

Stream: Inspiring Critical Thought
2018, Vol 10(1), 39–51
© The Author(s), 2018
<http://journals.sfu.ca/stream>

I/O: Reinforcing Newsmaking Practices Through Algorithmic Media

Luciano Frizzera

Department of Communication
Concordia University

Abstract

Recent developments in communication and information technology have disrupted the long-established dominance of mass media over the production and distribution of news. As an effort to reclaim their role of society's information gatekeeper, media companies absorb digital technology as instruments of institutional power to reproduce its own logic in the digital space. This paper discusses two interrelated modalities of algorithmic news: economically efficient production, where news outlets utilize quantitative metrics to improve content effectiveness and desirability; and shared-gatekeeping, where visibility and distribution of information are contextual and based on users' behaviour. The paper proposes that algorithmic media hides under its supposed objectivity and neutrality to become a new gatekeeper "organism", which not only regulates flows of information, but also interprets and negotiates both public interests and the value of the news.

Keywords

algorithmic Media, Newsmaking, Journalism, Gatekeeping, Big Data, Social Media

Introduction

The production of the news follows a simple process: selection of events, construction of a narrative that reports on the main features of these events, and distribution of these stories. These operations are carefully conducted by journalists and editors, and managed through a generally agreed-upon collection of rules and protocols through which the goal is to represent reality and inform the public, but also functioning as information gatekeepers (Wolf, 1987). However, digital technologies have disrupted this paradigm: blogs, social networks, mobile media, and search engines unlock and "democratize" the means of news production, distribution, and consumption, enabling anyone with the technological means to become a "journalist". Consequently, newspapers have had their role as gatekeepers questioned while the very concept of news has been challenged (Bentes, 2015; Carlson, 2007; Ramonet, 2013).

In response to this threat, and as a strategy to reclaim their role as gatekeepers, media companies began to integrate computer algorithms in their practice (*e.g.*, The Boston Globe, The Huffington Post, Vox Media, Associated Press). As a result, a clear trend in newsmaking has emerged: technology development is driving toward the personalization of our news experience (Gillespie, 2014; Napoli, 2014; O'Neil, 2016). Everything we click, read, search, and watch is the result of some delicate optimization process that uses our own social interactions to determine what we see in our newsfeed (McKelvey, 2014)¹. This is the building block of the so-called social media filter bubble

(Pariser, 2012), in which people are exposed only to information from like-minded individuals, amplifying confirmation bias.

One of the particularities of the “algorithmic turn” in the media sector is related to the intensive use of digital technologies and large volumes of data (*i.e.*, big data) to enhance decision-making about the production of content and the preferences of the audience. According to O’Neil (2016), the general assumption is that algorithmic media has more objective and efficient ways to deliver content to the public using a user-centric model: the news becomes contextualized in relation to the reader. It is a response to “growing quantities of available data, as well as a motivator for media organizations to gather ever more data from every available source to feed into massive processing capacities of these algorithms” (Napoli, 2014, p. 340). However, “algorithms are encoded procedures for transforming input data into a desired output, based on specified calculations” (Gillespie, 2014, p. 167). That is, algorithmic media are becoming the fundamental principle that governs the flows of information on which we depend.

This paper proposes that algorithmic media is a new form of gatekeeper “organism”, in which they not only regulate flows of information, but also interpret and negotiate both news value and public interests. To demonstrate that algorithmic media is not disrupting journalism practices and values, but rather reproducing them in the digital space, this paper examines two interrelated modalities. First, it considers the efficiency of news productions, where media companies utilize automation and A/B testing to improve effectiveness and desirability of content. Next, it analyzes the visibility and distribution of information based on user behaviour, location, and other non-disclosed attributes, as a form of shared gatekeeping.

Media Institutions: Gatekeeping and Newsmaking

It is important to consider the constitutive practices and functions of journalism before jumping into the examination of how digital tools are changing newsmaking. Journalism is traditionally understood as an institution residing at the intersection of complex economic organizations and an equally complex set of norms and procedures related to its professional practice (*e.g.*, public interest, objectivity) (Napoli, 2014). Typically, news production is carefully executed by journalists and editors through a process that involves selection, report, and distribution of information, following a generally agreed-upon collection of rules and protocols in which the goal is to represent reality and inform the public (Deuze, 2005; Wolf, 1987).

