Digital Deceptions:

Unveiling the Impact of Pseudo-Local News on Democracy and Crafting Countermeasures

(Metric Media Case Study)

André-Joseph (AJ) Cordeiro

A Thesis

In

The Department

Of

Journalism

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Arts (Digital Innovation in Journalism) at

Concordia University

Montreal, Quebec, Canada

March 2025

Concordia University

School of Graduate Studies

This is to certify that the thesis prepared

By: André-Joseph (AJ) Cordeiro

Entitled: Digital Deceptions: Unveiling the Impact of Pseudo-Local News on Democracy and Crafting Countermeasures (Metric Media Case Study)

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Digital Innovation in Journalism)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair

Dr. Andrea Hunter

_____ Examiner

Dr. David Secko

_____ Examiner

Dr. Magda Konieczna

_____ Supervisor

Dr. James McLean

Approved by _____

Chair of Department or Graduate Program Director

_____ 2025 _____

Dean of Faculty

Abstract

Unveiling the Impact of Pseudo-Local News on Democracy and Crafting Countermeasures (Metric Media Case Study)

AJ Cordeiro

Metric Media is known for its vast network of US-based pseudo-local news sites propagating specific ideologies, operating beyond traditional mass media manipulation techniques. By micro-targeting societal segments on divisive issues, Metric Media presents a nuanced form of content dissemination, raising questions about its impact on public opinion and welfare. The case study employed sentiment analysis, topic modelling, and social media engagement metrics to analyze content from July 2019 to January 2024. It offered a comprehensive view of Metric Media's thematic focus and the resonance of its narratives among communities. Additionally, it explored the role of data scraping and visualization in highlighting the organization's strategic approaches to news dissemination. This investigation not only elucidated Metric Media's use as a political tool across ideologies but also underscored the importance of data-driven methodologies in enhancing public awareness of media influence. The findings contribute to the discourse on digital media manipulation, emphasizing the need for critical engagement with news sources.

Keywords: local news, open source intelligence (OSINT), data journalism, propaganda

Acknowledgements

Every finished piece of work—especially one with a spine, page numbers, citations, and footnotes—tends to have a single name on the cover. That name, in bold, suggests ownership, authorship, authority. The myth of the solitary scholar and journalist persists: a lone figure hunched at their desk, lit by the blue glow of a laptop screen, fueled by little more than caffeine and stubbornness.

But it is, of course, fiction.

This work is not mine alone. It is the culmination of a thousand kindnesses, a decade of conversations, and a lifetime of quiet, often invisible support. And, as many stories worth telling do, it begins with family.

My parents are the kind of people who do hard things quietly. They immigrated to Canada to pursue something they could not name but could feel: possibility. They did not come for glory, or prestige, or praise. They came for us, my two siblings and me. Every sacrifice they made—each dollar saved, each comfort foregone, each racist insult endured—was calibrated to give their children a better life. They gave me a foundation. But more than that, they gave me an engine—a sense that success is not about spotlight, but about who and how many you bring along with you. Thank you, Papa. And thank you, especially, Mama—for the heavy lifting you did daily, without fanfare and credit. In this world, it is easy for work like yours to go unrecognized. In the world I now inhabit, we believe citations matter. Consider this one long overdue.

To my bonus family—my in-laws—thank you for your patience. You've tolerated dinner-table monologues about journalism and unsolicited commentary on the six o'clock news when all you wanted was to check the weather. For nearly two decades, you've endured an off-duty journalist who, truthfully, never is. Thank you for making space, for nodding along, and for understanding that sharing a life doesn't always mean sharing a field.

To my friends and extended family who stayed close through five long, often strange years—thank you. You didn't ask about footnotes or formatting. You asked the real questions. Are you sleeping? Have you eaten? When was the last time you saw the sun? To Eva, Marco, Cynthia, and Salim—your presence was ballast. You were lifelines, sounding boards, and a reminder that a well-rounded life is built outside the works cited. You showed up. And that matters.

