a direction

would seem particularly important during this period of what appears to be fairly rapid

institutionalization of algorithmically-driven decision-making in the media sector. We know

little, at this point, for instance, about the organizational dynamics surrounding the adoption and

usage of algorithmic tools in the media sector. Are there intra-organizational tensions; and if so,

how are they being resolved? How are established professional norms, identities, and practices

adapting? How are algorithmic tools becoming legitimized in organizational processes?

This last question seems particularly important in relation to the institutionalization of

algorithms. As many observers have pointed out, the internal operation of algorithms is a

combination of complexity and intentional opacity (in order to protect competitive advantages

and to prevent "gaming" of the systems) (see, e.g., Elgesem, 2008). Such characteristics would

seem to potentially serve as a significant roadblock to institutionalization. However, as Berger

and Luckmann (1966) note, "The objective reality of institutions is not diminished if the

individual does not understand their purpose or their mode of operation" (Berger & Luckmann,

p. 60). Certainly, the might be said for the role that algorithms play in the increasing automation

of media production and consumption.

Bibliography

Ahlkvist, J.A. (2001). Programming philosophies and the rationalization of music radio. *Media, Culture & Society, 23*(3), 339-358.

Alchian, A.A., & Demsetz. H. (1972). Production, information costs, and economic organization. *American Economic Review, 62,* 777-795.

Anderson, C. (2006). *The long tail: How the future of business is selling less of more.* New York: Hyperion.

Anderson, C.W. (2011a). Deliberative, agonistic, and algorithmic audiences: Journalism's vision of its public in an age of audience transparency. *International Journal of Communication, 5,* 529-547.

Anderson, C.W. (2011b). Between creative and quantified audiences: Web metrics and changing patterns of newswork in local U.S. newsrooms. *Journalism: Theory, Practice, Criticism, 12*(5), 550-566.

Anderson, C.W. (2011c, October). Understanding the role played by algorithms and computational practices in the collection, evaluation, presentation, and dissemination of journalistic evidence. Paper presented at the 1st Berlin Symposium on the Internet and Society, Berlin.

Anderson, C.W. (in press)). Towards a sociology of computational and algorithmic journalism. *New Media & Society,*

Bakker, (2012). Aggregation, content farms, and Huffinization: The rise of low-pay and no-pay journalism. *Journalism Practice, 6*(5-6), 627-637.

Barocas, S., Hood, S., & Ziewitz, M. (2013, May). Governing algorithms: A provocation piece. Paper prepared for the Governing Algorithms conference, New York.

Battelle, J. (2006). *The search: How Google and its rivals rewrote the rules of business and transformed our culture.* New York: Portfolio.

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society, 11,* 985-1002.

Bennett, J., & Lanning, S. (2007). The Netflix prize. Paper presented at the KDDCup 2007 conference, San Jose, CA.

Benson, R. (2006). New media as a "journalistic field": What Bourdieu adds to new institutionalism, and vice versa. *Political Communication, 23,* 187-202.

Berger, P.L., & Luckmann, T. (1966). *The social construction of reality: A treatise on the sociology of knowledge.* New York: Anchor Books.

Bucher, T. (2012a). Want to be on top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society, 14*(7), 1164-1180.

Bucher, T. (2012b). A technicity of attention: How software "makes sense." *Culture Machine, 13,* 1-13.

Carr, D. (2013, February 24). Giving viewers what they want. *New York Times.* Retrieved February 24, 2013, from: http://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html?pagewanted=all&_r=0.

Cho, J., & Roy, S. (2004, May). Impact of search engines on Web page popularity. Paper presented at the WWW2004 Conference, New York.

Constant, E.W., II (2012). The social locus of technological practice. In W.E. Bijker, T.P. Hughes, & T. Pinch (Eds.), *The social construction of technological systems: New directions in the sociology and history of technology* (anniversary ed.) (pp. 217-236). Cambridge, MA: MIT Press.

Cook, T.E. (2005). *Governing with the news: The news media as a political institution* (2nd ed.). Chicago: University of Chicago Press.

Cook, T.E. (2006). The news media as a political institution: Looking backward and looking forward. *Political Communication, 23,* 159-171.

Cosley, D., Lam, S.K., Albert, I., Konstan, J.A., Riedl, J., (2003, April). Paper presented at the CHI 2003 conference, Fort Lauderdale, FL.

Couldry, N. (2008). Actor-network theory and media: Do they connect and on what terms? in A. Hepp, F. Krotz, S. Moores & C. Winter (Eds.). *Connectivity, networks and flows: Conceptualizing contemporary communications* (pp. 93–111). Cresskill, NJ: Hampton Press.

Danaher, P., Lee, J., Laoucine, K. (2010). Optimal Internet media selection. *Marketing Science, 29*(2), 336-347.

DiMaggio, P.J., & Powell, W.W. (1991a). Introduction. In W.W. Powell & P.J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 1-38). Chicago: University of Chicago Press.