For Lesage and Hackett (2014), journalism is a response to a “problem of scale” in society: “people, as part of a ‘self-informing populace’, are unable to consider distant current events and so turn to journalism as a way of understanding what is happening in the present-day world” (p. 41). This resonates with what Lippmann (1997) labels intelligence work: “not to burden every citizen with expert opinions on all questions, but to push that burden away from him towards a responsible administrator” (p. 399). Indeed, according to Deuze (2005), one of the journalism’s core values is to “provide a public service (as watchdogs or ‘newshounds’, active collectors and disseminators of information)” (p. 447), that is, to keep the public informed about what is current and important. Consequently, newspapers and media companies in general, are gatekeepers that control flows of information, and have the power to decide what can be said and who is authorized to say it.

The concept of the gatekeeper was developed by Lewin (1947) when he was studying the social dynamic presented within a group of people, in particular with problems related to food habits. He observed that information flows through different channels which can also function as gates or filters. Generally, in a process of communication, there are individuals, or groups of individuals, who hold the power to control the flow of information — the gatekeepers (Wolf, 1987). This concept

was later applied by White (1950) to mass media, in which he showed that the selection and filtering of news places lower priority on individual subjectivity, and higher priority on a collection of norms and values that include both professional and organizational criteria, such as efficiency, speed, ideology, and public interest. Fifty years later, Shoemaker, Eichholz, Kim, and Wrigley (2001) confirmed the gatekeeping phenomenon in a more contemporaneous newsroom. Singer (2006), on the other hand, claims that online journalism began to move away from these practices; her research, however, was conducted before the rise of social and mobile media, and the adoption of algorithms by news companies.

The gatekeeping function in mass media is exercised in decisions involving selection of topics and sources, message encoding and formatting, and news distribution and visibility. It is precisely the gatekeeping function of newsmaking—and the routines and protocols of everyday news production—that gives newspapers and media organizations the power and authority to represent social reality to the public. The organization, structural requirements, and the techno-expressivity characteristics inherent in each mass communication medium are fundamental elements in determining the reproduction of social reality provided by mass media (Wolf, 1987). After all, the main purpose of any news organization is to provide reports of meaningful events (Deuze, 2005). However, one may ask which events are considered sufficiently interesting, meaningful, and relevant to become news.

As a complex institution, journalism follows a collection of systematic protocols to make the news. Journalists cannot redefine the criteria for selecting the news each time a new event occurs; this would make their work impractical. Hence, criteria of relevance were developed to define the “newsworthiness” and “news value” of each event; that is, the capability of an event to become a news event. These are a series of paradigms and protocols (parameters, task, pathways, values) which in turn, are naturalized through the professional practice, producing a journalistic ideology: practical rules embedded in professional knowledge (including “involuntary distortions”) that implicitly, and sometimes explicitly, guides newsmaking operations and procedures (Deuze, 2005). These protocols, the conventions of professional journalism, are created not only to determine what should be considered the news, but also to legitimate the news production process: from the selection of sources and topics, to the way the content is crafted. The main objective of these protocols is, therefore, to create a procedure for newsmaking in order to make the journalism practice feasible and manageable (Wolf, 1987).

However, the ‘news value’ operates in a peculiar manner. On one hand, news selection is a decision process made at speed, thus, the set of criteria must be quick and easy to apply in order to ease the cognition burden on the professional. On the other hand, these criteria should be flexible enough to adapt to any circumstance and to a variety of different events. They should allow quick comparisons between other available events, which is key in determining whether to add or suppress the news from the newspaper’s edition and other streams of information (*i.e.*, TV, radio, Internet). Ultimately, these criteria should be oriented to guarantee efficiency within the system, keeping the flow of information at a steady pace with a minimum time expenditure, effort, and resources. Wolf (1987) argues that the rigour of the ‘news value’ is not, therefore, an abstract, well structured, and theoretically coherent classification system; rather, it is a logic of classification that aims to achieve practical purposes in a planned manner and which is designed, above all, to enable the repetition of certain procedures.

Algorithmic Media

Recent developments in information and communication technology, particularly the introduction of algorithmic media, have changed practices and values in the newsmaking industry. Gillespie (2014) argues that algorithms play a key role in “selecting what information is considered most relevant to us, a crucial feature of our participation in public life” (p. 167). Napoli (2014) suggests that they are the “base structures and parameters that regulate the production, distribution, and consumption of content” (p. 343). In fact, when algorithms are considered as a collection of rules and predefined tasks and procedures, their functionalities and effects map quite closely alongside the practices of journalists and media institutions. Algorithms enact many kinds of control and fit in what Franklin (1999) describes as a prescriptive technology. These computational routines involve matters of representation and inclusion: they define a certain way — a process or a prescription — to do something. These “protocols” (Galloway, 2006) are a formalized set of rules and values that not only makes the results of any activity more predictable, but also excludes all other possibilities to engage on such activity.

