

EDITORIAL ANALYTICS: HOW A U.S. NEWSPAPER
APPLIES DATA TO MATCH TARGET AUDIENCES

A Thesis

presented to

the Faculty of the Graduate School
at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

DARIYA TSYRENZHAPOVA

Professor Mike McKean, Thesis Supervisor

MAY 2018

© Copyright by Dariya Tsyrenzhapova 2018

All Rights Reserved

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

EDITORIAL ANALYTICS:

HOW A U.S. NEWSPAPER APPLIES DATA TO MATCH TARGET AUDIENCES

presented by Dariya Tsyrenzhapova,

a candidate for the degree of master of arts,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Mike McKean

Professor Cristina Mislán

Professor Amy Simons

Professor Richard Johnson

This thesis is dedicated to my mother who through her love and care brings out the best in me.

ACKNOWLEDGEMENTS

This thesis would not be possible without the expertise, advice and support of my long-time mentors, friends and colleagues. I am forever grateful to Mizzou for letting me become more of myself in the fulfillment of my life goals and intellectual endeavors.

I am thankful to my thesis supervisor Professor Mike McKean who was the very first to support my interest in pursuing this topic and motivated me to execute it with novelty and skill. Professor Cristina Mislan had followed along from the very start by sharing her thoughtful expertise in critical political economy and Big Data. I commend Professor Amy Simons's industry-driven insights from her previous professional work experience as an editor at the Chicago Tribune. Thanks to the Media Entrepreneurship class taught by Professors Randall Smith and Richard Johnson who had inspired me to approach journalism from a standpoint of business sustainability.

My mentor Mike Jenner deserves a special recognition for helping me find a newsroom that not only shared their daily editorial practices with honesty and candor, but also opened access to their web analytics without contemplating that such information is proprietary.

My journey to this master's degree would not be complete without Professor Fritz Cropp and his fatherly support and encouragement all along. The same is true for my long-time college mentor Professor Berkley Hudson who continues to hold my hand as I move forward. And here is to Professor David Herzog for teaching me hard data skills that had put me on a new career path in journalism.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	iii
ABSTRACT	iv
Chapter	
1. INTRODUCTION	1
2. LITERATURE REVIEW	6
Actor-network theory	6
Newsrooms as actor-networks	8
Social platforms as actor-networks: Impact on news consumption	12
Audiences as actors	16
Audience data as a technological actant	22
Quantifying audience	25
3. METHODOLOGY	29
Sample	31
Procedure	32
4. FINDINGS	34
5. DISCUSSION	56
6. CONCLUSION	62
Limitations	62
Directions for future research	63
BIBLIOGRAPHY	64

LIST OF FIGURES

Figure	Page
1. Actor-network theory: A holistic matrix of “Four A’s”	9
2. Actor-network theory: Competing interests.....	9

EDITORIAL ANALYTICS: HOW A U.S. NEWSPAPER APPLIES DATA TO MATCH TARGET AUDIENCES

Dariya Tsyrenzhapova

Professor Mike McKean, Thesis Supervisor

ABSTRACT

This research is an in-depth case study of a major regional U.S. daily embracing audience data under the pressure of limited resources and shrinking advertising budgets. The legacy news operation observes analytics through “the rearview mirror,” believing that it is neither a true measure of audience engagement nor journalistic quality. Experiencing first-hand the implications of a commodified consumer attention, newsmakers find themselves conflicted between exercising their traditional role of public service and them leaning toward “soft” news to drive higher page-view revenue. By applying the actor-network theory as its central theoretical framework, the study addresses an intricate interplay of day-to-day editorial decision-making, Big Data analytics and the market economics of evolving digital news business.

Key words: audience engagement; editorial analytics; commodification of consumer attention; Big Data; journalistic role conceptions; actor-network theory; technology; journalism studies

INTRODUCTION

News organizations worldwide increasingly recognize the growing power of web and social media traffic data, which is intended to offer some insight into users' behavior. Without analytics software, embedded in newsrooms' daily routines and operations, who would have known that The New York Times' dialect quiz, which has generated more than 21 million page views, was "the most popular piece of content in the paper's history" (The New York Times Innovation Report, 2014, p. 37). The New York Times' 2020 Report, released in January of 2017, reaffirms the same trend that "the richest community engagement right now is mainly in nooks and crannies of the site": The Opinion section's The Stone, a discussion board on philosophical issues "both timely and timeless," the crosswords column, called The Wordplay, along with the stories of cancer survivors and the cooking recipe for chocolate chip cookies (The Report of the 2020 Group, 2017).

As the audience engagement metrics continue to evolve beyond just page views, the analytics tools cannot yet yield a comprehensive outlook into the audience's news consumption behavior. James G. Webster, the author of *The Marketplace of Attention*, argues the volumes of traffic data may provide "a fair approximation of what's commanding public attention, but [they] can rarely tell us why" (p. 88). Current metrics can only capture "traces of (parts of)" how people consume news online (Cherubini & Nielsen, 2016, p. 37). The absolute numbers of page views are almost meaningless: those "old metrics" do not fully capture the depth of social interactions as opposed to new, supposedly better, measures like "engaged time" and "loyalty" (Kormelink & Meijer, 2017). With that in mind, The

Financial Times takes a step further by using more elaborate measures, such as “time spent, recirculation, volume of articles read per visit, and number of comments” (Cherubini & Nielsen, 2016, p. 15).

The quality of data in digital analytics will continue to improve at a faster pace along with the growth of the global industry of Big Data and business analytics software, which is expected to reach \$187 billion by 2019. This is a more than a 50 percent increase from the \$122 billion in 2015; software development, alone, will generate more than \$55 billion, according to International Data Corporation estimates (Worldwide Semiannual Big Data and Analytics Spending Guide, 2016).

Media companies stand on the receiving end of technological innovations. But now more than ever newsrooms have strong business incentives to make sense of audience data. Zelenkauskaite (2016) describes news organizations as owners of “big data repositories” that are comprised of online multi-platform infrastructure (websites, mobile apps, pages on social media, etc.). They are updated live as new content is uploaded, and the audience begins to comment, tweet, like and share news online. But when content gets shared on social media as opposed to a news organization’s website, content distribution turns into a tradeoff between “reach” and “monetization” (Cherubini & Nielsen, 2016).

As of 2015, social media drove 31 percent of all referral traffic (DeMers, 2015). The 2017 Reuters Institute Digital News Report ranks the top three leading social networking sites (SNS) in the U.S. for news consumption with Facebook running the lead (48 percent +3), followed by YouTube (20 percent +1), and Twitter (15 percent +5) (p. 102-103). A decade after the launch of Facebook, social media

for the first time in 2016 had surpassed television, previously ranked as the No. 1 source of news consumption. A sample of 2,197 respondents in a survey by the Reuters Institute found that 28 percent of 18- to 24-year-olds now rely on social media as their primary source of news consumption as opposed to 24 percent still depending on television (Reuters Institute Digital News Report, 2016, p. 8).

Social media analytics now measure the success of television shows and the cash flow of advertising dollars across different programs: “a new market information regime” is defined by a new rating system – “social TV analytics” (Kosterich & Napoli, 2015, p. 254). As the news consumption experience is becoming increasingly social and participatory (Messing & Westwood, 2014), a traditional dichotomy between social media and news media is blurring. In fact, they are now co-dependent. Original stories are still generated by traditional media: 40 percent of Twitter’s traffic during peak hours is about television (Vanderbilt & Paskin, 2013). But Twitter has faced its first year-over-year decline in revenue: it did not grow its monthly user base in the second quarter of 2017 despite the high political climate during the Trump presidency (Heath, 2017).

News organizations remain as the primary content producers, but both Facebook and Google have mobilized their attention and resources around the news industry (Gibbs, 2016). They not only create “mechanisms by which audiences use media,” they also shape “representations of [that] media usage” (Napoli, 2011, p. 118). The central question is now about how technology is being used to facilitate user access and how to navigate audiences around content visibility settings,

powered by algorithms and machine learning. This becomes even more critical, especially in the culture of fake news.

The audience marketplace has always been “a contested space with a diverse array of institutional interests at stake” (Napoli, 2011, p. 118). It is no longer just about advertisers and media companies competing for public attention. The big social media entrants such as Facebook and Google that actually create new technology and have the ability to analyze Big Data streaming through their online platforms, now direct news organizations’ interactions with their audiences.

Online traffic translates into additional revenue streams, but news organizations’ growing tendency to chase online traffic may cause them to cover more soft news (Lee, Lewis & Powers, 2012; MacGregor, 2007; Sumpter, 2000). Therefore, within the framework of this study, it will be interesting to explore: 1) whether audience data affects the quality of editorial news judgment; 2) how that data guides editorial decisions in curating news content on social media and 3) whether audience data dictates story placement across various platforms (Singer, 2013; Lee et al, 2012). How does chasing for traffic factor into which news content type receives greater visibility across multiple news platforms?

The literature review will summarize scholarly and non-scholarly discourse analysis on news organizations’ evolving audience engagement practices that media professionals adopt through a process of constant negotiation and decision-making in their day-to-day interactions with analytics technology (Domingo, Masip, & Costera Meijer, 2015). Research on algorithms and user news consumption

behavior will inform and steer the discussion toward recognizing the agency of social media platforms in the public discourse.

By applying the actor-network theory (ANT) as its central theoretical framework, the study will trace the power dynamics between various human and non-human actors, actants and activities that shape the evolving journalistic professional practices, rooted in powerful cultural ideologies (Turner, 2005; Van Dijk, 2009). Using ANT as the vehicle to address “the organizational interplay of editorial, technology and business” will help to expand “a sociotechnical emphasis in journalism studies” (Lewis & Westlund, 2015, p. 21). Thus, this study will address the blind spots in previous communications research by “revealing nuances in the relationships among human actors inside the organization, human audiences beyond it, and the nonhuman actants that cross-mediate their interplay” (p. 21).

LITERATURE REVIEW

Actor-network theory (ANT)

Previous scholarly research exploring the impact of online analytics on editorial news judgment utilized the gatekeeping theory (Shoemaker & Vos, 2009) as the central theoretical framework (Deluliis, 2015; Singer, 2013; Sumpter, 2000; Tandoc, 2014; & Vu, 2014). Defined by a psychologist Kurt Lewin in 1943, it conceptualizes newsmakers as the primary decision-makers, knowing what becomes news, thus, deserves public attention.

In a survey of 318 editors often responsible for both print and online news operations, Vu (2014) had found that editors were willing to adjust their decisions based on web metrics (p. 1105). In fact, Singer (2013) had found that audience data affects story placement on the homepage. Although users seem to decide with their “clicks” which news story deserves greater public attention, editors still have the ability to tweak online visibility settings in order to highlight or to “de-select” a certain news item on the homepage (Tandoc, 2014).

Given the success of previous research, the gatekeeping theory could be a viable proxy to trace the impact of technology on news production, but it might suggest a rather linear approach that does not take into full consideration the power dynamics of the new technological forces, such as algorithms (Webster, 2014) and filter bubbles (Pariser, 2012) along with the audience’s changing news consumption behavior (Hermida; Fletcher; Korell, & Logan, 2012; McKenzie; Lowrey; Hays; Chung; & Woo, 2011).

First introduced by French science and technology scholar Bruno Latour in 1978, and then developed by sociologists John Law and John Hassard in the early 1980s, the ANT (“*acteur reseau*” in French) has drifted from sociology to science and technology. The theory denotes a network of relationships and associations among various actors and actants that are assembled and maintained to achieve a specific objective (Law & Hassard, 1999). Most recently, the ANT has entered the realm of communications research to theorize news production practices (Turner, 2005), the effects of digitization and Information and communications technology (ICTs) on journalism (Plesner, 2009), the rise of social media platforms as distinct “microsystems” in the social web (Van Dijck, 2013), production of television programming (Teurlings, 2013), the diffusion of mobile live-streaming apps (Stern, 2016), etc.