To Dr. James McLean, my thesis supervisor and sometimes lifeguard, thank you for wading into the deep end with me. I arrived with an unwieldy topic and an archive of scattered thoughts, and you met every rambling email with calm clarity and razor-sharp insight. There were, unquestionably, too many semi-coherent questions from my end, what often could be charitably described as "earnest but ill-considered.". You answered each one as if it were the most reasonable query in the world. If coherence exists here, it is due to your logic. If style occasionally rises above "functional," it is due to your faith in letting me explore.

To the professors I encountered during my MA—thank you for your generosity. You modelled not just scholarship but integrity. You gave me permission to be unsure and the tools to navigate uncertainty with curiosity instead of fear. You opened the door and let me wander in. You gave me room to ask questions, including the wrong ones, which is the mark of real mentorship.

To my colleagues, the journalism students I've worked alongside over the past ten years—thank you for teaching me why we teach and how to be a better journalist. You reminded me that learning isn't something we give to others; it's something we do with them. I wish I could name every one of you. Know that you're there—in the margins of this work, and in every moment I share in the CJ. Your questions made me sharper. Your ideas kept me honest. You made me want to be better.

Robin Della Corte and Cristina Sanza—what a gift to teach alongside you. Thank you for the camaraderie, the encouragement, and for being the kind of colleagues who show up, over and over, in big ways and small. Through the chaos and quiet, through deadlines and doubts, your presence was grounding. You made the work easier. You made the hard days lighter.

To Dr. David Secko and Dr. Brian Gabrial: thank you for believing in my work even before I did. You provided both structure and spark, and your support reverberated long past the classroom. And to the inimitable Professor Linda Kay, my first supervisor at Concordia—thank you for never hoarding your wisdom, offering guidance freely and warmly, and making the world of journalism feel not just possible but essential.

A quiet but emphatic thank you to Benjamin Languay, whose gentle persistence may very well be why this thesis met its deadline (and, of course, his comment was: five years was the deadline?). He was not on my supervisory committee, but functionally? He was the deadline whisperer.

To *the Arsenal*: thank you for showing that excellence is not just an end goal but a way of building. From academy to title winners—there is no better metaphor for mentorship and leadership, whether on the pitch or in a newsroom. AWFC, U-18, U-21, AFC. Victoria Concordia Crescit. And to the Habs—*tricolore jusqu'au bout*. Les Canadiens toujours.

And then there is Peter Downie. There are teachers, there are journalists, and there are mentors, and then there is Peter—who somehow manages to be all at the highest level and something more. Over a decade ago, when I was still figuring out what kind of journalist I wanted to be, Peter saw something I did not. He encouraged me to apply to the MA program back when the phrase "graduate school" felt aspirational at best, laughable at worst. I said (and still say) that I am no academic and not suited to this world. Peter said otherwise. He was the first to bridge journalism and academia in a way that felt real, rigorous, and necessary. His belief in professional development—actual development, not just performance—helped shape my own approach to working with others, including our TA teams. And leading by example, his boundless curiosity and openness, and his refusal to disengage from a messy, complicated world, remind me that optimism is not naïveté. It is a choice. And it is one I try, daily, to make. We are not Concordia Journalism without Peter Downie. We have no Newsroom without Peter Downie. We are not journalists without Peter Downie.

And finally—Val.

You are the axis on which this whole operation turns. You are never "supporting" from the sidelines; you are in it, with me, always. From introducing you to data science to then seeing you become mine and others' data guru, you made it legible, but more importantly, you made the life around the thesis bearable. When the research was messy, you brought order. When I doubted, you didn't flinch. You brought a steadiness that I borrowed, often and without asking. You are the person behind every sentence, every idea, every finish line I managed to cross. And as Peter (and anyone else with eyes and a brain) would say: you're a saint. If there's anything in this thesis—or in this world—that is kind or good or worthwhile, it's all because of you. ILYWAMH. Five Fingers. Always.