The power of digital media is not on what they present in a visible format, but in its intricate network of codes and protocols, that is, in the algorithms that collect, transform, and generates the content. Algorithms are unreachable and hidden not only inside of highly complex calculation but also protected by intellectual property laws (Bucher, 2012). These codes are the ‘formalization’ of hegemonic practices, a set of strict and proper rules based on a specific system of ideological conventions that determines its power relations (Franklin, 1999). Prescriptive technologies aim for predictability in its power relations — a way to not only anticipate results, but the order of things. In fact, digital technology becomes an agent of ordering and structuring: every act becomes stored as footprints of social transactions and available for future reference. It is used to regulate, include/exclude actors, open/close doors, as well as provide means to know more about the world, how to organize this knowledge, and who is authorized to access it.

In fact, algorithms cannot be disconnected from the processes that create them. They are not just pieces of code running on machines making automatic decisions, but also an intermingling of pre-established technical protocols, professional practices, and social-economic interests. Moreover, they should be thought of as actors in a complex network of communication. Latour, Jensen Venturini, Grauwin, and Boullier (2012) define actors as not only humans and human groups, but also nonhumans operating on equal footing to affect social conditions: institutions, biological elements, technological artefacts, and, of course, machines. In this sense, the concept of algorithmic media seeks to avoid overemphasizing the importance of any single algorithm, software, or platform. They are systems composed of many different algorithms that cooperate and compete with one another (McKelvey, 2014), and have the capability to shape user behaviours, influence preference formation, and impact content production decisions. That is, algorithmic media mediate flows of information between people, institutions, and other social agents.

Considering that algorithms are a collection of codes of conduct, communication protocols, and idiosyncratic codes (Galloway, 2006), they can be seen as the base layer of any professional practice, including journalism. Just as in a microcomputer the operating system is the basic framework that prescribes the proper rules of operation to a word-processing program, in media institutions, the professional rules and codes of conduct define the rules of operation for their activities (Napoli, 2014). Algorithmic mediation is, therefore, a translation of a set of rules and values already established by traditional practice into another milieu. It is a “formalization” of current practices — written and non-written — materialized into computer codes: a routine, as it is called by programmers.

When algorithms are used to control operations of communication, particularly when they take over the main functions of newsmaking and journalism (selection, production, distribution, and reception of the news), they can be a powerful tool for regulating information flow (Gillespie, 2014).

Moreover, they can be used to govern users' behaviour in ways that are not always obvious since the code is usually protected, opaque, and not easy to read or even to be stopped. Thus, we can think of algorithmic media as a new form of gatekeeper "organism" that regulates flows of information with the capacity to influence public opinion.

Efficient Production

Media organizations are increasingly turning to data and algorithms to find efficient ways to produce and deliver content (C. W. Anderson, 2013). One common practice is to learn what their audience wants next. Netflix, for instance, has been investing in the creation of original content based on the users' behavioural data which the company harvest continuously from their system (ratings, number of shows viewed, hours spent on each show, etc.) (Vanderbilt, 2013). Transferring production decision to the machine may seem to be a disruption in the realm of the motion pictures, long praised as a form of art. Rather, it is a continuity in the logic of capital. Systematization and optimization are not strange to capitalist society (Dyer-Witheford, 1999), since companies are continuously looking for the most effective and profitable way to produce a new product. The next sections explore how this tendency can be problematic in the newsmaking domain since it deals not only with the distribution of information as a public good but also with the formation of opinions and ideologies.

A/B Testing

To an increasing extent, news organizations are becoming dependent on algorithmic analyses of various forms of user behaviour and feedback data to better define their news gathering and reporting activities (Napoli, 2014). In this case, the problem lies in the role that algorithms play in making sense of this data, and how these algorithmic analyses then affect content decision-making. Consider the A/B testing, a practice being implemented in many editorial companies, including The Boston Globe, The Huffington Post, Vox Media, and BuzzFeed, to optimize headlines and images in their stories online (Soberman, 2013). While this technique has been used in scientific experimental research and for marketing purposes for decades (O'Neil, 2016), it is only recently that news organizations began to use A/B testing to understand how readers engage with the news.

From an editorial perspective, increasing reader engagement could potentially correlate with a growing informed audience, which is commonly said to be part of a news organization's mission (Deuze, 2005). Not surprisingly, this technique is used especially in headlines, since they are the selling point of the news. Traditionally, most of the content decisions were made by editors and reporters following professionally-established criteria (newsworthiness and news value) about audiences' informational needs in order to be better-informed citizens: a gatekeeping process in which the audience had little input (Wolf, 1987). In their most commercial facet, news corporations need to capture readers' attention, engage the audience with the news piece in order to "sell" the story, and, as a consequence, attract advertisement, and make a profit.