DiMaggio, P.J., & Powell, W.W. (1991b). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. In W.W. Powell & P.J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 63-82). Chicago: University of Chicago Press.

Douglas, M. (1986). *How institutions think.* New York: Syracuse University Press.

Elgesem, D. (2008). Search engines and the public use of reason. *Ethics and Information Technology, 10,* 233-242.

Epstein, R., & Robertson, R.E. (2013, May). Democracy at risk: Manipulating search rankings can shift voting preferences substantially without voter awareness. Paper presented at the annual meeting of the Association for Psychological Science, Washington, DC.

Ettema, J.S., & Whitney, D.C. (1994). The money arrow: An introduction to audiencemaking. In J.S. Ettema & D.C. Whitney (Eds.), *Audiencemaking: How the media create the audience* (pp. 1-18). Thousand Oaks, CA: Sage.

Flanagin, A.J., Flanagin, C., & Flanagin, J. (2010). Technical code and the social construction of the Internet. *New Media & Society, 12*(2), 179-196.

Galperin, H. (2004). Beyond interests, ideas, and technology: An institutional approach to communication and information policy. *The Information Society, 20,* 159-168.

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration.* Berkeley, CA: University of California Press.

Gillespie, T. (2011, October). Can an algorithm be wrong? *Culture Digitally.* Retrieved April 22, 2013, from: http://limn.it/can-an-algorithm-be-wrong/.

Gladwell, M. (2006, October 16). The formula: What if you built a machine to predict hit movies? *The New Yorker.* Retrieved April 23, 2013, from: http://www.newyorker.com/archive/2006/10/16/061016fa_fact6.

Goldman (2006). Search engine bias and the demise of search engine utopianism. *Yale Journal of Law & Technology, 9,* 111-123.

Grimmelman, J. (2008/2009). The Google dilemma. *New York Law School Law Review,* 53, 939-950.

Guzman, T. (2005). The little theatre movement: The institutionalization of the European art film in America. *Film History, 17,* 261-284.

Halavais, A. (2009). *Search engine society.* Cambridge, UK: Policy Press.

Helft, M. (2009, December 30). YouTube's quest to suggest more. New York Times. Retrieved December 30, 2009, from:

http://www.nytimes.com/2009/12/31/technology/internet/31tube.html.

Herbst, S. (1993). *Numbered voices: How opinion polling has shaped American politics.* Chicago: University of Chicago Press.

Himdman, M. (2009). *The myth of digital democracy.* Princeton, NJ: Princeton University Press.

Hrynshyn, D. (2008). Globalization, nationality and commodification: The politics of the social construction of the Internet. *New Media & Society, 10*(5), 751-770.

Hughes, T.P. (2012). The evolution of large technological systems. In W.E. Bijker, T.P. Hughes, & T. Pinch (Eds.), The social construction of technological systems: New directions in the sociology and history of technology (pp. 45-76). Cambridge, MA: MIT Press.

Jepperson, R.L. (1991). Institutions, institutional effects, and institutionalism. In W.W. Powell & P.J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 143-163). Chicago: University of Chicago Press.

Jiang, M. (in press). The business and politics of search engines: A comparative study of Baidu and Google's search results of Internet events in China. *New Media & Society.*

Jones, J. (2012, October). Creating networks through search: PageRank, algorithmic truth, and tracing the Web. Paper presented at the annual meeting of the Association of Internet Researchers, Salford, UK.

Kaplan, R.L. (2006). The news about new institutionalism: Journalism's ethic of objectivity and its political origins. *Political Communication, 23,* 173-185.

Katzenbach, C. (2011). Technologies as institutions: Rethinking the role of technology in media governance constellations. In M. Puppis & M. Just (Eds.), *Trends in communication policy research* (pp. 117-138). Bristol, UK: Intellect.

Keating, G. (2012). *Netflixed: The epic battle for America's eyeballs.* New York: Portfolio.

Kendall, G. (2004). Towards a sociology of nonhumans: Technology and creativity. Paper presented at Social Change in the 21st Century Conference, Queensland, Australia.

Kim, J. (2012). The institutionalization of YouTube: From user-generated content to professionally generated content. *Media, Culture & Society, 34*(1), 53-67.

La Torre, M. (2010). Law as institution. *Law & Philosophy Library, 90,* 97-134.

Laidlaw, E.B. (2008). Private power, public interest: An examination of search engine accountability. *International Journal of Law & Information Technology, 17*(1), 113-145.

Langlois, G., Elmer, G., McKelvey, F., & Devereaux, Z. (2009). Networked publics: The double articulation of code and politics on Facebook. *Canadian Journal of Communication, 34,* 415-434.

Latour, B. (2005). *Reassembling the social: An introduction to actor network theory.* Oxford: Oxford University Press.

Leonard, A. (2013, February 1). How Netflix is turning viewers into puppets. *Salon.* Retrieved April 24, 2013, from: http://www.salon.com/2013/02/01/how_netflix_is_turning_viewers_into_puppets/.