PS (because I don't have any copy editors telling me to cut down this section!):

I want to thank my colleagues I have worked with on TA teams over the past ten years for teaching me more than any course ever could. Your curiosity and creativity remind me daily why we do this. This includes but is not limited to (and in no particular order): Autumn Darey, Jared Lackman-Mincoff, Geneviève Sylvestre, Conor Tomalty, Emily Pasquarelli, Sarah-Maria Khoueiry, Tristan McKenna, Marcus Werner, Evan Lindsay, Cedric Gallant, Kaity Brady, Matthew Skelhorne, Anthony Issa, Gabriel Guindi, Clément Lechat, Esteban Cuevas, Hadassah Alencar, Talia Kliot, Marieke Glorieux-Stryckman, Emily Palmer, Kaaria Quash, Benjamin Languay, Camille Perrin-Ridgway, Luca Caruso-Moro, Laurence Brisson Dubreuil, Katherina Boucher, Calvin Cashen, Sara Chahrour, Caitlin Yardley, Marie-Josée Kelly-Lelièvre, Michelle Pucci, Leah Batstone, Audrey Folliot, Natasha Taggart, Vanik Kechian, Joseph Arciresi, Kathleen Sheridan, Sara Siobhan King-Abadi, Coraline Mathon, Evgenia Choros, Diandra D'Alessio, Julian McKenzie, Kassandra Nadeau-Lamarche, Arian Roz Zarrinkoub, Eva-Marie Kratochvil D'Amico, Kelsey Litwin, Nikolas Litzenberger, Matthew Shanahan, Salim Valji, Christopher Aitkens, Julia Bryant, Brian Lapuz, Elysia-Marie Campbell, Alexander Cole, Patrick Mocella, Jessica Ramirez Jaschek, Alyssia Rubertucci, Emmanuel Barbier, Rafael Figueroa-Goetschi, Erdene Batzorig, Carl Bindman, Simon New, Shakti Langois-Ortega, Frank Pavan, Luca Caruso-Moro, Savanna Craig, Nicholas Di Giovanni, Shanellie Desparois, and Diona-Rosan Macalinga.

Table of Contents

Chapter 1: Introduction……...01

Chapter 2: Literature Review……..05

Chapter 3: Methodology……...32

Chapter 4: Findings……...63

Chapter 5: Discussion…..……....90

Chapter 6: Conclusion…..……..114

Chapter 7: Works Cited…….……..118

1. Introduction: Micro-Targeted Propaganda and the New Crisis of Local Journalism

Journalism has traditionally played a pivotal role in informing communities about matters of public concern and fostering debate in the public sphere. A 2022 UNESCO report highlights journalism as "a necessary condition of an open society and healthy democracy" (UNESCO, 2022, p.18). However, this essential democratic function has never been entirely immune to negative influences. Historically, ultra-wealthy individuals have manipulated the media to serve as instruments of propaganda. One prominent example is William Randolph Hearst's newspaper chain, which significantly influenced cannabis prohibition in the early 20th century by linking marijuana use to racial fears in a bid to protect his paper mill investments (French & Manzanarez, 2004, p.129). This illustrates how media can be weaponized to serve private interests at the expense of the public good.

In an age dominated by data, algorithms, and fractured information ecosystems, journalism faces a new adversary: micro-targeted propaganda disguised as local news. Organizations like Metric Media exploit digital tools not only to spread ideologically driven messages but to do so under the veneer of community reporting. This emerging phenomenon challenges the foundations of public trust, democratic deliberation, and journalistic integrity.

The rise of pseudo-local news operations represents a fundamental shift in the information environment. Once anchored in the principles of accountability and objectivity, local journalism is increasingly crowded out by content farms engineered to mimic its tone, format, and authority. Metric Media is emblematic of this transformation. With over a thousand domain-specific news websites across the United States, it delivers ideologically

selective content cloaked in the aesthetics of hyperlocal reporting. Often automated and data-driven, these platforms deliver targeted messages designed to influence specific demographic groups while evading traditional forms of journalistic scrutiny.