A simple A/B testing would have a news report released with two competing versions of its title: half of the audience sees one version and the other half sees the other. The testing period is usually short, but can vary from a few minutes to many hours. Only a portion of the readers will participate in this blind test, though it can be extended by the editor and combined with other forms of personalization, such as targeting specific audiences with words that they favour the most. In the end, the title with the best performance, usually by number of clicks, is chosen for the final and stable version of the news. The presumption of this test is that the *most interesting* titles correlate with the

one chosen by the majority of the readers. However, while this can work well to increase readership and sales, it does not necessarily deliver the most interesting content. The question is not qualitative, but a quantitative one, since readers do not directly compare and evaluate the quality of two titles and choose the best one. In fact, the best title is the one that receives more clicks: *not the most interesting, but the one with the best performance*.

The headlines are not, however, just revolving doors that screen and count readers as they enter the story; they also carry discursive meaning in themselves (Andrew, 2007; Dor, 2003). A/B testing can be used to manipulate the audience to increase revenue reproducing the so-called “click bait,”² as well as to reinforce cultural and social biases. Indeed, what are the implications, biases, and tendencies that might emerge from a title selected by the first wave of readers? What happens, for instance, when a newspaper does an A/B testing at 7 am in Montreal when people in Vancouver are still asleep? Moreover, the participants-readers are not aware of the experiment, which can cause distortions in the way that an event is represented: readers cannot know if they are getting the right information or even the same information as other readers. It is possible that a series of blind tests and the use of black box algorithms in news practices can provoke serious consequences, including alternative ways to describe an event that reaffirms social bias and delivers what the audience want to see, not what the news is about.

Automated News

Algorithms are also used to automatize content creation, which can reduce and, in some contexts, eliminate human input in the process. While artificial intelligence can be programmed to generate automated reports based on quantitative data, machine learning can be used to build robots (or “bots”) that can “learn” from and make predictions based on collected data. Today, the technology necessary to replace reporters and editors is still in its infancy, but there are some experiments around this theme. For example, the Los Angeles Times developed Quakebot, an algorithm to produce simple reports on earthquakes in southern California (Walker, 2014). In 2014, the Associated Press was the first news agency to adopt software to produce more complex automated stories (Marconi, Siegman, & Machine Journalist, 2017). Consider also the case of Quill, developed by the start-up company Narrative Science. It can generate complete news stories based on raw data, such as sport events scores/stats, financial reports, or housing and survey data (Narrative Science, n.d.).

Indeed, these well-structured quantitative reports can easily be automated into the traditional structure of a news story. The summary news lead style, a common and widely-used method of writing news stories, is a highly structured and systematized way to deliver the most important information at the beginning of the news piece — who, when, where, why, what, and how are addressed in the story’s first paragraph. Developments in Natural Language Processing and Artificial Intelligence have already enabled computers to produce meaningful and readable information from chunks of text or datasets, just like its human counterpart would do.

These bots are programmed to find patterns in datasets and craft meaningful information to be distributed and consumed by the public. However, in principle, data has no particular quality or meaning; newsworthiness, as well as any other meaningful value that may be harvested from the dataset, is a social construct and has to be carefully inserted into the algorithm. In the case of Quakebot, the algorithms “knows” that “a magnitude less than 3.0 means it’s probably not worth freaking out about, a lesson that over-eager wire reporters don’t always grasp” (Oremus & Brogan, 2014, n.p). That is, the threshold of what is noteworthy is hard-coded in the system. While Quakebot can only work with the data generated and collected by the U.S. Geological Survey, Narrative Science (n.d.) claims that their software uses specific sets of data to answer questions about the

state of the world and frame stories that best suit the audiences. In other words, beyond simply reporting on the data, the machine interprets it for us.

What becomes evident by the rhetoric of the companies that produce these tools is the promise that these bots will write more comprehensive and objective stories than any human reporter. Not surprisingly, algorithmic media operates under the same assumption as more traditional media. It is perceived as objective and reliable to produce representations of reality, following the same values long attached to the institution of journalism. Additionally, these examples demonstrate technologies used to produce short reports on subjects that are highly quantitative, such as natural disasters, financial updates, and sports events. In the long run, however, the civic impact of such technologies may be more problematic. What happens when media companies start to use these methods to report on qualitative events that are harder to measure, or at least controversial by nature, such as political and social issues? What specific sets of data would interpret the world for us? It is important to understand who is going to define the parameters of these machines, based on what kind of epistemology and political ideology they will be built on, and for what purpose.