Theories have limitations, and the ANT is no exception. While recognizing the power dynamics “emerging from the activity in the network,” it disregards the pre-existing power relations that are critical for understanding legal and economic conditions that contribute to the current state of digital market infrastructure (Van Dijck, 2012, p. 151). However, this question traditionally lies in the theoretical scope of political economy, which is not a direct research interest of this study. In other words, instead of looking at production as “an economic process determined by property relations,” this study will rather take an epistemological approach to conceptualizing the networks of actors and actants in the newsroom as social constructs that are being maintained through a set of values and standards of professional news practice (Teurlings, 2013, pp. 102-103). Following the cultural

economy tradition, which blends together political economy and cultural studies, the ANT will serve as a proxy to inform a scholarly discourse about production practices specifically within media organizations that by default have “the explicit aim of ‘improving’ the populations” by shaping and reinforcing existing public ideologies (Teurlings, 2013, p. 102).

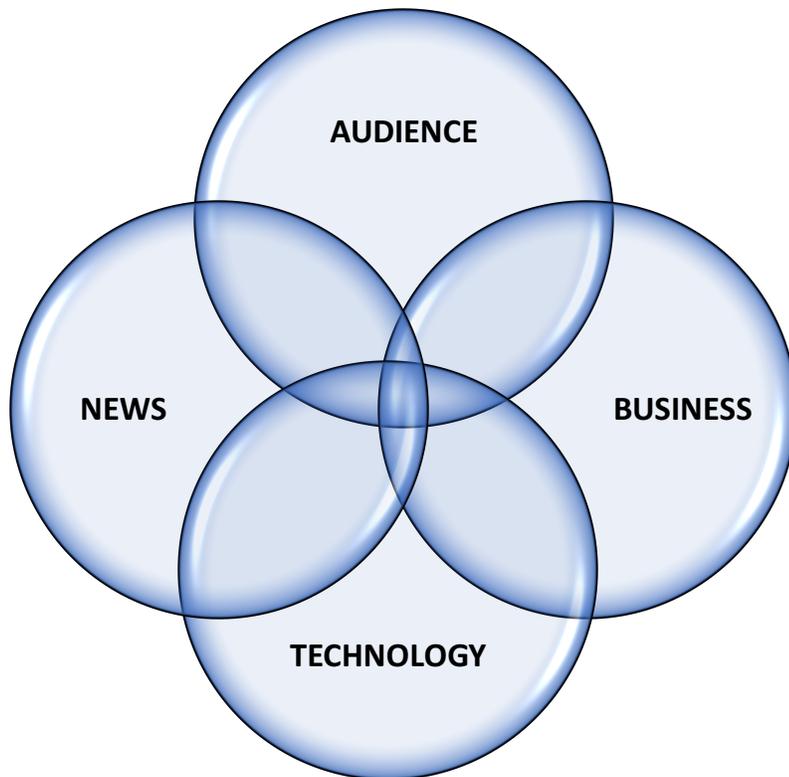
Newsrooms as actor-networks

The concept of *movement* is essential for theorizing a network of relations and associations among actors and actants. To trace news as *a discourse*, the ANT applies “news networks” as its primary analytical tool to unpack a full set of professional *practices* (news production and circulation) that are shaped by “professional ideals,” “symbolic constructions [of] newsworthiness, shares, ratings.” The distribution of those messages occurs via “material artifacts,” such as news websites and social media platforms (Domingo, Masip, & Costera Meijer, 2015, p. 56). Those news activities are understood as “effects” and “outcomes” that are accomplished through the continuous performance of relations among different actors and actants (Law & Hassard, 1999, p. 4).

Lewis & Westlund (2015) apply a holistic matrix of “Four A’s” to bring actors, actants, activities and audiences under the same theoretical umbrella. (See **Figure 1**: *Actor-network theory: A holistic matrix of “Four A’s”*).

ACTORS	Newsroom & Audience
ACTANTS	Technology
ACTIVITIES	News

Acknowledging the social construction of technology in a cross-media news work, the model points at the power dynamics in an interplay of editorial, technological and business interests. (See **Figure 2: Actor-network theory: Competing interests**)



Within the context of media organizations, actors in the newsroom are interconnected by a series of non-human actants, such as data, technology, cultural norms and professional practices. Within the framework of this research, journalists

and editors belong to the primary group of actors whom Weiss & Domingo (2010) also define as professional “communities of practice,” or Zelizer (1993) describes as “interpretive communities” that share collective interpretations of reality based on their shared norms of news values and professional journalistic standards. Aside from the traditional newsmakers, another important set of actors consists of audience engagement editors, technologists, and business people who begin to earn greater prominence in the newsroom as the journalism industry adopts “audience-centric” values; thus, more heavily relies on audience tracking technology and Big Data.

While journalists and editors alike increasingly try to engage in the act of social listening in order to tap into consumers' needs and preferences, news organizations still find it hard to overcome the existing “news gap” (Boczkowski & Mitchelstein, 2013), or an ideological dilemma between “what audiences want to read and what journalists feel they should read” (Lee, Lewis, & Powers, 2012, p. 519). A comparative textual analysis of the top 10 editors' and readers' stories from NYTimes.com, NYPost.com, and NYDailyNews.com found that editors prefer “hard” news while their audiences gravitate to “soft” news (Lee et al., 2012). This finding points to an interesting facet of audience behavior that news organizations can no longer disregard.

The essential strength of the ANT in conceptualizing this phenomenon is twofold. It challenges institutional functionalism (Couldry, 2008) and technological determinism (Weiss & Domingo, 2010). Moving away from a rigid dichotomy of producers and consumers, the anti-functionalist approach recognizes audience's

growing power to create and circulate news across various platforms. The denial of technological determinism stresses the editorial agency in deciding what is newsworthy; thus, commanding public attention. The editorial decision-making is traced by means of material artifacts (news organizations' websites and mobile apps) and reflected in editorial routine practices of changing story placement and tweaking headlines, as the analytics software may suggest. In the meantime, these traces, left behind by the actors' activities, are important for researchers to follow in order to better understand how powerful "occupational ideologies" of newsworthiness (Deuze, 2005, p. 443) and agenda-setting (McCombs, Shaw, & Weaver, 2014) are being persistently reaffirmed and reinforced in "heterogeneous actor-networks" (Plesner, 2009, p. 605).

Most recently, Facebook has transformed into "a non-traditional media company," as the company CEO Mark Zuckerberg, himself, alluded to in a video chat with the company's COO Sheryl Sandberg in December of 2016, by saying: "You know, we build technology, and we feel responsible for how it's used" (Gibbs, 2016). Therefore, within the ANT framework, it is reasonable to define Facebook as a non-human actor, instead of labeling it as "a technological actant." Van Dijck (2013) sees social platforms as "sociotechnical ensembles and performative infrastructures" (p. 26) that "create and mediate a new type of social capital: connectivity" (Van Dijck, 2012, p. 150). Facebook does exemplify a new type of actor that over time has transformed from an intermediary into "a producer of sociality," "a mediator" of public discourse (Van Dijck, 2012). The platform's algorithms propels the business of "directing human sociality" (Van Dijck, 2013, p. 13-14).

To expand our understanding on how actors in the newsroom perform their relations with other actors and actants technologically embedded in news networks, Couldry (2008) applies a concept of “online liveness” – understood as an extension of traditional liveness across media – could be used as an analytical tool to spotlight the patterns of how real-time audience data leads editors to tweak headlines and adjust story placement across news platforms to increase traffic. This approach will fulfill Van Dijck’s (2013) suggestion not only “to analyze technology, users and content in close alignment,” but will also validate his argument that social media platforms are “techno-cultural constructs,” exercising a higher level of agency (p. 28).

Social platforms as actor-networks: Impact on news consumption

The term “actants” usually refers to “any actor, human or non-human that is engaged in a networked system” (Weiss & Domingo, 2015, p. 23). Actants could be placed either internally or externally in relation to a newsroom, undertaking the action under scrutiny. The multiplicity of those actants involved in news production, circulation and audience reception, highlights an interconnected tension between symbolic social constructs (journalistic texts) and artifacts (the evolving metrics for evaluating audience reception), traveling in a technological environment that is regulated by algorithms (Webster, 2014) and filter bubbles (Pariser, 2012). Through the lens of the ANT, Plesner (2009) interprets the Internet as “a huge storage space for informational traces” (p. 615). However, this vision falls short in addressing the distinct force of machine learning, which has the technological power

to determine what news content we consume and when we should consume it (Pariser, 2012).

Webster (2014) defines an algorithm as “a computer program that runs data through a series of instructions or decision points to reach some ideal result” (p. 84). As for being human creations, they are prone to “behavior,” “personalization” and “popularity” biases (p. 85). By analyzing audience logs for 22,000 Facebook users’ posts, Bernstein, Bakshy, Burke, & Karrer (2013) found that, on average, only 35 percent of Facebook users actually see a status update posted within their network of friends and followers. With this knowledge in mind, *The New York Times* now persistently scrutinizes social media algorithms in order to ensure that their award-winning journalism reaches its audiences (The New York Times Innovation Report, 2014).

Algorithms are real, and they perpetuate personalized “filter bubbles,” a concept famously advocated and explained by internet activist Eli Pariser (2012):

First, you are alone in it [filter bubble]. Second, the filter bubble is invisible. Google doesn’t tell you who it thinks you are or why it’s showing you the results you’re seeing. You don’t know if its assumptions about you are right or wrong – and you might not even know it’s making assumptions about you in the first place. (p. 9)

Given that personalization is the key strategy for the top-five social platforms, Yahoo, Google, Facebook, YouTube and Microsoft Live, users’ browsing history becomes even more critical because it determines what content becomes available first and how it is packaged together (Pariser, 2012). Similarly, Google

uses users' location, past clicker-behavior and browsing history to bundle content into the top search results. In fact, this is a constantly evolving practice because Google changes its search algorithms 500 times a year (Webster, 2014). Thus, following Van Dijck's (2013) argument that the social web is "an ecosystem of connective media," it is clear that any "changes in other parts of the ecosystem," provoke action within the network of actors: "If Facebook changes its interface settings, Google reacts by tweaking its artillery of platforms; if participation in Wikipedia should wane, Google's algorithmic remedies could work wonders" (p. 21). The sophisticated "platform tactics" guided by algorithms should not be discounted while assessing what information the web feeds its users in top searches or in trending topics (Van Dijck, 2013, p. 21).

Pointing to similar implications, Joseph Turow (2011), the author of *The Daily You*, talks about "reputation silos" that are being constructed as a result of data mining users' profiles: "If the receivers of those customized mosaics of the world do notice consistent patterns in their worldviews, they will have no way of figuring out why the specific tiles were presented to them" (p. 137). The consequences could be far-reaching and even detrimental to one's information diet: from something as relatively benign as receiving customized ads to having articles that fit users' personal interests appear in the prime spots of a news site's homepage. That same logic can play out the other way around when certain types of news stories that a person has hardly clicked on or spent little time on eventually either disappear from the view on the homepage or are being pushed down the screen from the user's main view (Turow, 2011, p. 126-127).

Imagining the human mind behind a computer screen, it becomes obvious that none of the algorithms are “value-neutral” (Fairfield & Shtein, 2014, p. 47). The ultimate goal of algorithms and machine-learning technology is to optimize, predict and present media that would be “a perfect reflection of our interests and desires” (Pariser, 2012, p. 12). But in reality, “the Internet’s distorted picture of us” recreated under the control of algorithms can reshape individuals’ identities by feeding them a media diet, filtered in an algorithmic fashion. Because technology has no way to distinguish the true intent and motivation behind users’ clicking behavior, whether it is driven by “compulsion” or “general interest,” personalized filters are likely to generate nothing more than “compulsive media” that attract most clicks and page views (Pariser, 2012).

The fact that Google News’ feed and Facebook’s trending topics are both “surfaced by algorithms” (Tufekci, 2016) is already old news. Personalization technology in the publishing industry is on to the next level: NBC, BBC, *The Washington Post*, and Quartz have been experimenting with AI-enabled chatbots that upon request can search for “the most important news story” of the day and deliver it instantly through mobile messaging services (Barot, 2016; Mitchell & Holcomb, 2016). This new technology is even more nuanced and sophisticated from what Pariser (2012) has imagined – when regular news sites deliver a personalized homepage with fully customized news content re-arranged in a particular fashion to fit someone’s supposedly personal interest.