This thesis examines how Metric Media functions as a case study in weaponizing digital technologies to simulate journalism while enacting ideological manipulation. It investigates the organization's structural, technological, and rhetorical strategies to reveal how micro-targeted disinformation is embedded in the mechanics of modern media production. In doing so, it explores a broader crisis in the media ecosystem: the erosion of credibility, the co-opting of local trust networks, and the algorithmic optimization of propaganda.

The research is grounded in three intersecting theoretical frameworks: Herman and Chomsky's Propaganda Model, McCombs and Shaw's Agenda-Setting Theory, and Shoshana Zuboff's Surveillance Capitalism. The Propaganda Model illuminates how economic and political interests shape ostensibly neutral news; Agenda-Setting Theory provides a lens for understanding how selective emphasis frames public priorities; and Surveillance Capitalism underscores the monetization of personal data to engineer behavioural prediction and influence. Together, these frameworks offer a comprehensive lens to understand Metric Media's operation as both a media enterprise and a political actor.

The central research objective of this study is to analyze the strategies through which Metric Media manufactures legitimacy, micro-targets audiences, and shapes ideological narratives. Specifically, the study seeks to:

1. Investigate the methods used by Metric Media to create, host, disseminate, and monetize pseudo-news platforms.

2. Evaluate the effectiveness of automated content distribution, social media engagement, and geo-specific messaging.
3. Explore practical strategies for journalists, scholars, and the public to detect and counteract covert propaganda.
4. Examine the potential to repurpose similar digital tools to identify and expose manipulative content.

To guide this investigation, the following research questions were developed:

- **RQ1: How does Metric Media construct and disseminate its pseudo-news ecosystem?**
- **RQ2: What are the identifiable patterns in its content strategy, particularly in sentiment, topic emphasis, and platform targeting?**
- **RQ3: How can digital journalism practitioners and the public detect, deconstruct, and resist such micro-targeted propaganda?**
- **RQ4: Can data-driven methods used for manipulation also be mobilized for accountability?**

Methodologically, this research adopts a multi-method digital humanities approach. It combines web scraping, natural language processing, sentiment analysis, topic modelling, and network visualization to examine over 7,500 posts collected from Metric Media websites and affiliated Facebook pages. These computational techniques are not merely tools for data analysis; they are instruments of critique used here to reverse-engineer algorithmically enhanced ideological messaging tactics. In doing so, this thesis not only documents the influence strategies of Metric Media but also models a reproducible workflow for scholars investigating digital propaganda in future contexts.

This study is significant in two ways. Academically, it contributes to the expanding field of computational journalism studies by operationalizing classic media theory through new media analytics. Practically, it equips media professionals, policymakers, and the public with insights into the emerging tactics of digital disinformation networks. At a time when democratic institutions face existential threats from within the information sphere, such research is urgently needed.

Chapter Two surveys the relevant literature, identifying the theoretical and empirical contributions of propaganda theory, agenda-setting, and surveillance capitalism to our understanding of digital news manipulation. Chapter Three outlines the study's computational methodology, including data acquisition, cleaning, analysis, and visualization. Chapter Four presents the empirical findings, focusing on sentiment distribution, topic clusters, and platform engagement. Chapter Five interprets these findings through the theoretical frameworks, identifying implications for democratic discourse, media policy, and the future of journalism. Finally, Chapter Six concludes by summarizing key contributions, limitations, and areas for future research.

This thesis argues that pseudo-local news networks like Metric Media represent a critical juncture in the evolution of propaganda. By merging community journalism's aesthetics with surveillance capitalism's infrastructure, these platforms weaponize trust for ideological ends. Kovach and Rosenstiel's seminal work, The Elements of Journalism, underscores journalism's core commitment to serving citizens by providing content "without fear or favor" (American Press Institute, 2013). Exposing and countering such strategies is not only an academic exercise but a democratic imperative.