Shared Gatekeeping

In the digital age, the amount of generated data available to the reader has grown exponentially, but our capabilities to absorb this information have not increased. According to Bozdag (2013), it is precisely “because the mind’s information processing capacity is biologically limited, we get the feeling of being overwhelmed by the number of choices and end up with ‘bounded rationality’” (p. 211). Not surprisingly, news organizations began to use algorithms to assist audiences navigating the increasingly complex and fragmented media environment: search engine, recommendation systems, aggregators, data curation, and customization to facilitate search, navigation, and selection in a content overloaded environment (C. Anderson, 2006). To a certain extent, content has become commodified, with its true value residing in the system built to aid users to navigate and select from the profusion of available content (Napoli, 2014), and create and share more information.

Performing all of these actions manually without the help of technology is no longer practical nor effective. On the consumption side, news organizations, including those that insist upon avoiding such categorization (Herbst, 2016) such as search engines and social media, use algorithms in two non-exclusive ways: selection and visibility. These two interrelated functions often occur in a sequence and can be understood as a set of filters. Furthermore, they resonate the roles of journalism identified earlier: in response to a ‘problem of scale’ (Lesage & Hackett, 2014) it provides reports about meaningful and interesting events (Wolf, 1987).

Selection

It is commonsensical to think that with the increasing popularity of online media and social networks the audience gains greater control over news selection and can focus on issues that they consider more relevant, which in turn empowers audiences and reduce the degree of editorial influence over the news production. However, Bozdag (2013) affirms that the “gatekeeping process in online information services is more than a simple transition from editor selection to audience selection or from biased human decisions to neutral computerized selections” (p. 214). Online platforms, such as Google and Facebook, follow the same protocols developed by news organizations: selecting, organizing, distributing, and displaying information according to predefined and well-curated criteria. They usually do not provide equal channels for every user, and since they work as commercial

ventures following certain social and cultural ideologies, they are prone to biases as much as any other previous mass media channel.

In the selection stage, a search engine will automatically crawl the web, while a social network site will collect information produced by its users. It appears that the outcome — a collection of links and social media posts — is representative of reality. However, these results have a number of limitations, such as: off-line information is certainly absent; not all digitally-available information is accessible to the algorithm (*e.g.*, newspapers, and scientific journals paywalls); privacy protection in social media content can prevent the search engine from collecting information. Moreover, if a source has a bad reputation, like an illegal website, or has copyright infringements, it can be excluded or blocked from the search selection by the search engine owners. However, because information filtering is an automated process, it might be manipulated by third-party activities, such as the case of Search Engine Optimization (SEO) techniques, or “like farms” on Facebook (De Cristofaro, Friedman, Jourjon, Kaafar, & Shafiq, 2014). To avoid and inhibit these strategies, commercial organizations have a tendency toward internal process opacity, creating protocols to prevent others from “gaming” the system, to protect competitive advantages, as well as to promote their own services and products. This is precisely the reason that Google puts its own services at the top of the search results and that Facebook prioritizes its own video platform (Bucher, 2012; Bozdag, 2013).

Sorting and prioritizing this information is another crucial step, resembling the concept of news value (Wolf, 1987). Popularity is the most-used metric in classifying information: while a search engine gives more weight to information coming from popular websites in order to support majority interests and values, a social media platform prioritizes engagement with the content. A Google search result, for instance, only shows ten websites at a time and tends to favour popular sources of information which have more hyperlinks leading to these sites, are optimized for mobile phones, and are paid to be at the top of the list. Some authors, like Flichy (2008), have called this phenomenon “googlearchy” (Google + hierarchy). However, similar to the A/B testing, what the audience clicks the most is not always the most interesting or newsworthy piece of information. If in the past newsworthiness was defined by journalistic deontology, the opacity of algorithmic media makes it more complicated to define what is newsworthy in a fast, dynamic, and very personalized environment.

Visibility

The media industry is built around parameters of visibility (Bucher, 2012; Wolf, 1987). While printing press journalism carefully organizes physical spaces through a human editorial process, digital media revolves around algorithmically-tailored topics based on demography, ideology, commercial interest, and user preferences. That is, while in traditional media the likelihood that an event will be reported depends on the current public agenda followed by each news outlet (Wolf, 1987), in digital media, the audience has a role in defining this agenda: individuals and institutions mutually produce the media environment (Bucher, 2012). Napoli (2014) argues that this process leads to important patterns, such as a “certain amount of reflexivity that is inherent in much algorithmically driven media consumption” (p. 346). For instance, in Bucher’s (2012) examination of Facebook’s GraphRank algorithm, she shows that the algorithm monitors users’ behaviour 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” (p. 14).