Audiences as actors

News organizations can approach audiences simultaneously as “recipients, commodities, and active participants,” which at the same time denotes “normative, commercial and cultural functions” (Lewis & Westlund, 2015, p. 26-27). Recognizing its growing participatory nature and the power to influence editorial decision-making, the ANT could be used as a framework to understand how audience’s online behavior shapes daily (inter)actions and practices within news networks (Domingo, Masip, & Costera Meijer, 2015). Therefore, audiences legitimately could be seen as an evolving set of actors that media organizations are increasingly trying to be mindful of.

In their attempts to quantify audiences, media organizations rely on a notion of an “institutionalized audience” (Napoli, 2011). The implications of how news organizations “perceive” and “imagine” their audiences (Loosen & Schmidt, 2012; Bernstein et al., 2013) are rooted not only in the economics of the media business, but also in editorial decisions about what becomes news. According to one analysis, journalists as part of their occupational ideology rely on five “ideal-typical values:” “public service, objectivity, autonomy, immediacy, and ethics” (Kovach & Rosenstiel, 2007). In a national survey of 546 U.S. journalists and editors, Holton, Lewis & Coddington (2016) contend that conceptualizing professional roles through the traditional ideals of public service is rather prevalent among journalists working for legacy news organizations as opposed to their colleagues in digital-native media outlets.

A study by Harcup & Neill (2016) found that for a news story to be selected

for publication, it needs to satisfy at least one of these 10 requirements: a story about the power elites, celebrities or entertainment; contains some element of a “surprise,” “good news,” “bad news,” “magnitude,” “relevance,” “follow-up” or fulfills a news organization’s agenda (p. 2). Such powerful ideologies, or “mental representations” of the public’s wants and preferences are being “reproduced [and reinforced] by text, talk and communication” (Van Dijk, 2009, p. 192). McChesney (2013) takes the discussion a step further and argues that media firms have no incentives “to upgrade the taste” of their audiences because “they take the market as they find it” (p. 76).

The participatory nature of news consumption via social media has shifted news organizations’ established notions of audiences and their collective roles in the public domain. Recognizing audiences’ both “participatory” and “generative” nature (Anderson, 2011, p. 557), news organizations have re-conceptualized “audience members” into “content providers” (Napoli, 2011). The traditional dichotomy of news production and consumption rooted in a paradigm “we select and we write, you only read” has shifted toward “...we select, we write, we read, and we share” (Renés-Arellano, Hernandez-Serrano, Graham & Greenhill, 2017).

Most recently, as Big Social Data began gaining traction in the news industry, audiences are seen as those generating viewership figures and user data (Cockayne, 2016, p. 2) that is an object of commodification (Turow, 2011). In fact, audiences simultaneously could be seen as recipients of information, commodities, and active participants, depending on what value news organizations ascribe to them (Lewis & Westlund, 2015).

The Big Data industry with its infrastructure geared toward quantifying audiences and their behaviors propels and amplifies the capitalism monopoly that media organizations, audiences and advertisers participate in. The notion of seeing audiences as commodities is not novel. Dallas Smythe's (1981) "On the Audience Commodity and its Work" lays out a conceptual framework for assessing the value of "audience power" as a distinct commodity that is being "produced, sold, purchased and consumed" (p. 233).

Bringing in a concept of "a free lunch" metaphorically equated to non-advertising content by A. J. Liebling, Smythe (1981) highlights the media's intermediary role in contributing to "capitalism monopoly." Embracing a critical political economy perspective, he concludes that media artifacts – "messages," "information," "images," "meaning" – do not constitute an actual product, but rather "its effects, or purpose" (p. 231).

According to Smythe, the commodity is audience power. To access it, advertisers pay almost like "an opportunity cost." For buying bulk, the price is set at a discounted rate, which compensates for those audience members might not get exposed to the message. It is a cost-efficient bargain based on probability and the law of large numbers that yields profit.

The way media professionals think of their audiences affects the decisions behind what eventually becomes news (Turow, 2011). The growing power of algorithms makes it even more difficult to discern what actually steers the news coverage: traffic data guided by algorithms or collective journalistic standards of newsworthiness, cultivated in a newsroom culture? The latter are shaped by

powerful institutional actors and their routine professional practices, continuously cultivated by the fluid journalistic norms and standards of newsworthiness. In fact, that “lack of fixity at the core of this conception of culture” is “the origin of [its] power,” but its dynamics could be traced in the day-to-day editorial decision-making practices (Cohen, 2012, p. 17).

Couldry (2011) suggests that seeing audience *as practice* provides theoretical grounds to tap into “how people understand what they are doing, and what ‘practices’ their actions comprise” (p. 217). The idea of “audience *as practice*” extends a scholarly discourse in the direction of audience’s news consumption behavior, including broader issues, such as preferred platforms for accessing news content as well as habits and daily rituals associated with consuming news online.

In a constantly changing digital environment, social media platforms are critical “anchoring device[s] in the media field” that denote “the act of being part of an audience itself, as a form of ‘necessary’ attention” (Couldry, 2011, p. 218). Online participation has hardly been quantified and is yet open to debate. However, rough estimates mentioned in (Napoli, 2011) provide some insight into how people interact with media online. From the entire online population, he (she?) says 90 percent are best described as “lurkers”; nine percent occasionally engage in some sort of online participation and creation of user-generated content (UGC), and only the remaining one percent actively “engage in online participation and content creation” (p. 106). The combined 10 percent of the audience sustain “a sharing economy” by “creating and sharing media” (Webster, 2014, p. 52).

Hierarchically, out of the four levels of news engagement, defined by Ha et al. (2016), the act of sharing news content on social media belongs to the third tier, preceded by the basic linear news consumption and relying on news from diverse media sources. The fourth (highest) level denotes one's continuous commitment to producing citizen journalism (p. 3-4). However, this is a rather schematic vision, which does not grant any depth to a comprehensive understanding of news consumption behavior online.

According to Ha, Xu, Yang, C., Wang, Yang, L., Abuljadail, ... & Gabay (2016), a broad definition of news engagement implies "the effort made in *obtaining* and *utilizing* the news content among the audience" (p. 3). Batsell (2015) expands this definition by stressing the idea of creating and promoting "engaged journalism" – a product purposefully integrated across platforms to further a media company's primary journalistic purpose and its financial mission.

Understanding the patterns and rituals surrounding audience's media consumption behavior are key to solving a puzzle that goes beyond just capturing the eyeballs and extends further toward "harness[ing] [audience's] collective intelligence" and "tap[ping] their capacity to circulate messages" (Green & Jenkins, 2011, p. 126). The "many-to-many model" reveals the audience's agency in "spreading media" within an established network of friends and followers that directly competes with news organizations' ability to reach audiences (Green & Jenkins, 2011).

A poll of 1,600 Canadians shows that two out of five (43 percent) of those surveyed by Hermida et al. (2012) receive their news from their Facebook friends

and followers while only one-fifth rely on media companies and individual journalists whom they follow on social media. The remainder is split between 18 percent who receive Twitter updates from friends and followers, 10 percent who rely on tweets of news organizations and individual journalists, and nine percent who marked their responses as “unsure.” The habit of consuming and sharing news on social media based on recommendations of your immediate connections, significantly reduces the need to seek out news via search engines or by visiting the homepages of different news organizations.

Content popularity on social media is often equated with the sheer number of “likes” and “shares.” Those measures, however, are not backed up by the “quality assessment” of actual user preferences because “likes” or the other five Facebook emoji reactions launched in February of 2016 (Chowdhry, 2016) can only favor “instant, gut-fired, emotional” evaluations (Van Dijck, 2013, p. 13). Moreover, “popularity as a coded concept,” according to Van Dijck (2013), is subject to algorithmic manipulation. For instance, Facebook’s updated News Feed algorithm will have the capabilities to downgrade web pages that load slowly on mobile (Perez, 2017). In the name of accessibility and better user experience, this measure aims to encourage publishers to partner with Facebook’s Instant Articles mobile service, which is designed to keep users in the native Facebook app without the real need to visit the site of the content’s source. This highlights Facebook’s strategy of moving from “a gateway” to becoming “a destination – a sit-down restaurant, not a take-out counter” (Constine, 2015). Twitter’s Moments and Snapchat’s Discover sections use a similar logic and rationale (Segall & Stetler, 2015).

In an active digital environment with a very competitive marketplace for attention, technology and user agency are “inseparable” (Van Dijck’s, 2013, p. 32). Google and Facebook not only create “mechanisms by which audiences use media,” they also shape “representations of [that] media usage” (Napoli, 2011, p. 118). Thus, audiences more rarely consume media as “unified media products or programming flows”: it is rather individual articles, songs and program episodes often pushed through social media (Turow, 2011). This shifting paradigm of audience interaction with news content, reinforced by the culture of “social sharing,” raises questions not only about media organizations weakening gatekeeping function, but also brings in concerns about “limited exposure to a variety of news sources” (Hermida, et al., 2012, p. 822).

Audience data as a technological actant

Technology propels media innovations. Weiss & Domingo (2015) describe newsrooms’ content management systems (CMS) as “technological actants” that are “encoded with journalistic news values” (p. 24). The software is developed and further directed by humans to fulfill journalistic purposes of storytelling, content monetization and multi-platform distribution. Although technology overwhelmingly defines and sustains journalistic workflow, editors still firmly believe that it cannot substitute for human judgment, but it can certainly help them simplify complex tasks. Despite the wealth of traffic data available throughout a 24-hour news cycle, MacGregor (2007) asserts that news organizations use audience data only as a means to double-check their editorial instincts. A *Financial Times* staff member, who

was interviewed among 18 other journalists in MacGregor's qualitative study, describes the ongoing dilemma: "Sheer adherence to chasing user numbers or numbers of clicks is criticized also for loss of editorial independence, or "slavery" to the audience. The notion of following clicks is so crude as to be 'dangerous'" (p. 291).

Blom & Hansen (2015) define clickbait as a form of forward-referencing headlines to exploit users' curiosity and lure them into "soft" news content. Clickbait headlines generate additional traffic to the site, which translates into commercial revenues, but they negatively affect source credibility (Hurst, 2016). From a technical standpoint, because clickbait content, by nature, has high Click Through Rates (CTR), it has a higher position in the search stream. A regular user might not differentiate it at first glance from quality journalism (Biyani, Tsioutsoulouklis & Blackmer, 2016).

Changing industry dynamics force news organizations to revisit their traditional institutional role as gatekeepers (Shoemaker & Vos, 2009) and agenda setters (McCombs, Shaw, & Waeber, 2014). Over time, their collective editorial decisions, routinely guided by a shared understanding of news values, ultimately have shifted toward recognizing audiences' *needs* and *wants*. This reflects *The Financial Times'* new goal of transforming the newsroom into "audience-first journalists" as defined by their Head of Audience Engagement (Cherubini & Nielsen, 2016). The goal is to streamline "engagement objectives into the commissioning and production process," which also involves posting stories during peak traffic hours (p. 12). The latter vary for each social media platform: Facebook has slots at 9 a.m., 1

p.m. and 3 p.m. while Twitter's peak cycles start at noon, followed by 3 p.m. and a window between 5 and 6 p.m. (Kolowich, 2017).

This new approach leans on technology and is rooted in an organizational structure and newsroom culture that employs data-informed decision-making through tailored editorial analytics tools. Part of the move toward conquering the audience marketplace is to use tools like Chartbeat (65 percent), NewsWhip (15 percent) or Parse.ly (12 percent) that have been designed with the editorial priorities in mind, or even relying on tailor-made home-grown analytics systems, as indicated by 45 percent of editors and news media CEOs, surveyed by the Reuters Institute in 2016. Both CNN and AOL have in-house software, which measures a percentage of clicks in real time, so does *The Guardian* that has been relying on online traffic data technology since 2005 (MacGregor, 2007). However, unlike the more generic tools – Google Analytics, Facebook Insights, Adobe Marketing Cloud (Omniure – prior to 2012), Twitter Analytics – those journalism-specific software packages allow the publisher to modify the structure of the homepage in real time; apply A/B headline testing; follow trending topics on social media, and even identify news stories users responded to in the past and what they read next (Cherubini & Nielsen, 2016, pp. 24-27).