2. Literature Review

Rapid advancements in digital technologies have irrevocably transformed the landscape of journalism. Traditional media platforms—print newspapers, radio, television—remain relevant, but the evolution of digital networks has introduced new opportunities and challenges for news production and consumption. At the heart of these changes is the concept of propaganda, which has been reconfigured in expression and impact by the digital environment. This literature review examines key theoretical frameworks in media and communication–propaganda, agenda-setting, and surveillance capitalism – and integrates emerging concepts like data colonialism, algorithmic curation, and platform governance. Throughout, the U.S.-based case of Metric Media, a network of pseudo-local news sites, serves as a central example of contemporary propaganda and digital disinformation.

Propaganda in Journalism: Theoretical Foundations

Propaganda is traditionally understood as the manipulation of public opinion for specific ideological, political, or commercial ends (Smith, 2015). It represents a form of communication aimed at influencing attitudes and behaviours to further the intent of the message source. Scholars have long debated precise definitions. Jowett and O'Donnell (2012) offer a widely cited definition: "Propaganda is the deliberate, systematic attempt to shape perceptions, manipulate cognitions, and direct behavior to achieve a response that furthers the desired intent of the propagandist" (p.7). This definition emphasizes intentionality (deliberate, systematic) and the multi-faceted targeting of perception, thought, and action.

In contrast, Jacques Ellul (1965) approached propaganda as a pervasive sociological phenomenon ingrained in modern society rather than merely a series of deliberate acts. Ellul argued that in technologically advanced societies, propaganda becomes an ambient process that shapes virtually all biased messaging, often blurring the line between propaganda and ordinary social discourse (Ellul, 1965). He suggested propaganda leverages the prevailing social environment—what people share in motivations and myths—such that "the individual never is considered as an individual, but always in terms of [their average traits shared with others]." (Ellul, 1965, p. 90). In Ellul's view, propaganda's potency lies in its ubiquity and ability to fuse with truth (using facts, half-truths, or lies as needed) to the point that people almost need propaganda to navigate mass society (Ellul, 1965).

The interplay between propaganda and journalism has presented enduring ethical dilemmas. Journalism traditionally values objectivity, truthfulness, and serving the public interest, whereas propaganda often prioritizes a partisan goal over unbiased truth. Where does persuasive yet ethical journalism end and propaganda begin? Scholars note that this line can be nebulous. Jowett and O'Donnell's definition implies that propaganda involves an intent to deceive or manipulate for the propagandist's benefit, whereas honest journalism seeks to inform for the public's benefit (Jowett & O'Donnell, 2012). However, another might defend what one observer labels propaganda as responsible advocacy or editorial slant. A report viewed as propagandistic by one scholar may be seen as legitimate advocacy by another. Marlin (2013) argues that the line between persuasion and propaganda is frequently blurred, particularly in democratic societies where strategic communication is central to public discourse. Moreover, Bakir et al. (2018) note that in the current media environment—characterized by information abundance and fragmented publics—the distinction between advocacy, journalism, and propaganda becomes even more difficult to

maintain. This subjectivity complicates scholarly consensus on identifying propaganda in news.

Classical Propaganda Theories and News Media

Early propaganda studies emerged from wartime and political contexts but laid foundational concepts for understanding media influence. During World War I, governments demonstrated how mass media could be mobilized to shape public sentiment on a large scale (Taylor, 2003; Welch, 2014). Edward Bernays' concept of the "engineering of consent" illustrated the persuasive power of propaganda when media channels are saturated with coordinated messaging (Bernays, 1947). Harold Lasswell's seminal analysis of World War I propaganda emphasized how symbols, slogans, and emotional appeals effectively mobilized public opinion and legitimized state actions (Lasswell, 1927; Sproule, 1997). These historical developments set the stage for analyzing propaganda not only in times of conflict but also within peacetime media systems, where subtle forms of influence persist (Jowett & O'Donnell, 2012).