Lessig, L. (2006). *Code and other laws of cyberspace 2.0.* New York: Basic Books.

Lohr, S. (2011, September 10). In case you wondered, a real human wrote this. *New York Times.* Retrieved April 22, 2013, from: http://www.nytimes.com/2011/09/11/business/computer-generated-articles-are-gaining-traction.html?pagewanted=all&_r=0.

Lowrey, W. (2011). Institutionalism, news organizations and innovation. *Journalism Studies, 12*(1), 64-79).

Lowrey, W., & Woo, C.W. (2010). The news organization in uncertain times: Business or institution? *Journalism & Mass Communication Quarterly, 87*(1), 41-61.

Mager, A. (2012). Algorithmic ideology. *Information, Communication & Society, 15*(5), 269-787.

Magno, T., & Sable, C. (2008, September). A comparison of signal-based music recommendation to genre labels, collaborative filtering, musicological analysis, human recommendation, and random baseline. Paper presented at at the ISMIR 2008 conference, Philadelphia, PA.

Mayer-Schonberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think.* New York: Houghton Mifflin Harcourt.

McKelvey, F. (2010). Ends and ways: The algorithmic politics of network neutrality. *Global Media Journal, 3*(1), 51-73.

Moe, (1990). Political institutions: The neglected side of the story. *Journal of Law, Economics, and Organizations, 6,* 213-253.

Moe, H., & Syvertsen, T. (2007). Media institutions as a research field. *Nordicom Review, 28,* 149-167.

Napoli, P.M. (1997). A principle-agent approach to the study of media organizations: Toward a theory of the media firm. *Political Communication, 14*(2), 207-219.

Napoli, P.M. (2003). *Audience economics: Media institutions and the audience marketplace.* New York: Columbia University Press.

Napoli, P.M. (2011). *Audience evolution: New technologies and the transformation of media audiences.* New York: Columbia University Press.

North, D. (1981). *Structure and change in economic history.* New York: Norton.

Pasquinelli, M. (2009). Google's PageRank algorithm: A diagram of the cognitive capitalism and the rentier of the common intellect. In K. Becker & F. Stalder (Eds.), *Deep search: The politics of search beyond Google.* Innsbruck, Germany: Studien Verlag

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you.* New York: Penguin.

Paterson, C., & Domingo, D. (2008). *Making online news (vol. 1).* New York: Peter Lang.

Paterson, C., & Domingo, D. (2011). *Making online news (vol. 2).* New York: Peter Lang.

Plesner, U. (2009). An actor-network theory perspective on changing work practices: Communication technologies as actants in newswork. *Journalism, 10*(5), 604-626.

Rogers, R. (2004). *Information politics on the Web.* Cambridge, MA: MIT Press.

Roth, D. (2009, November). The answer factory: Demand Media and the fast, disposable, and profitable as hell media model. *Wired.* Retrieved April 24, 2012, from: http://www.wired/com/magazine/2009/10/ff_demandmedia.

Ryfe, D.M. (2006). New institutionalism and the news. *Political Communication, 23,* 135-144.

Schudson, M. (2002). The news media as political institutions. *Annual Review of Political Science, 5,* 249-269.

Scott, W.R. (1991). Unpacking institutional arguments. In W.W. Powell & P.J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 164-182). Chicago: University of Chicago Press.

Scott, W.R. (2008). *Institutions and organizations: Ideas and interests* (3rd ed.). Los Angeles: Sage.

Sparrow, B.H. (1999). *Uncertain guardians: The news media as a political institution.* Baltimore, MD: Johns Hopkins University Press.

Sparrow, B.H. (2006). A research agenda for institutional media. *Political Communication, 23,* 145-157.

Steiner, C. (2012). *Automate this: How algorithms came to rule our world.* New York: Portfolio.

Suddaby, R. (2010). Challenges for institutional theory. *Journal of Management Inquiry, 19*(1), 14-20.

Tartakoff, J. (2011, February 24). In latest anti-"content farm" move, Google changes its search algorithm. *paidContent.* Retrieved April 29, 2013, from: http://paidcontent.org/2011/02/25/419-in-latest-anti-content-farm-move-google-changes-its-algorithm/.

Uricchio, W. (2011). The algorithmic turn: Photosynth, augmented reality and the changing implications of the image. *Visual Studies, 26*(1), 25-35.

Vaidhyanathan, S. (2011). *The Googlization of everything (and why we should worry).* Berkeley, CA: University of California Press.

Van Couvering, E. (2007). Is relevance relevant? Market, science, and war: Discourses of search engine quality. *Journal of Computer-Mediated Communication, 12,* 866-887.

Webster, J.G. (2010). User information regimes: How social media shape patterns of consumption. *Northwestern University Law Review, 104*(2), 593-612.

Webster, J.G. (2011). The duality of media: A structurational theory of public attention. *Communication Theory, 21,* 43-66.