Communication scholars today are still in the early stages of exploring and discerning the real impact of user behavior on editorial decisions (Vu, 2014; MacGregor, 2007; Lee, Lewis, & Powers, 2012; McKenzie et al., 2011). Partly, this is because the phenomenon of real-time audience data, itself, is relatively new. For companies that understand the value of mining and decoding online analytics, Big

Social Data is “the new oil,” according to Derrick Harris, a technology writer at Gigaom.com. This analogy, however, mainly applies to larger news organizations that make the extra effort now to figure out the “ins” and “outs” of news personalization technology. For those media with far fewer resources allocated to social media outreach, online analytics is more like “gasoline,” which sustains their online presence in a competitive media environment (Harris, 2015).

In their constant tries to disrupt a print-centric mindset, media organizations – both large and small – are making a substantial effort to rethink their organizational and editorial workflow. For instance, at *The Huffington Post*, one of the most successful digital-only start-ups in recent years, “the article **begins** (emphasis added) its life when you hit publish,” says co-founder Paul Berry (The New York Times Innovation Report, 2014, p. 24). Thinking through SEO and storytelling and making sure that a story is published at the right time for the intended audience has become a part of “a digital hygiene,” said Carla Zanoni, the Executive Emerging Media Editor at *The Wall Street Journal’s* audience team (Cherubini & Nielsen, 2016, p. 29).

Quantifying audience

Online analytics present “an even more powerful form of quantitative journalistic decision making” that is also increasingly becoming audience-centric (Anderson, 2011a, p. 536). Audience metrics are gradually transforming into a distinct technological actant that actors in the newsroom use, on one hand, as a radar to “prove their public relevance” and to back up their professional legitimacy

in news coverage and on the other, as a signal, or warning that important stories did not reach the audience. Editors can then take further action (Karlsson & Clerwall, 2013).

Having received less scholarly attention, audience measurement systems represent a distinct type of innovation that shapes “a new representation of the audience marketplace; a new ‘market information regime’” (Napoli, 2011, p. 121). Mainly the challenge is centered around “definition,” “measurement,” and “data quality” (Cherubini & Nielsen, 2016, p. 40), which points to the fact that any metrics, old or new, are “human creations” and are subject to “bias and abuse” (Webster, 2014, p. 94).

Scholars and many practitioners say that page views, for example, “are flawed” and “meaningless” in terms of capturing audience’s keen interests and preferences (Kormelink & Costera Meijer, 2017). In fact, there is a spectrum of “cognitive,” “affective” and “pragmatic” considerations that propel users to click or not to click on a news story; therefore, editors’ strict reliance on page views has the potential to harm the democratic value of journalism and to trivialize news as a professional practice.

Analytics companies and news organizations alike revisit metrics to keep up with advancing new media technology. According to Batsell (2015), Chartbeat redesigned a dashboard that allows publishers to monitor their site visitors by segments – “new,” “returning,” and “loyal.” Cherubini & Nielsen (2016) have outlined a list of 18 analytics that cover a wide range of measures: from the rudimentary ones, like page views that technically signify a mere act of “hitting

news items or links for more information” (Meijer & Kormelink, 2015, p. 672), to more sophisticated ones, such as “scroll depth” and “recirculation” (p. 34-35). This reflects a subsequent move toward more complex “exposure-based dimensions” that encompass “loyalty and attentiveness” (Napoli, 2011, p. 90). Internalizing that same trend, news aggregator Upworthy.com in 2014 adopted so-called “attention minutes” (Batsell, 2015). A year later, *The New York Times* began ranking its most popular stories by the total combined time spent on a news item instead of tracking the number of unique visitors (Cherubini & Nielsen, 2016, p. 36).

These metrics are not perfect, nor is the data itself, which is generated to fit the existing analytics architecture. According to *The New York Times’* audience insights group under Laura Evans, a new, more sophisticated alternative to page views could “measure an article’s value to attracting and retaining subscribers” (The Report of the 2020 Group, 2017). But as “an object of commodification and consumption” (Puschmann & Burgess, 2014, p. 1691), the industry of Big Data is persistently driven by commercial and advertising priorities with little consideration toward journalism. News organizations are in the midst of trying to understand how to integrate metrics from different platforms (websites, mobile apps and social media) along with translating that use across multiple devices (Cherubini & Nielsen, 2016). Based on the traffic data, it is difficult to link users to specific demographics without knowing their login information. It is also hard to track whether the conversion from regular to loyal user and subscriber makes one more informed or engaged (Cherubini & Nielsen, 2016).

RESEARCH QUESTIONS

The study posed the following four research questions:

RQ1: How do employees of an innovative news organization perceive their relationship with audiences both in financial and journalistic terms?

RQ2: How do real-time audience data and a reliance on social media relate to those perceptions?

RQ3: How do those perceptions relate to editorial decision-making?

RQ4: What kinds of audience data might effectively measure journalistic concerns such as learning and citizen participation?

METHODOLOGY

Using the methodology of semi-structured interviews (Sumpter, 2000, MacGregor, 2007; Anderson, 2011; Usher, 2013; Tandoc, 2014), the study collected industry insights about how a legacy print news organization applies editorial analytics to better serve their every-day readers with timely quality news content as well as to continue growing its existing audience base that is a lucrative asset for advertisers.

An in-depth case study was conducted at a major regional daily, which ranks in the top-30 largest newspapers nationwide. Its corporate owner runs nearly 50 dailies nationwide, reaching an audience of more than 30 million a month. The studied newspaper accounts for a 23 percent share of the total audience for the company's publications. The insights collected over nine hours of interviews mirrors a narrative that is prevalent among traditional metropolitan U.S. dailies, trying to make sense out of audience data under the pressure of limited resources and shrinking advertising budgets.

Since the aforementioned studies were conducted, the digital environment where news businesses operate had changed significantly. Social media platforms had transformed into powerful technological actors, directing media usage and audience attention alike. Traffic data is no longer just a barometer of audience engagement. It is a commodity sold for higher advertising dollars. The studied newsroom provides an example of an enduring dilemma between two competing interests – audiences and advertisers. Interviews with the primary actors in the newsroom – both on the production and editorial side – reflect that ongoing tension.

Unlike surveys in studies by Singer (2013) and (Vu, 2014), using interviews as the primary research method allowed the researcher to collect the necessary detail and depth of an insider's perspective while giving an opportunity to "test the meaning of responses" (Whyte, 1982, p. 114). The interviews offered participants a fair amount of flexibility and freedom to express their vision and to introduce story examples, backed up by data from their analytics dashboards. Also, asking several informants the same set of questions, as suggested by (Berger, 2000), helped to trace common patterns in their perceptions of journalistic values of newsworthiness and how they differ from audience's needs and wants.

Methodologically, previously cited studies often supplemented interviews with elements of newsroom ethnographies. For instance, Anderson (2011) spent more than 300 hours of observation and conducted 60 semi-structured interviews with journalists, editors, bloggers and activists from three Philadelphia-based newspapers to show how web analytics increasingly supplements editorial news judgment. However, for the purposes of this research, the possibility of conducting a newsroom ethnography was constrained by access and the limited availability of each interviewee during the scheduled two-day newsroom visit.

Qualitative studies on a similar issue (Sumpter, 2000; Usher, 2013; Tandoc, 2014) drew findings from at least a dozen interviews conducted in each newsroom. According to Guest, Bunce, & Johnson (2006), 12 is the optimum number of interviewees, sufficient for accomplishing "theoretical saturation" – when any new information collected through another interview becomes insignificant to the overall findings.

To accomplish “data triangulation,” which aspires to “get a ‘true’ fix on a situation” (Silverman, 2013, p. 287), or what Whyte (1982) calls “a fairly approximation of reality” (p. 117), the interviews were analyzed in reference to the newsroom’s audience analytics, which offers insight into real-time as well as historical data, going several years back. Industry reports allowed the researcher to place the news publisher within a larger context as well as to gather additional insight into industry best practices.

Sample

Relying on “criterion sampling” (Yin, 2014), the newsroom was selected from a group of national and regional news outlets with designated audience engagement departments that utilize editorial analytics to sustain readership as well as to increase digital revenue. With a busy editorial workflow, the newsroom offered insight into the behind-the-scenes dynamics of a major metropolitan news website with a weekly audience of about 2 million unique visitors. Thus, the operational size of the newsroom and its vast online audience reach both facilitated for the findings to be “genuinely based on critical investigation of all the data and not depend on a few well-chosen ‘examples’” (Silverman, 2013, p. 286).

To address the questions posed in this study, the researcher conducted a total of 11 semi-structured, in-depth interviews, in many instances lasting for up to an hour. The study participants were selected voluntarily, based on their professional occupations, requiring them to interact with audience data at various

stages of a news cycle – from primary gatekeeping and initial storyboarding to production and cross-platform distribution.

The sample included a wide array of the primary actors in the newsroom, such as the editor in chief, the Vice President for Digital Content and Strategy, the digital editor, the reader engagement editor, the breaking news editor, three department editor and three front-line reporters. Adequate sampling had insured the gathering of quality data, contributing to overall “thematic exhaustion and variability” (Guest, Bunce, & Johnson, 2006, p. 65). The sample includes five males and six females. Their work experience within the studied newsroom ranged anywhere from a few months to almost three decades.

Recruitment letters were sent out three weeks prior to the site visit. Interview recordings were transcribed, yielding 93 pages of single-spaced 12 pt. text. Transcripts are not included in the study to preserve interviewees’ professional privacy along with the newsroom’s organizational confidentiality.

Procedure

Prior to obtaining an approval from the Institutional Review Board (IRB), the researcher conducted a trial run of the following interview questions with an engagement editor at the *Columbia Missourian* to make sure that they are received and interpreted as intended:

Audience perceptions and journalistic norms:

1. Describe your audience. Who are they? What do they want?

2. What do you want the audience to get from your news content? Why?
3. How do you interact directly with users? What is the role, if any, of online comments? Social media accounts?

Measuring audience:

1. What tools do you use to measure how the audience interacts with your content?
2. What do your analytics tools tell you about your audience?
3. What measures are most important to your editorial decision-making?
4. What weight do you assign to analytics when making changes to content or content placement on a news platform?
5. What role do analytics play in evaluating your and your colleagues' job performance?
6. How well do your newsroom's analytics measure the quality and impact of your journalism? How might you measure quality and impact more effectively?

FINDINGS

Audience attention is a commodity that news organizations produce and sell to advertisers. Numbers, reflected in audience data, matter though newsrooms tend to focus too much attention on sheer volume and they worry about traditional journalistic values. While generating commercial revenue, the newsroom studied is conflicted between chasing numbers and fulfilling their professional journalistic creed rooted in the idea of informing and serving the public. The current audience measurement system provides only a footprint of readers' online behavior and yet falls short in assessing the quality and the impact of journalism being produced every day. Journalists, themselves, firmly rely on their editorial news judgment to decide what is newsworthy. At the same time, if there is an external pressure to boost traffic, the newsroom knows how to make it happen. Insight that audience data offers is used to address short-term and immediate editorial goals but not to drive news coverage.

Newsroom's perceptions of their audiences

The newsroom of nearly 100 employees serves a local online audience of more than six million unique visitors a month along with a million print subscribers. As the largest metro daily in the region, the core readership takes up 45 percent of the local audience, according to the company's scorecard. This exceeds the viewership of local television affiliates of several leading network stations. "We are the biggest local media – you can combine all three TV stations in town, and their audience is smaller than ours," said the Vice President for Digital Content and

Strategy, a veteran journalist who has spent almost 30 years working in the same newsroom.

It's "a town square" and "the place that defines what's going on in the city," he said, emphasizing the newspaper's online community outreach through the newspaper's live forums. Unlike prominent national news outlets that recently shut down their commentary sections, shunting public conversations to social media, this newsroom finds it critical to sustain and cultivate its role as a public forum. But this effort lacks a built-in monetization strategy for actively engaged users that statistically account for 10 percent of online surfers.

Growing an audience at a steady pace has been a challenge. "It's hard with a significant [audience] base," the editor-in-chief said. Several interviewees indicated that for a local daily there is a chance of hitting the ceiling, given that their core audience is constrained by the population size of their city and its metropolitan area. With limited resources and a downsized staff, the challenge is even greater. In the early 2000s, the newsroom employed 300 people. When the current editor-in-chief came in six years ago, he said, it was still a staff of 160. Today, it is a team of 100, which is still "empowering" to think that there are that many people covering the city. Still, "it's a double-edged sword," the Vice President for Digital Content and Strategy said.