Ørmen (2016) points out that technical settings like language and IP address matter when users are searching online. Digital platforms have access to a large number of different types of infor-

mation about their audience, which make it is possible to “target” individual readers. That is, to display or allow access to stories specifically tailored to users based on their locations, behaviour, and preferences. Consider the case of Twitter’s Trend topics that provides users with a list of the most popular subjects currently being discussed on the platform. This list is algorithmically tailored to each user based on their location, language, and the people they follow, as well as the topics’ “freshness”. The large number of variables results in a dynamism that can also lead to unpredictability of what is really a “trend” at any given time in the system, showing users only what they already expect to see. Ultimately, this raises questions of censorship, as in the controversy over the seemingly premature disappearance of the Occupy Wall Street movement from the Twitter Trends list (Gillespie, 2012). In this case, the term “trend” might even have lost its meaning, becoming more like a “click bait” based on the user’s bias, since the trends list distorts (and perhaps manipulates) the sense of reality, making people believe that some topics are more important than others in a given moment.

This is a major concern because of the influential role search engines and social media have today. They are perceived as objective and reliable representations of relevant online content, similar attribute long attached to journalism. In other words, “the algorithms that are at the core of search engines are functioning in a political capacity similar to established media institutions” (Napoli, 2014, p. 348). Not surprisingly, news organizations create and use algorithms that embed their institutional nature, including political and ideological inclinations. For instance, election results can be affected by manipulation of search engine rankings. It is a common practice now to look up candidates’ profiles using search engines and social media. Depending on how the algorithms rank positive and negative news stories on the screen, it can have an enormous influence on the way people vote: it is estimated that Google could determine the outcome of upwards of 25 percent of all national elections (Epstein & Robertson, 2015; Pfeiffer, 2014). These examples put algorithmic media protocols in resonance with traditional media practice in relation to the presence or lack of media coverage over specific issues within political contexts, which also reaffirms their role as gatekeepers.

Subjectivity

Decision-making is often permeated with idiosyncrasies due to the heavy involvement of human subjectivity in the various processes of production. Traditional media developed newsmaking protocols such as news value and newsworthiness to improve efficiency and pursue objectivity. Online web services frequently claim that such human bias does not exist in their systems, as their core operations are completely automated (Herbst, 2016). Yet, this is a problematic claim, since these systems are still driven by editorial decisions made by humans. While Google has admitted that the company manually demotes websites (Metz, 2011), Facebook has been accused several times of censorship due to its (opaque) policy of removing posts labeled as offensive, such as the case of a picture of women breastfeeding (“FB vs. Breastfeeding” group on Facebook), or two men kissing (Bozdog, 2013). To solve the centralized bias problem and avoid being classified as editors, these companies to make users accountable for these decisions. Facebook, for example, argues that it is not against breastfeeding in particular; it just follows what its users label as inappropriate (Facebook, n.d).

Participatory audience and crowd-sourcing are concepts usually coupled with decentralization of the mediation process and democratization of communication. Indeed, more participation can reduce the effects of technical bias induced by algorithms in online intermediaries (Bentes, 2015; Carlson, 2007), but they can also be used to enhance the “gatekeeping” effect. Media corporations

are heavily investing in “profiling” users to reach a higher level of personalization, which can reduce the structural biases inherent to popularity-based metrics. That is, by combining various sources of data, such as demography, location, technical information captured by electronic devices, personal preferences, and inferred behaviour derived from online activities, media companies can tailor and deliver the “most interesting” news according to the user’s definition of interesting.

However, this is quite opposite from reporting public interest stories: the importance to report on events that somehow impact the public life, including those that are unknown or exceptional (Bucher, 2012; Wolf, 1987). Algorithmic media have the potential to lock individuals in a “social bubble” in which people are exposed only to information from like-minded individuals, amplifying confirmation bias (Pariser, 2012). Or, as Bozdog (2013) puts it, “online services can cause citizens to be ill-informed about current events and may have increasingly idiosyncratic perceptions about the importance of current events and political issues” (p. 218). Personalization improves accuracy at the expense of serendipity. The problem is that in the long run, it will assist us to unconsciously avoid facts and opinions we disagree with, potentially undermining deliberative democracy by limiting contradictory information.