In the late 20th century, Edward Herman and Noam Chomsky's propaganda model (1988) provided a structural theory of how news media can serve propaganda functions even in liberal democracies. In Manufacturing Consent, Herman and Chomsky argue that systemic factors–notably media ownership, advertising revenue, dependence on official sources, flak (negative feedback to discipline the press), and prevailing anti-ideologies–act as filters that shape news output (Herman & Chomsky, 1988, pp. 2-3). Together, these filters result in news content that tends to favour the interests of powerful political and corporate actors, effectively propagating elite viewpoints to the public (Herman & Chomsky, 1988, pp. 2-3). For example, concentrated private ownership of media and the profit motive

can limit critical reporting on corporate or government malfeasance, even without overt censorship. The propaganda model is methodological in nature: it was formulated through institutional analysis and case studies of media coverage (e.g., disparities in reporting genocides or elections) (Herman & Chomsky, 1988, pp. 2-3). Its theoretical contribution highlights that propaganda is not only about falsehoods but also about framing and filtering–the subtle privileging of certain narratives–under the guise of news.

However, a criticism of the propaganda model is that it downplays instances of media autonomy and investigative journalism that run counter to elite interests. Critics argue that while structural constraints strongly influence media content, they do not wholly determine it; moments of journalistic resistance and investigative reporting have emerged, particularly during high-profile political scandals or crises (Cottle, 2006; McChesney, 2004). Nonetheless, Kellner (1992) was less hopeful, noting that the rise of television and now digital media has deeply blurred these lines, allowing for propaganda to operate through the medium's spectacle and affective appeal rather than traditional didactic forms (p. 65). Contemporary scholars have similarly argued that the emotive and immersive nature of digital platforms reinforces propaganda's reach by integrating it seamlessly into entertainment, personalization algorithms, and user-generated content (Fuchs, 2018; Freedman, 2014).

Other theorists have examined propaganda through different lenses. Jowett and O'Donnell's own work not only defined propaganda but also differentiated it from persuasion, stressing that propaganda is often one-way, manipulative communication lacking the mutuality of true persuasion. They advocate analyzing propaganda by examining the institutions and techniques behind messages – for example, identifying the propagandist, the structure of the propaganda organization, media utilization, and

audience targeting (Jowett & O'Donnell, 2012, pp. 19-23). This analytical framework has been methodologically useful for case studies ranging from political campaigns to public relations spin. Scholars like Leo Bogart (1995) studied propaganda in government information agencies, focusing on the message strategies and how they aim to sway public opinion.

Florian Zollmann (2017) offers a contemporary view linking these ideas to modern journalism. He contends that "propaganda is not just about spreading falsehoods; it's about framing the discourse in a way that favors certain political outcomes" (Zollmann, 2017, p.92). Zollmann's study, which examines media coverage of political interventions, underscores that modern propaganda often operates by emphasizing or omitting certain facts to shape narratives rather than blatantly lying (Zollmann, 2017, p. 108). This perspective aligns with agenda-setting and framing theories, bridging classical propaganda theory with modern media effects research.

Propaganda in the Digital Age: Metric Media as Example

Digital media have both amplified the reach of propaganda and obscured its source. In the current era, propaganda can manifest as seemingly grassroots or "organic" content online, complicating efforts to recognize it. Metric Media is an illustrative case: a network of websites that present themselves as local news outlets but, in fact, disseminate politically loaded content on a massive scale. These sites, described as "pink slime" journalism by researchers, blur the distinction between genuine local journalism and orchestrated propaganda (Begani, 2020). Bengani (2019, 2020, 2021) documented how Metric Media and affiliated networks launched hundreds of local-looking news sites that algorithmically generate stories and promote partisan talking points under the veneer of community news.

In line with propaganda, the intent is to shape public perceptions at the local level (for instance, casting certain candidates or policies in a favourable or unfavourable light) while concealing partisan sponsorship.