Audiences and journalistic role conceptions

Despite transforming into a digital-first news operation, the legacy of being a print publication did not fade away. Partly, this is because the company's revenue

streams revolve around print subscribers that comprise more than a quarter of its overall readership, according to the editor in chief. “We can’t abandon that. It pays a lot of our bills. [...] It’s still the biggest part of our revenue,” he said. In the meantime, the digital revenue has been growing steadily, according to the Vice President for Digital Content and Strategy. But the growth was not fast enough to offset the print.

The economics of a news business spinning around print dollars and digital pennies continues to drive a wedge between two news platforms that editors and journalists alike treat as separate entities, each having its own audience. “They’re different audiences. We have some people who use both, but, largely, it’s a split still,” said the editor in chief. “There are some people who use both. There’s no doubt – we would like to grow that,” he added.

The tension between print and digital is self-evident. Without discrediting the importance of going digital, some readers continue to manifest themselves primarily through the print product. Thus, print subscribers are attributed to the core readership. Their relationship with the newsroom is seen as valuable, perpetuated through incoming letters and phone calls. “We talk to those people on the phone. [...] I feel like we kind of know our audience from that,” said the features editor.

Journalistically, the newsroom holds strong on reinforcing its public service role to audiences they serve. According to the editor in chief, their job in serving the community relates to the primary journalistic premise – serving and engaging audiences through quality by providing readers with information that they can

share along, saying: “‘Hey, mom. Hey, dad. Hey, sister, here's something you need to take a look at. It is important.’ I feel like that's the kind of connection we need.”

Thematically, daily news content revolves around life within the city. It breaks local news in business, politics, crime and sports. The newsroom overwhelmingly believes that their audience genuinely cares about their city, striving to make it a better place to live. Thus, the news organization takes great effort to produce high-utility database projects with a shelf life extending for up to a year. The audience reciprocates: whether it be a city school guide, a homicide map tracking crime by neighborhood or a guide to “100 Best Restaurants,” each generates more than a million page views.

Commodification of audience data

The common area of an open-space office layout, located on multiple floors, has several big screens showing real-time audience data from Google Analytics and Parse.ly. If the newsroom has been relying on Google Analytics since 2015, Parse.ly is their brand-new tool, first introduced on January 8, 2018, a week prior to the researcher’s site visit. Before Parse.ly, the newsroom used Chartbeat and, before that, Omniture.

If Parse.ly tracks users currently present on the website and throughout the day, the newsroom refers to Google Analytics for so-called “yesterday numbers.” But every time there is a switch to a new tool, the clock for how far back historical data can go is being reset to zero, which could be a downside if there is ever a need

to compare traffic's "highs" and "lows" for significant news stories spread across several years.

The digital editor also sends out daily traffic reports and announces hottest headlines during an 11 o'clock budget meeting. The overall vibe is that traffic numbers are becoming increasingly important to justify the success of the product – both in journalistic and financial terms. Not yet uniformly settled on a new reality, the newsroom grapples with internalizing its audience through the prism of sheer numbers. The breaking news editor who directly oversees a team of six journalists observes the tension first-hand. "Analytics scares some reporters," he said. "They see it like, 'Oh, I'm going to be judged on what people click on.'"

The hottest headlines usually associate with soft news. Thus, to newsmakers who aspire to produce quality content, it may appear that they need to downgrade their professional standards of newsworthiness in order to chase numbers. Speaking of the newsroom's ongoing race for higher traffic, one of the interviewees said:

We do so much more talking about the sheer numbers than we do what's in the story – what's behind this number and that number. This is tough. [...] It's difficult some days to just walk out of here. Journalistically, did I do my best or did I look at the scoreboard all day? Just moved the headlines around and to try generate numbers?

Competing interests and external pressures

The last quarter of 2017 the newsroom remembers as nothing more than a hasty chase for page views. The two, five-weeks periods, which interviewees casually label with the television ratings term “sweeps,” turned into an onerous team effort. It is an internally set requirement intended to generate additional page view revenue.

The first sweeps period came around mid-September to late October. The projected increase in year-over-year estimates was set to 15 percent. The task primarily fell on the shoulders of the digital editor who is also the main lead behind the homepage:

We're just and literally busting our butts to come up with additional news stories or photo galleries, repackaging stuff – whatever we can think of to figure it out. Ok, stuff that's popular and it has some value. It's not just the junk food of news. So, we cranked it out. [...] I think we've got an extra 15 million page views during the first sweeps period, which was a lot. When you're averaging about two million [page views] a day, and all of a sudden, you go up by 15 million, that's like more than two weeks of traffic.

The reader engagement editor was in the exact same boat, trying to think of ways to package news stories and further promote them via 17 Twitter and 13 Facebook accounts. “It's exhausting to think about content that way,” she said. “I don't like calling it content, but that's what it is. How can we turn this article into a

slideshow? What can we resurface? How can I best link to it? It's a thought process that we don't do all the time the way that we do during sweeps."

Chasing traffic during sweeps could be extremely hard. First off, traffic is seasonal. News is its key driver and catalyst. But for a local newspaper to be in the epicenter of big news is "a stroke of luck": "You can't be guaranteed massive flooding on the [river]. It's like Jesus. [...] You can't get blood from a stone. I mean, you can't manufacture the news," the digital editor said.

No matter how "aggressive" or "crazy" it may sound, having an extremely rare natural phenomenon as a total solar eclipse and a verdict in a police shooting case close to one another, in fact, made it easier to pull a 15 percent traffic bump.

But the second sweeps period was a far more difficult challenge, coming right before the winter holidays – from early November to late December. The goal for growth in page views was set to 25 percent, running up against the end of 2016 figures. "To say no – it's kind of setting yourself up for failure in a way," the digital editor said. They fell short by a narrow margin, hitting a 22 percent increase. "That came on top of another sweeps period. I'm fine with this. We are fantastic. But it's exhausting."

Editors found sweeps being extra stressful. Some saw them as "an arbitrary goal." Missing the second time around, the reader engagement editor said, felt "really shitty":

But it's worrying that we're constantly on this cycle when we have to increase page views. Like at a certain point, do we reach max audience? Is there a point where we can't reach any more users, page

views or stock or we have to continually generate more and more and more? [...] If we have a banner year this year, how are we going to be held to that next year? I think it's just continually going to be... where does it stop? But that's a really long-term question, and at the moment we just want the ad revenue.

Setting both sweeps periods for the last quarter of a calendar year does not seem to be an arbitrary decision. From a commercial standpoint, it is a last-minute chance to make target audiences a more marketable commodity for advertisers (McManus, 2009). On the other hand, being on the receiving end of this challenge, the newsroom had to figure out ways to satisfy the needs of two separate markets – advertisers and the readership – while still producing the same commodity. But the newsroom had a hard time relating to this goal, given their primary journalistic commitment to inform and serve the public.

Seeing audience through analytics

Due to a set organizational goal to increase page views, and analytics reports used as indicators of that growth, the newsroom was pressed to evaluate the audience's reception of media messages exclusively through quantitative measures. But most interviewees said page views should be the least important way to gauge interaction. While knowing that page views have a direct influence on advertising revenue, newsroom management understands the flaws of this measure in assessing the quality of journalism and its reach: "There is some monetization there, but it's also not always the truest test of engagement. There are ways you can move

the page-view dial that doesn't really reflect true engagement for your site," said the Vice President for Digital Content and Strategy.

When asked to define audience engagement, both reporters and editors mentioned several metrics that help them interpret reader behavior. Referring to metrics on an analytics dashboard, they primarily pointed to users and average time spent on the page. Therefore, when evaluating audiences in journalistic terms, their core logic boils down to the following: how many people read a story and for how long?

The online sports editor, whose department is the leading force in generating traffic, mainly goes by users. "It's sheer numbers," he said. "I feel like the more people you would get in there, if it's worthy content, the better chance it gets spread out to all of their friends." Generally, if a story attracts anywhere more than 25,000 users, it's considered to be "pretty high," but "a low" could be as few as 500.

As the newsroom begins to invest its efforts and resources into data-informed audience engagement, they envision looking more closely into "high-" and "low-quality clicks" with a greater emphasis on "sessions from the heaviest users." This seems a reasonable next step as the newsroom sets the ground for introducing a paywall within the next year. In the meantime, the day-to-day priority is to focus on page view revenue, with a target growth set to 15 percent, compared to the previous year.

But high traffic is not a true measure of journalistic quality, effort and time spent in production. Audience data is only "one aspect" of readers' feedback, several

journalists and editors across the newsroom stated. A breaking news reporter who on a busy day produces up to eight stories, said:

It's not end-all and be-all of what we do. It's one measure of who is reading your work, how long they're spending on your work, how much influence is having what you are writing. But the most high-read stories that I have produced are not the stories I am proud of.

A 200-word blurb about a woman flashing at a baseball game took the breaking news reporter 15 minutes to write, but it turned out to be one of the most-read stories of the year. It only caught traffic because thousands of people at the game, then, millions of television viewers were Googling it. "Is it the story that I am most proud of? No," she said.

What journalistically is an important, high news-value story may not even strike a chord with the audience. Journalists' "mental representations" of what is newsworthy (Van Dajk, 2009, p. 192) may not even align with people's interests, at times, receiving as little as 500 views. An education beat reporter was frustrated to see her recent story on changes in state funding of public colleges not getting any traction online. "Are you kidding me?! It's so important, right? [...] Some story that someone else wrote about a dog that did something... will get 20,000, 50,000 – whatever. You know, that can be really disheartening," she said. But a different newspaper found the story to be important enough to publish it as a center piece in its Sunday edition.

Audience data is elusive, not always pointing in the direction of journalistic quality. But there are certainly times when things do match up – a highly important

breaking news story generates tremendous traffic in excess over a million of page views. Aside from that story about a woman flashing at a baseball game, the No. 1 most read story in 2017 was an investigation about police officers using tear gas during street protests. It is one of the instances when wide audiences demonstrated high interest in the story, recognizing its extremely high news value.

But audience data cannot predict what will go viral. When headlines start shooting up, editors say, it could be that Drudge or Reddit picked it up and it got further shared on social media. That is how a three-year-old story about a guy dying from a snakebite all of a sudden made it to the top headlines. “We couldn't figure it out,” the breaking news editor said. “What does that data tell me? I don't know. I can't make other three-year-old stories go viral. [...] There are things like that happen, that you just can't control.”

He takes traffic numbers with a grain of salt. “We don't want to do clickbait here,” he said. “But that doesn't mean you can't learn something from the data to be used to take your serious journalistic story that is important and get clicks for it. There is no shame for that.”

Editorial decision-making and audience data

Audience data is not perfect, and it can be hard to interpret. It is an insight the newsroom looks at in “the rearview mirror,” supplementing it with journalistic news judgment and expertise. It informs short-term editorial decisions in content packaging and its placement across the website, but not as much as in the realm of

directly influencing daily news coverage. If any significant impact persists, it is likely to occur during the primary gatekeeping process:

Sometimes, you've got to say, "Is it worth covering every one of these potent crimes? Are these smaller communities having a vote on this thing? Is it really interesting enough to a wide enough audience? Is it worth our time doing?" the breaking news editor said.

In their daily interactions with analytics, the editorial staff has adopted three distinct, data-driven decision-making strategies, primarily directed at resolving short-term traffic goals. Each of those changes is rather cosmetic, often resulting in a traffic bump, unless "it's a pretty thin story to begin with" without any compelling elements to draw audience's attention.

There is no longer such a thing as one audience, or one front page. It is "a lot of different streams" of people accessing a mix of stories at different times throughout the day through different devices and various platforms. In curating web traffic, the analytics team tries to spread content among a wide range of people – those "high-power users" who have the homepage bookmarked along with casual so-called "social users" who find their stories by browsing on Google, Facebook and Twitter.

Typically, the goal is to "move people in a circle," allowing them to come into that circle at various points to see "the new stuff," said the reader engagement editor. But as the wheel keeps rotating, in a few hours, things can get old, the editor in chief said. Then, he pondered: "How can we seize that moment? How do we get them to stick, so that they would want to spend time?"