Conclusion

The news report is not simply a collection of random symbols; it obeys some sort of rule. Media organizations encode information on behalf of the individual based on routine decisions and according to a set of parameters (language, practices, goals). As Hall (1980) states, “the event must become a ‘story’ before it can become a communicative event” (p. 129). At that moment, the rules of society act upon the event; there are social and political consequences when an event is designated as important or interesting: production constructs the message.

Algorithms are not neutral and objective machines. Quite the opposite; they reflect a certain ideology embedding a set of rules and practices: they prescribe a way to do things. Like any other institution, news organizations are ideological and have specific social, cultural, economic, and political agendas. In this domain, algorithms serve as prime examples of constructors and codifiers of knowledge, particularly in the context of search engines, social media, and news portals, which play a central institutional role in aggregating, categorizing, organizing, and broadcasting information. Consequently, the adoption of algorithms in newsmaking practices by traditional and new media is not disrupting journalism practices and values, but rather reproducing them in the digital space.

Perhaps what is different about the algorithms adoption in the newsmaking practice is the level of automatization, since they are built to run without any human assistance. As a consequence, it seems that we are losing our ability and even our wiliness to inquire about the world. Algorithmic media reads data without asking any questions; it just repeats what is embedded in its codes. They rely on data and internal assumptions, both subject to biases and errors, which might produce unexpected and unintended consequences. Yet, they go beyond their assumed role and function of aiding in media production and consumption, as they serve as instruments of institutional power at the same time that they are shaped by the already-defined rules and protocols of the institutions that use them. In this sense, algorithmic media are nothing less than, and indeed no different from any attempt to formalize idealized newsmaking practices and protocols to establish the authority of hegemonic power to control flows of information.

Notes

1. I do not mean to argue that technology has absolute power over the user, but to emphasize the accelerated feedback loop in which digital technology is built upon. Users often do have some level of agency to actively avoid algorithmic logic, as well as to reappropriate, hack, and hijack technology in different ways through everyday practices (see De Certeau, 2002; Proulx, 2009).
2. Click bait is an advertising strategy to attract attention and encourage visitors to click on a link that usually has little or no connection with the content being linked. It is often paid for by the advertiser or generates income based on the number of clicks.

References

- Anderson, C. (2006). *The long tail: Why the future of business is selling less of more* (1st ed.). New York, NY: Hachette Books.
- Anderson, C. W. (2013). Towards a sociology of computational and algorithmic journalism. *New Media & Society*, 15(7), 1005–1021.
- Andrew, B. C. (2007). Media-generated shortcuts: Do newspaper headlines present another roadblock for low-information rationality? *The International Journal of Press/Politics*, 12(2), 24–43.
- Bentes, I. (2015). *Mídia - Multidão*. Rio de Janeiro, BR: Mauad.
- Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3), 209–227.
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14(7), 1164–1180.
- Carlson, M. (2007). Order versus access: news search engines and the challenge to traditional journalistic roles. *Media, Culture & Society*, 29(6), 1014–1030.
- De Certeau, M. (2002). *The practice of everyday life*. (S. F. Rendall, Trans.) (2nd ed.). Berkeley, CA: University of California Press.
- De Cristofaro, E., Friedman, A., Jourjon, G., Kaafar, M. A., & Shafiq, M. Z. (2014, October 4). *Paying for likes? Understanding Facebook like fraud using honeypots*. Retrieved from <http://arxiv.org/abs/1409.2097>
- Deuze, M. (2005). What is journalism? Professional identity and ideology of journalists reconsidered. *Journalism*, 6(4), 442–464.
- Dor, D. (2003). On newspaper headlines as relevance optimizers. *Journal of Pragmatics*, 35(5), 695–721.
- Dyer-Witheford, N. (1999). *Cyber-Marx: Cycles and circuits of struggle in high technology capitalism* (1st ed.). Urbana, IL: University of Illinois Press.
- Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences*, 112(33), E4512–E4521.
- Facebook. (n.d.). *Does Facebook allow photos of mothers breastfeeding?* Retrieved from <https://www.facebook.com/help/340974655932193>
- Flichy, P. (2008). *Is the Internet an instrument of democracy?* Retrieved from <http://www.booksandideas.net/Is-the-Internet-an-instrument-of.html>
- Franklin, U. (1999). *The real world of technology* (1st ed.). Toronto, ON: House of Anansi Press.
- Galloway, A. R. (2004). *Protocol: How control exists after decentralization* (1st ed.). Cambridge, MA: The MIT Press.
- Gillespie, T. (2012). Can an algorithm be wrong? *Limn*, (2). Retrieved from <http://limn.it/can-an-algorithm-be-wrong/>