Following and meeting the audience where they are, capitalizing on the power of now, can pay dividends. After seeing a conversation on Twitter, the digital editor cobbled together in haste a news brief about Matt Lauer being fired from the Today Show. Being ahead of CNN and AP brought them 100,000 page views that morning, which is nearly five percent of a typical day's traffic.

Typically, we don't get a national story like that. It is not going to be our bread and butter – people come to us for local news. But we'll take the traffic where we get it. We don't turn it away – oh, no, don't read that. Come on, read everything here – we like that.

Editorial perceptions of content and audience are no longer monolithic, tied to a single platform or a slot on the homepage. The breaking news editor mirrors their daily editorial workflow beyond grids and platforms:

Ok, well, I can put it up now. Now is a good time, we'll get a little splash of traffic. But, then, you know what?! Maybe, this is something that goes back up in the hot spot at 7 p.m. when people are getting out their phones or iPads or are watching TV or chilling at home. Or, maybe, it's a story we dropped on Tuesday, but on Saturday some people might sit down to read it.

If some news organizations have seen their homepage die slowly due to readers increasingly finding news through social media and Google, this has not been the experience for this newsroom. The homepage remains “a reasonably healthy part of traffic,” the Vice President for Digital Content and Strategy said. Nearly one-third of their digital audience (30 percent) comes in through the front door. These are “local,” “traditional” or also “power users” who check in consistently

throughout the day, and they are “less likely to be one-and-done-type readers” randomly finding their stories on Facebook or Twitter, the breaking news editor said. However, according to the reader engagement editor, the homepage traffic has been dropping over time, which poses “a huge concern” unless social and other referral traffic increase at equal margins to offset the losses. The future outcome is uncertain, given that Facebook is “such a black box,” she said.

Until this March when Facebook eventually announced it would not separate news content into a brand-new, so-called “Explore feed,” the reader engagement editor was worried that their referral social media traffic would go down. But knowing that the experiment fell flat, she can breathe a sigh of relief. Facebook is the No. 1 source of their social media referral traffic, generating five times more page views than Twitter. They post to Facebook almost every half-an-hour, but she never knows who will see their content and when:

Sometimes, we'll get a lot of comments, and a lot of people liking, but, then, when you look at how many people have actually clicked through to read the story, it's not as high as the combined number of comments and likes. People will interact with content that they don't even bother reading.

Despite being dependent on social media for its referral traffic, for the online audience, the newsroom sees its homepage as the primary gateway to their news stories. Interviewees insist editorial news judgment is the absolute driver of any macro-level decisions, aimed to preserve the homepage and its journalistic integrity.

During her 10-hour work-day, the digital editor updates the homepage almost every 45 minutes or an hour. The featured story gets changed three-to-five times, depending on how busy the day is. If things are slow, stories may sit at the same spots a little longer. The digital editor admitted that to a certain extent audience data dictates content placement, but she said, she would not necessarily “go in and mess with something” just because it generates a lot of likes and shares. It is certainly “not the sole factor,” she added. Similarly, if something is shooting off and going viral on social, the breaking news editor says he does not always press to upgrade a story to a more prominent spot. It could be that the story is “doing really well on its own” and people are finding it on social media or elsewhere.

Although audience data provides a representative sample of what people click on, newsroom decision-makers say they exercise journalistic integrity, anchoring their news judgment through slots and time stamps. Depending on how newsworthy a story is, it receives a certain play on the website, resulting from a wide spectrum of subtle and nuanced editorial decisions. “Briefly, I may choose to have ‘I stabbed my wife [over changing a TV channel]’ up higher,” the digital editor said, “but, in the long run, I’m going to give way more attention and care to the [governor’s] story. [...] There’re different ways you can try and package that or promote it that will eventually strike a chord with readers.”

There are certainly times when the homepage leans toward “soft” news, featuring cute baby cheetahs, for example. Editors, themselves, know that it is “a little off beat.” It is a slippery slope, enticing clickbait behavior and feeding into the audience’s browsing mood. What people click on isn’t always what journalist

themselves consider important news,” the breaking news editor said. Indeed, if not, why did a shoplifter with a frying pan down her pants attract more than 173,000 visitors? On that same day, but a year later, an important A1 story about parents now being able to use their kids' 529 plans for private school expenses generated only 28,000 users.

In their constant pursuit of maximizing page views, the newsroom persists in trying to figure out what drives traffic. Stories about crime, a favorite baseball team or exquisite dining tend to generate consistently high numbers. But headline tweaking in real time can boost traffic from 30 to 300 percent, editors say. In nine out of 10 instances, traffic goes up, the online sports editor said. Because the website updates slowly, they usually have to wait for about 15 minutes before a headline starts taking off.

To the editor in chief, the headline on its own can be a mark of quality journalism: an art form that translates into higher audience engagement beyond a single page view. He expects his staff to write headlines “with clarity and spark,” but in reality, the digital editor said, they often come last when there is neither time nor energy left. Greater time and effort go into crafting “the first two paragraphs of the story,” but those do not translate into traffic, she said.

The headline is “the curb appeal,” a sell-line to grabbing audience’s attention. As one of the editors put it: “It's like you're selling your house, and you don't bother putting down a fresh welcome mat, or paint over the scuffs on your door.” If a story is not trending, it could be that the headline is flat or unclear. “There is no science to a good headline. You know it when you see it,” the digital editor said. The trick of

tweaking is about making it “more specific,” “a little more active” and “social-friendly” – as something people would click on while browsing Facebook.

Sometimes, it is as simple as mentioning the name of a favorite baseball team.

For the digital editor, it is about finding that “sweet spot for all different audiences” or “a common ground” that readers can relate to, which is “not always easy.” For hard news, “if you hit it right off the bat with the right tone,” or in soft news, relating to personal experiences can drive the message home:

Yesterday, someone got stabbed at a restaurant, and the headline said,

‘A man stabbed at a restaurant.’ A man stabbed at Steak ‘n Shake.

That’s better. You're going to read more about a guy who got stabbed at Steak ‘n Shake than a restaurant! Because everyone is like, “Steak ‘n Shake! I’ve been to Steak ‘n Shake!”

The newsroom has been consistently using content packaging as a supplementary strategy for increasing traffic to their stories, both old and new. Even during sweeps, editors said, they go after stories while always thinking: “Is there a way to package this or present it that will automatically help boost traffic?” The trick could be as simple as building content streams of related stories to more sophisticated ways such as exploring formats that go beyond a straightforward inverted pyramid text piece. But the latter, by default, often leads to building slideshows, given the persistent pressure from advertisers to generate page views.

Audience changes every day, the digital editor said. This insight draws her to mix and match content, sometimes even pulling it from the shelf and bringing it back for the readers who missed it previously. “Our audience and our readers have

no clue. They don't remember," she said. Right around Christmas, she invited her audience to check out a report card for the city's favorite baseball team to track players' progress during the past season. By repurposing "old" content, she generated an additional 100,000 page views.

The reader engagement editor knows that for their mobile users, their slideshows seem clunky – swiping left and right with a finger is inconvenient. "People just don't like it," she said. "People arrive at our site, and they see that it's a slideshow, and they're like, 'Here's another page clickbait.'" But when it comes to analytics, the arithmetic is simple and straightforward: 20 photos generate 20 page views whereas one text piece counts as one page view. Obviously, not every story can become a gallery due to a lacking visual appeal, but the temptation to produce slideshows over and over again may erode the quality of journalism.

Editors are learning to unbox the metrics, with a better sense of what variables trigger a desired outcome. Even "a corporately done thing" like a Pearl Harbor gallery, consisting of 70 photos, generated over a million page views after being re-shared by a state senator and numerous veteran groups. But a high number of page views is not a reflection of true engagement. Average time spent on the page could be as low as 30 to 40 seconds, the engagement editor said. Similarly, a story with over a million page views could in fact attract just about 100,000 users. Although this is still tremendous traffic, a large percentage of those clicks could be "pretty low-quality." What if many of them are just driving by? How many of them are actually local audience?

Toward better measures of engagement

Audience data has yet to become a uniformly powerful actant. Editors, who interact with data more closely on a daily basis, share a deeper understanding and connection with audience metrics. At the same time, they are not fully sold on the value proposition of Big Data technology because it does not tell the full story about who the users are. The digital editor said, she would benefit from having “one-stop shopping” to look and say: “Oh, wow...I thought this could do really well with the millennials or with baby boomers. Instead, it's doing its best with people who are under 20 or whatever it might be.”

The analytics software out on the market does not have the means to connect user demographics information with their browsing behavior unless it is some in-house, custom-made programming tool that The Washington Post has been relying on. According to the company CIO and CTO Shailesh Prakash, their newsroom uses five Big-Data tools – Clavis, Varality, Bandito, Headliner, Heliograf – in order to decode audience behavior online. This technology has a wide range of capabilities from helping to understand who the readers are, predicting popularity for a story, suggesting “the best performing combination” of headlines and images while taking in real-time audience feedback, automatically generating headlines based on story content to automatically writing stories using artificial intelligence (AI) based on structured data (“How Washington Post’s data-driven product development engages audiences”, 2017).

In terms of relying on a different metric for analyzing and interpreting audience, the researched newsroom, overall, did not provide concrete answers to

whether one measure would be better than the other. Instead, several interviewees expressed a genuine interest in learning case studies of other newsrooms, trying to make sense of messy audience data.

Following the example of Slate, the newsroom could, for example, redefine content success around more meaningful measures, like “engaged time” and “loyalty” as opposed to seeing their audiences through old and rudimentary metrics, such as “page views” and “visitors” (Cherubini & Nielsen, 2016). But there is no single metric that could assess the quality of online behavior. Even “engaged time,” which seems to offer a more sound measure of audience engagement, could be an inflated number. Parse.ly, the analytics software the newsroom has been relying on for a few months now, addresses this issue by calculating the engaged time in increments of “heartbeats.” The clock is ticking only when a user is clicking, scrolling, swiping, tapping on the page. A built-in timer sends heartbeat-like signals back to the server every 15 seconds. Based on the duration of that engagement, users are further broken down into four distinct categories – “mis-clicks” (less than 15 seconds), “passer-bys” (15 to 30 seconds), “short stays” (30 to 60 seconds) and “long stays.” The latter’s presence on the page exceeds 60 seconds in duration. An analysis of 1.6 billion sessions centered around 18 million publisher posts showed that an average time for “a good visit” amounts to 81 seconds (Parse.ly, 2017).

The newsroom would benefit from moving away from page views because those show nothing but volume without any assessment of how much interest the audience demonstrates. At its worst, a single page view could be just a mere glance

on the page for less than 15 seconds, not followed by a single action, even a single click (Parse.ly, 2017).

Given the news organization's future plans to establish paywalls, the newsroom management sees audience data as a valuable insight into increasing customer loyalty with a further conversion of audience members into subscribers. The editor in chief sees the conversion occurring if they continue to deliver a quality product. "We can't lose track of the fact we're doing journalism," he said. "That can't be trumped by the different gimmicks. We are not going to gimmick our way to success. We have to have good content."

The newsroom insists that it is maintaining its editorial integrity. Analytics as a dynamic technological actant firmly embedded in the editorial workflow rather serves, they say, as a measuring stick for double-checking professional gut instincts, often set to prove journalistic value and relevance. It could also be a reflection of a persistent presence of a print-centric mentality. "Everyone is still like, "Ah, I got a story on A1!" the digital editor said. "Great! But, you know, how much traffic did it get online? Oh, really? That's all? Maybe, it wasn't the greatest story you think. But it's hard to change that mindset."

As one reporter put it, "I know that print is increasingly dying, as people would say. "But you don't know how many people picked up the paper and read your story."

Some reporters and editors emphasized that they still value personal feedback in the form of hand-written letters, emails or comments rather than fully trusting machine-generated traffic reports. "Editors like to trust their gut instincts in

the way they've done things," the Vice President for Digital Content and Strategy said. "A lot of times you hear about – I got two calls about this thing. Well, they will let two phone calls dictate their coverage more than a Google sample of five million people."