- Gillespie, T. (2014). The Relevance of Algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: essays on communication, materiality, and society* (pp. 167–193). Cambridge, MA: The MIT Press.
- Hall, S. (1980). Encoding/Decoding. In S. Hall (Ed.), *Culture, Media, Language* (pp. 128–138). London, UK: Hutchinson.
- Herbst, J. (2016, April 13). The algorithm is an editor. *The Wall Street Journal*. Retrieved from <http://www.wsj.com/articles/the-algorithm-is-an-editor-1460585346>
- Latour, B., Jensen, P., Venturini, T., Grauwin, S., & Boullier, D. (2012). 'The whole is always smaller than its parts' – a digital test of Gabriel Tarde's monads. *The British Journal of Sociology*, 63(4), 590–615.
- Lesage, F., & Hackett, R. A. (2014). Between objectivity and openness—The Mediality of Data for Journalism. *Media and Communication*, 1(1), 39–50.
- Lewin, K. (1947). Frontiers in group dynamics: II. Channels of group life; Social planning and action research. *Human Relations*, 1(2), 143–153.
- Lippmann, W. (1997). *Public Opinion*. New Brunswick, NJ: Transaction Publishers.
- Marconi, F., Siegman, A., & Machine Journalist. (2017). *AP - The future of augmented journalism: A guide for newsrooms in the age of smart machines*. Retrieved from <https://insights.ap.org/uploads/images/the-future-of-augmented-journalism-ap-report.pdf>
- McKelvey, F. R. (2014). Algorithmic media need algorithmic methods: Why publics matter. *Canadian Journal of Communication*, 39(4), 597–613.
- Metz, C. (2011, February 18). Google opens curtain on “manual” search penalties. Retrieved from http://www.theregister.co.uk/2011/02/18/google_on_manual_search_penalties/
- Napoli, P. M. (2014). Automated Media: An institutional theory perspective on algorithmic media production and consumption. *Communication Theory*, 24(3), 340–360.
- Narrative Science. (n.d.). Retrieved from <https://www.narrativescience.com>
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy* (Reprint edition). New York, NY: Broadway Books.
- Oremus, W. (2014, March 17). The First News Report on the L.A. Earthquake Was Written by a Robot. Retrieved from http://www.slate.com/blogs/future_tense/2014/03/17/quakebot_los_angeles_times_robot_journalist_writes_article_on_la_earthquake.html
- Ørmen, J. (2016). Googling the news: Opportunities and challenges in studying news events through Google Search. *Digital Journalism*, 4(1), 107–124.
- Pariser, E. (2012). *The filter bubble: How the new personalized web is changing what we read and how we think*. (Reprint edition). New York, NY: Penguin Books.
- Pfeiffer, S. (2014, June 6). Could Facebook or Google manipulate an election? Retrieved from <http://radioboston.wbur.org/2014/06/06/zittrain-google-election>
- Proulx, S. (2009). Can the use of digital media favour citizen involvement? *Global Media and Communication*, 5(3), 293–307.
- Ramonet, I. (2013). A Explosão do Jornalismo na Era Digital. In M. Denis de & I. Ramonet (Eds.), *Mídia, Poder e Contrapoder - da Concentração Monopólica À Democratização da Informação* (E-reader version). São Paulo, BR: Boitempo Editorial.
- Shoemaker, P. J., Eichholz, M., Kim, E., & Wrigley, B. (2001). Individual and Routine Forces in Gate-keeping. *Journalism & Mass Communication Quarterly*, 78(2), 233–246.
- Singer, J. B. (2006). Stepping back from the gate: Online newspaper editors and the co-production of content in campaign 2004. *Journalism & Mass Communication Quarterly*, 83(2), 265–280.

- Soberman, J. (2013, August 15). Designing from data - How news organizations use A/B testing to increase user engagement. Retrieved, from <http://knightlab.northwestern.edu/2013/08/15/designing-from-data-how-news-organizations-use-ab-testing-to-increase-user-engagement/>
- Vanderbilt, T. (2013, August 7). The science behind the netflix algorithms that decide what you'll watch next. Retrieved from <https://www.wired.com/2013/08/qq-netflix-algorithm/>
- Walker, A. (2014, March 19). Quakebot: An algorithm that writes the news about earthquakes. Retrieved from <http://gizmodo.com/quakebot-an-algorithm-that-writes-the-news-about-earth-1547182732>
- White, D. M. (1950). The "gate keeper": A case study in the selection of news. *Journalism & Mass Communication Quarterly*, 27(4), 383-390.
- Wolf, M. (1987). *Teorias da Comunicação*. (M. J. V. de Figueiredo, Trans.). Lisboa, PT: Presença.