DISCUSSION

The power dynamics between newsmakers and their audiences in the contested “marketplace of attention” remains an unsettled question, pointing at a greater co-dependence. The relationship is no longer linear, going top-down from the traditional “information gatekeepers” (Shoemaker & Vos, 2009) and “agenda-setters” (McCombs, Shaw, & Waeber, 2014) straight to news consumers. The exchange is rather sophisticated and nuanced, no longer fixed on a single channel of distribution. As observed in the newsroom studied, while the newsroom holds strong to its role of public service, it also persistently engages in acts of social listening, often adopting “audience-centric” values (Anderson, 2011a). Pressed by corporate interests to generate advertising revenue, the newsroom ends up chasing page views that are meaningless for journalistic quality.

Journalists’ drifting role conceptions

In theorizing a network of relations and associations performed by actors and actants, the ANT employs the concept of movement, which is essential for capturing the nature of the observed day-to-day interactions as newsmakers connect with their audiences. The impulse is to break through the clutter and to relate without being too salacious while holding strong to their beliefs of professional practice, shaped by “symbolic constructions [of] newsworthiness” (Domingo, Masip, & Costera Meijer, 2015). While entering a contested digital environment, the newsroom studied faces an ongoing challenge of being steered away from its traditional role of public service in order to satisfy audience’s

personal interest for soft news (Lee et al, 2012). This internal dilemma reflects “a paternalistic function of journalism” that is prevalent among legacy news organizations as opposed to their digital-native counterparts (Holton et al, 2016).

Interviewed reporters and editors come with a strong news background, which translates into decades of working in print publications. Representing professional “communities of practice” (Weiss & Domingo, 2010), or what Zelizer (1993) describes as “interpretive communities,” they strive to preserve their journalistic integrity in deciding what warrants public attention.

Being heavily dependent on the audience’s reciprocity, which brings in advertising revenue, the newsroom falls into the mode of following their readers’ browsing behavior. That action can result in clickbait content (Hurst, 2016). “There are times when the readers might want to click on the NFL cheerleaders gallery all day long, but that's not what we're here for,” said the breaking news editor, calling for journalism’s core mission to inform, rather than to amuse and entertain. This unsettled relationship points to diverging interests between consumers’ preferences and journalistic news judgment, widely conceptualized as “a news gap” (Boczkowski & Mitchelstein, 2013).

Audience as a commodity

News organizations’ enduring struggle to juggle competing interests and expectations between serving the public and generating advertising revenue is not new. Audience measurement systems allow to track online behavior on site in real time, which adds a considerable pressure and provides an incentive to monitor

engagement. Within the marketplace of attention (Webster, 2014), the latter directly correlates with the value of “audience power” that the newsroom produces by serving news to their target demographics (Smythe, 1981). Audience data is a new form of industry ratings, which dictates an asking price for advertising rates.

The newsroom studied experiences the described tension firsthand, imposed by the demand from the corporate office to generate page views during “sweeps.” When the overall mentality is based on the assumption that the newsroom is intended to produce “messages,” “information,” “images,” “meaning,” but not audience power per se in economic terms, this perception creates and triggers a mismatch within the system of journalistic beliefs and professional expectations. According to Smythe (1981), all these concepts are “subjective mental entities” that imply “effects, or purpose,” but do not constitute a product, or a commodity (p. 231).

The commodification of audience data increasingly prompts the newsroom to perceive that doing engaged journalism is the same as running the business of creating audiences around valuable local breaking news content. But that business-like mindset is particularly common among editors on the audience engagement side of news operations that spend time scrutinizing traffic data. Frontline reporters, though, who bring the tapestry of news content to the table tend to personify their audiences from a human-interest standpoint, which is reflected in their commitment to public service.

Conflicting expectations for audience data

Audience as an entity transformed into a powerful actor becomes a higher-value commodity for advertisers. Consumer attention is reflected in “informational traces” (Plesner 2009) that are being tracked in real time by analytics software. Audience exposure to media messages is being amplified and filtered through Facebook and Twitter, each serving as an independent “mediator” of public discourse (Van Dijck, 2012).

As observed in the newsroom studied, the audience’s response does not always satisfy journalists’ professional expectations, leading to skepticism about Big Data’s ability to reasonably assess the quality of journalism. As a result, its designated role in the editorial workflow is being reduced to a technological actant set to prove and reaffirm the validity of journalistic gut instincts.

But advertisers view traffic data as a footprint of audience attention, primarily quantified in page views, which the newsroom, in fact, perceives as an untrue measure of audience engagement. It conflicts with their journalistic duty to cultivate a personal relationship with the audience. That exchange is built on a premise that the readership continuously spends time on site interacting with news and engaging in a public conversation. Page views capture only the volume of that interaction (Cherubini & Nielsen, 2016), which, within the context of journalism norms, could be interpreted as catching noise when the expectation is to aim for a signal.

Redefining audience engagement

During the first five-week sweeps period, the newsroom generated about 15 million page views, almost doubling its monthly supply. But higher traffic does not necessarily translate into greater audience engagement. Knowing exactly which variables drive traffic, the newsroom set itself on a path of engineering clicks. Journalistically, the news production got pushed into a revolving cycle of pursuing slideshows over the long-form text pieces, which count as only one page view. Several interviewees provided examples of soft news generating much higher traffic, highlighting an ongoing tension between “hard” and “soft” news. However, in order to provide a qualitative assessment of how much of its coverage the newsroom shifted to soft news during sweeps, warrants a separate study that would employ textual analysis as its primary methodological approach.

The newsroom’s reliance on audience data for guiding short-term editorial decisions in content packaging and its placement on the website mirrors a common trend across many news organizations nationwide. A survey of 178 editors working for U.S. dailies found that out of 62 percent of editors who daily track analytics, only seven percent say they apply it for news selection and coverage (Lowrey and Woo, 2010).

The newsroom has yet to realize the full potential of editorial analytics. The audience engagement team represents a distinct group of powerful actors, who hold knowledge and understanding of analytics. Further democratization of data among traditional newsroom actors, including front-line reporters, is essential for deeper understanding of audience metrics. Sharing traffic reports could serve to redefine

and reach new editorial goals that would not put journalistic quality at odds with advertising revenue.

Before the newsroom introduces a paywall, it could be beneficial to shift the key emphasis from users and page views toward engaged time. This will help to reset a standard for measuring quality engagement over sheer volume. Time is “a new universal currency” that is increasingly becoming the most important metric in the digital media ecosystem. This mirrors the ongoing media industry trend, aimed at capturing consumer attention (time) through higher quality programming as rigorously being pursued by Netflix and Amazon (Kaplan, 2017).

CONCLUSION

The interviews collected and analyzed here shed light on how journalists at a legacy news operation try to hold strong to their professional role of public service while being pushed into the realm of “soft” news. Proponents of analytics say data produces insights into increasing revenue streams by supplying real-time editorial guidance for boosting the audience’s exposure to media messages. In reality, the value proposition is not as promising. Audience measurement technology is imperfect and, sometimes, poorly understood, providing the rationale for journalists to fall back on their news judgment and expertise, hoping to rely on their traditional roles as “gatekeepers” (Shoemaker & Vos, 2009) and “agenda-setters” (McCombs, Shaw, & Waeber, 2014).

Commodification of audience data has lead the news organization in this study to seek out ways to serve primary audiences and advertisers alike while producing the same news product. Speaking from their personal experiences, the interviewees found it challenging to generate bulk traffic while still trying to go after important news.

Limitations: This study has limitations. The qualitative research framework is not expected to provide generalizable results. However, this does not trump the substance of the findings. Each data unit, or an interview, gathered throughout the research, represents a personal reflection and assessment of how journalists negotiate their professional practices using a footprint of real-time audience behavior. A supplementary analysis of industry reports along with online traffic

reports from the participating newsroom provided additional insights into audience metrics and their measurement.

Directions for future research: Scholars studying Big Data analytics advocate for a mixed-methods framework, which combines in-depth interviews (Sumpter, 2000; MacGregor, 2007; Anderson, 2011; Usher, 2013; Tandoc, 2014) and newsroom ethnographies (Singer, 2013; Vu, 2014; Belair-Gagnon, 2015) with a survey of several hundred respondents in order to arrive at richer results and mitigate limitations of each framework (Tufekci, 2014, p. 10).

Communications research could further benefit from a textual analysis of digital news content that would trace a relationship between external pressures to generate traffic and news organizations' growing tendency to cover "soft" news. This could further a discussion about the impact of "soft" news in perpetuating and amplifying cultural stereotypes.

Refining analytics is not sufficient to assess the quality of audience engagement. A footprint of audience interactions captured by the analytics software does not necessarily indicate the reader can recall key points in a story or apply what has been learned to effective citizen participation. Numbers not supplemented by critical thinking and expertise of human actors could be meaningless, leading to erroneous outcomes.

BIBLIOGRAPHY

- 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 (19), 529-547.
- Anderson, C. W. (2011b). Between creative and quantified audiences: Web metrics and changing patterns of newswork in local US newsrooms. *Journalism*, 12(5), 550-566.
- Barot, T. (2016). The botification of news. *NiemanLab*. Retrieved from <http://www.niemanlab.org/2015/12/the-botification-of-news/>
- Batsell, J. (2015). *Engaged journalism: connecting with digitally empowered news audiences*. New York: Columbia University Press.
- Belair-Gagnon, V. (2015). *Social media at BBC news: The re-making of crisis reporting*. New York: Routledge, 2015.
- Biyani, P., Tsioutsoulouklis, K., & Blackmer, J. (2016, February). " 8 Amazing Secrets for Getting More Clicks": Detecting Clickbaits in News Streams Using Article Informality. In *AAAI* (pp. 94-100).
- Chowdhry, A. (2016, February 29). Facebook emoji 'reactions': Are there ulterior motives? *Forbes*. Retrieved from <https://www.forbes.com/sites/amitchowdhry/2016/02/29/facebook-reactions/#6a1c8f4f1a62>
- Gibbs, S. (2016, December 22). Mark Zuckerberg appears to finally admit Facebook is a media company. *The Guardian*. Retrieved from <https://www.theguardian.com/technology/2016/dec/22/mark-zuckerberg-appears-to-finally-admit-facebook-is-a-media-company>

- Berger, A. A. (2000). *Media and communication research methods: An introduction to qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications, 2000.
- Bernstein, M. S., Bakshy, E., Burke, M., & Karrer, B. (2013, April 27 – May 2). *Quantifying the invisible audience in social networks*. Paper presented at CHI 2013: Changing perspectives, Paris, France.
- Blom, J. N., & Hansen, K. R. (2015). Click bait: Forward-reference as lure in online news headlines. *Journal of Pragmatics*, 76, 87-100.
- Boczkowski, P. J., & Mitchelstein, E. (2013). *The news gap. when the information preferences of the media and the public diverge. [electronic resource]*: Cambridge, Massachusetts : The MIT Press.
- Constine, J. (2015, October 15). Twitter and Facebook are turning publishers into ghost writers. *Techcrunch*. Retrieved from <https://techcrunch.com/2015/10/15/smart-pipes-and-dumb-content/>
- Cherubini, F. & Nielsen, R.K. (2016). *Editorial Analytics: How news media are developing and using audience data and metrics* (Reuters Institute Digital News Project). Retrieved from <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Editorial%20analytics%20-%20how%20news%20media%20are%20developing%20and%20using%20audience%20data%20and%20metrics.pdf>
- Cohen, J. E. (2012). *Configuring the networked self: Law, code, and the play of everyday practice*. Yale University Press.

- Costera Meijer, I., & Groot Kormelink, T. (2015). Checking, Sharing, Clicking and Linking: Changing patterns of news use between 2004 and 2014. *Digital Journalism*, 3(5), 664-679, doi:10.1080/21670811.2014.937149
- Couldry, N. (2008). Actor network theory and media: Do they connect and on what terms? in Hepp, A., Krotz, F., Moores, S. and Winter, C. (Eds.), *Connectivity, networks and flows: Conceptualizing contemporary communications*. Cresskill, NJ, USA: Hampton Press, Inc., 2008, pp. 93-110.
- Couldry, N. (2011). The Audience... and How to Research It. *The handbook of media audiences*, 4, 213-227.
- Deluliis, D. (2015). Gatekeeping theory from social fields to social networks. *Communication Research Trends*, 34(1), 4-23.
- Deuze, M. (2005). What is journalism? Professional identity and ideology of journalists reconsidered. *Journalism*, 6(4), 442-464.
- Domingo, D., Masip, P., & Costera Meijer, I. (2015). Tracing digital news networks: Towards an integrated framework of the dynamics of news production, circulation and use. *Digital Journalism*, 3(1), 53-67. doi:10.1080/21670811.2014.927996
- Fairfield, J., Shtein, H. (2014). Big data, big problems: Emerging issues in the ethics of data science and journalism. *Journal of Mass Media Ethics*, pp. 38-51.
- Green, J., & Jenkins, H. (2011). Spreadable media: How audiences create value and meaning in a networked economy. *The handbook of media audiences*, 109-127.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59-82.

- Ha, L., Xu, Y., Yang, C., Wang, F., Yang, L., Abuljadail, M., ... & Gabay, I. (2016). Decline in news content engagement or news medium engagement? A longitudinal analysis of news engagement since the rise of social and mobile media 2009–2012. *Journalism*, 1464884916667654.
- Harcup, T., & O'Neill, D. (2016). What is news? News values revisited (again). *Journalism studies*, 1-19.
- Harris, D. (2015, March 4). Data might be the new oil, but a lot of us just need gasoline. *Gigaom*. Retrieved from <https://gigaom.com/2015/03/04/data-might-be-the-new-oil-but-a-lot-of-us-just-need-gasoline/>
- Hassard, J., & Law, J. (1999). *Actor network theory and after*. Oxford; Malden, MA: Blackwell, 1999.
- Heath, A. (2017, July 27). Twitter didn't add any new users last quarter. *Business Insider*. Retrieved from <http://www.businessinsider.com/twitter-didnt-add-any-new-users-last-quarter-and-the-stock-is-diving-2017-7>
- Hermida, A., Fletcher, F., Korell, D., & Logan, D. (2012). SHARE, LIKE, RECOMMEND: Decoding the social media news consumer. *Journalism Studies*, 13(5-6), 815-824. doi:10.1080/1461670X.2012.664430
- Holton, A. E., Lewis, S. C., & Coddington, M. (2016). Interacting with audiences: Journalistic role conceptions, reciprocity, and perceptions about participation. *Journalism studies*, 17(7), 849-859.
- How Washington Post's data-driven product development engages audiences. (2017). *Reuters*. Retrieved from <https://agency.reuters.com/en/insights/articles/articles->

archive/how-washington-post-data-driven-product-development-engages-audiences.html

Hurst, N. (2016). *To clickbait or not to clickbait? An examination of clickbait headline effects on source credibility* (Doctoral dissertation, University of Missouri-Columbia)

Kaplan, Z. (2017, Jan. 9). In a digital media world, all that matters is your time. *Recode*. Retrieved from <https://www.recode.net/2017/1/9/14214322/digital-media-consumer-facebook-video-monetization-content-advertising>

Karlsson, M., & Clerwall, C. (2013). Negotiating Professional News Judgment and “Clicks”. *Nordicom Review*, 34(2), 65-76.

Kolowich, L. (2017, July 25). The best time to post on Instagram, Facebook, Twitter, LinkedIn, Pinterest, and Google+. *HubSpot*. Retrieved from <https://blog.hubspot.com/marketing/best-times-post-pin-tweet-social-media-infographic>

Kormelink, T. G., & Meijer, I. C. (2017). What clicks actually mean: Exploring digital news user practices. *Journalism*, 1464884916688290.

Kosterich, A., & Napoli, P. (2016). Reconfiguring the Audience Commodity: The Institutionalization of Social TV Analytics as Market Information Regime. *Television And New Media*, 17(3), 254-271, doi:10.1177/1527476415597480.

Kovach, B., & Rosenstiel, T. (2007). *The elements of journalism* (Rev. ed.), New York, Three Rivers Press.

Kuzel, A. (1992). Sampling in qualitative inquiry. In B. Crabtree & W. Miller (Ed.), *Doing qualitative research* (pp. 31-44). Newbury Park, CA: Sage.

- Lee, A., Lewis, S., & Powers, M. (2012). Audience clicks and news placement: A study of time-lagged influence in online journalism. *Communication Research*, 41(4), 505-530.
- Leech, B. L. (2002). Asking questions: Techniques for semistructured interviews. *PS: Political Science and Politics*, (4). 665.
- Lewis, S. C., Westlund, O. (2015). Big data and journalism. Epistemology, expertise, economics and ethics, *Digital Journalism*, Volume 3, Number 3, pp. 447-466.
- Loosen, W., & Schmidt, J. H. (2012). (Re-)discovering the audience: The relationship between journalism and audience in networked digital media. *Information, Communication & Society*, 15(6), 867-887.
- Lowrey, W. and Woo, C. W. (2010). The news organization in uncertain times: Business or institution? *Journalism and Mass Communication Quarterly*, 87: 1, pp. 41–61.
- MacGregor, P. (2007). Tracking the online audience. *Journalism Studies*, 8(2), 280-298.
doi:10.1080/14616700601148879
- McCombs, M., Shaw, D., & Weaver, D. (2014). New directions in agenda-setting theory and research. *Mass Communication And Society*, 17(6), 781-802.
- McChesney, R. W. (2013). *Digital disconnect: How capitalism is turning the Internet against democracy*. New York: The New Press, 2013.
- McKenzie, C., Lowrey, W., Hays, H., Chung, J., & Woo, C. (2011). Listening to news audiences: The impact of community structure and economic factors. *Mass Communication And Society*, 14(3), 375-395.
- McManus, J. (2009), The commercialization of news, in K. Wahl-Jorgensen and T. Hanitzsch (eds), *Handbook of Journalism Studies*, New York: Routledge, pp. 218-

33.

Messing, S., & Westwood, S. (2014). Selective Exposure in the Age of Social Media:

Endorsements Trump Partisan Source Affiliation When Selecting News Online.

Communication Research, 41(8), 1042-1063. doi:10.1177/0093650212466406

Meyer, R. (2016, May 26). How many stories do newspapers publish per day? *The Atlantic*.

Retrieved from <https://www.theatlantic.com/technology/archive/2016/05/how-many-stories-do-newspapers-publish-per-day/483845/>

Mitchell, A. & Holcomb, J. (2016, June 15). State of the news media 2016. Pew Research

Center *Journalism & Media*. Retrieved from

[http://www.journalism.org/2016/06/15/state-of-the-news-media-](http://www.journalism.org/2016/06/15/state-of-the-news-media-2016/?utm_content=buffer6871f&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer)

[2016/?utm_content=buffer6871f&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer](http://www.journalism.org/2016/06/15/state-of-the-news-media-2016/?utm_content=buffer6871f&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer)

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

Newman, N., Fletcher, R., Levy, D. A. L. & Nielsen, R. K. (2016). Reuters Institute Digital News Report 2016. Retrieved from

<https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital-News-Report-2016.pdf>

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L. & Nielsen, R. K. (2017). Reuters Institute Digital News Report 2017. Retrieved from

https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017%20web_0.pdf?utm_source=digitalnewsreport.org&utm_medium=referral

- Pariser, E. (2012). *The filter bubble: How the new personalized Web is changing what we read and how we think*. New York, N.Y.: Penguin Books/Penguin Press, 2012.
- Parse.ly (2017). The Authority Report: Redefining bounce rate with engaged time.
Retrieved from <https://www.parse.ly/resources/data-studies/authority-report-14/>
- Perez, S. (2017, August 2). Facebook will soon start ranking faster loading webpages higher in News Feed. *Technocrunch*. Retrieved from
<https://techcrunch.com/2017/08/02/facebook-will-soon-start-ranking-faster-loading-webpages-higher-in-news-feed/>
- Plesner, U. (2009). An actor-network perspective on changing work practices: Communication technologies as actants in newswork. *Journalism*, 10(5), 604-626.
doi:10.1177/1464884909106535
- Puschmann, C., Burgess, J. (2014). Metaphors of Big Data. *International Journal of Communication*, pp. 1690-1709.
- Renés-Arellano, P., Hernandez-Serrano, M., Graham, G., & Greenhill, A. (2017). From prosumer to “prodesigner”: Participatory news consumption. *Comunicar*, 25(50), 77-87.
- Segall, L. & Stetler, B. (2015, October 6). Twitter launches Moments. *CNNTech*. Retrieved from <http://money.cnn.com/2015/10/06/technology/twitter-moments/index.html>
- Shoemaker, P. J., & Vos, T. P. (2009). *Gatekeeping theory*. New York: Routledge, 2009.
- Silverman, D. (2013). *Doing qualitative research: A practical handbook*. SAGE Publications Limited.
- Singer, J. (2013). User-generated visibility: Secondary gatekeeping in a shared media space. *New Media & Society*, 16(1), 55-73.

- Smythe, D. W. (1981). On the audience commodity and its work. *Media and cultural studies: Keywords*, 230-56.
- Stern, R. (2016). Betting on technology: The diffusion of mobile live-streaming apps within news startup actor-networks (Master's thesis). Retrieved from MOspace Database. Retrieved from <https://mospace.umsystem.edu/xmlui/handle/10355/56193>
- Sumpter, R. S. (2000). Daily newspaper editors' audience construction routines: A case study. *Critical Studies In Media Communication*, 17(3), 334.
- Tandoc, E. (2014). Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society*, 16(4), 559-575.
- The Report of the 2020 Group. (2017, January). Journalism that stands apart. *The New York Times*. Retrieved from <https://www.nytimes.com/projects/2020-report/>
- Tufekci, Z. (2014). Big questions for social media Big Data: Representativeness, validity and other methodological pitfalls. In ICWSM '14: Proceedings of the 8th International AAAI Conference on Weblogs and Social Media, 2014.
- Tufekci, Z. (2016, May 19). The Real Bias Built in at Facebook. *The New York Times*. Retrieved from <https://www.nytimes.com/2016/05/19/opinion/the-real-bias-built-in-at-facebook.html>
- Turner, F. (2005). Actor-networking the news. *Social Epistemology*, 19(4), pp. 321-324.
- Turow, J. (2011). *The daily you: How the new advertising industry is defining your identity and your worth*. New Haven: Yale University Press, 2011.
- The New York Times Innovation Report. (2014). Retrieved from NiemanLab: <http://www.niemanlab.org/2014/05/the-leaked-new-york-times-innovation-report-is-one-of-the-key-documents-of-this-media-age/>

- Teurlings, J. (2013). Unblackboxing production: What media studies can learn from actor-network theory. In M. de Valck, & J. Teurlings (Eds.), *After the break: Television theory today* (pp. 101-116). (Televisual culture). Amsterdam: Amsterdam University Press.
- Usher, N. (2013). Al Jazeera English online: Understanding web metrics and news production when a quantified audience is not a commodified audience. *Digital Journalism, 1*(3), 335-351.
- Vanderbilt, T., & Paskin, W. (2013). The New Rules Of The Hyper-Social, Data-Driven, Actor-Friendly, Super-Seductive Platinum Age Of Television. *Wired, 21*(4), 90.
- Van Dijck, J. (2012). Facebook and the engineering of connectivity: A multi-layered approach to social media platforms. *Convergence, 19*(2), 141-155.
- Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press.
- Van Dijk, T. A. (2009). News, discourse, and ideology. *The handbook of journalism studies*, 191-204.
- Vu, H. (2014). The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism, 15*(8), 1094-1110. doi:10.1177/1464884913504259.
- Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. Cambridge, Massachusetts: The MIT Press.
- Weiss, A., & Domingo, D. (2010). Innovation processes in online newsrooms as actor-networks and communities of practice. *New Media & Society, 12*(7), 1156-1171.
- Westlund, O. (2013). Mobile news: A review and model of journalism in an age of mobile journalism. *Digital journalism, 1*(1), 6-26.

- Whyte, W. F. (1982). Interviewing in field research. In R. G. Burgess (Ed.), *Field research: A sourcebook and field manual* (pp. 111-121). London & New York: Routledge.
- Yin, R. (2014). *Case Study Research: Design & Methods*. Thousand Oaks, CA: Sage Publications, Inc.
- Zelenkauskaite, A. (2016). Remediation, convergence, and big data: Conceptual limits of cross-platform social media. *Convergence: The International Journal of Research into New Media Technologies*, 1-16, doi: 10.1177/11354856516631519.
- Zelizer, B. (1993). Journalists as interpretive communities. *Critical Studies in Mass Communication*, 10(3), 219.