What makes Metric Media a paradigmatic example of digital-age propaganda are the techniques employed. Over 90% of content across its sites was found to be automatically produced from public data or lightly rewritten from other sources (Begani, 2020). This low-cost, high-volume production leverages algorithms to flood information channels, echoing Ellul's notion that technological society enables propaganda to be pervasive and constant (Ellul, 1965). Moreover, these sites often carried a conservative bias in topic selection (e.g. emphasizing themes like government waste or the merits of deregulation). By flooding local information ecosystems—many of which are "news deserts" with few or no traditional local newspapers—Metric Media effectively frames which issues appear salient in those communities. This is a strategic exploitation of agenda-setting power. Crucially, Metric Media's operations were backed by political and ideological actors: investigations using open-source intelligence and IRS records uncovered funding ties to conservative advocacy groups and even a multi-million dollar campaign targeting the 2020 U.S. presidential race (Bengani, 2021). Such findings illustrate that behind the façade of local journalism, deliberate propaganda efforts are at work, aligning with Jowett and O'Donnell's emphasis on identifying the propagandist and their intent.

Methodologically, studies of Metric Media (e.g., Bengani 2019, 2021) combined content analysis (to detect bias and automation in the news output) with investigative research into funding and ownership. This mixed-methods approach is notable: it not only characterizes the content as propagandistic but also reveals the infrastructure and motive–essential for confirming something as propaganda rather than mere low-quality

journalism. The case of Metric Media underscores why updated frameworks are needed: traditional propaganda theory focuses on overt propaganda (e.g., government posters, broadcasted statements), whereas today, propaganda can hide in plain sight within what looks like independent journalism. This calls for integrating theories of propaganda with theories of digital media, algorithmic dissemination, and platform dynamics.

Agenda-Setting and Media Influence

Closely related to propaganda is the question of who or what determines the topics people think about. Agenda-setting theory, first formulated by McCombs and Shaw (1972), addresses the media's influence on issue salience in the public mind. In their seminal Chapel Hill study during the 1968 U.S. presidential election, McCombs and Shaw content-analyzed news and surveyed voters, finding a strong correlation between the issues that dominated media coverage and the issues voters identified as important (McCombs & Shaw, 1972). In other words, the media appeared successful in telling people what to think about, if not what to think (McCombs & Shaw, 1972, p. 177). This first-level agenda-setting effect has been widely replicated: the more coverage an issue receives, the more the public tends to rank it as a key concern (McCombs, 2005, p. 547). The consistent finding that media priorities and public priorities align suggests a significant media effect on public perception of reality (McCombs, 2005, pp. 546-547).

Agenda-setting is a foundational media-effects theory with important theoretical implications. It highlights the transfer of salience: media acts as a gatekeeper, filtering a vast array of news into a limited set of headlines and thereby shaping the perceived agenda of the day (McCombs & Shaw, 1972). This gatekeeping role gave rise to further concepts like second-level agenda-setting (how media not only tell us what to think about but also

how to think about it via framing attributes) (McCombs et al., 1997) and intermedia agenda-setting (how issues transfer between elite media, local media, and now social media) (Harder et al., 2017; Meraz, 2009).

One criticism of classic agenda-setting theory is the difficulty of disentangling causality—does media coverage shape the public's agenda, or does media simply reflect existing public concerns? Some scholars argue that the relationship can be bidirectional or mediated by real-world conditions, such as economic crises or natural disasters, which independently heighten public concern regardless of media emphasis (Wanta & Hu, 1994; Vliegenthart & Walgrave, 2008). Nevertheless, a substantial body of empirical research supports the notion that media significantly influence public issue salience, particularly for topics where individuals lack direct experience and must rely on mediated information (McCombs & Reynolds, 2002; Scheufele & Tewksbury, 2007). This influence is most pronounced in the first level of agenda-setting, which concerns issue salience but extends into the second level, affecting the perceived attributes of those issues.

In today's digital context, traditional media's agenda-setting role is increasingly challenged and supplemented by social media platforms and algorithmic curation. On the one hand, the fragmentation of information sources means the public's agenda can be influenced by a wide array of online voices—including trending hashtags, influencers, and partisan outlets like Metric Media—not just mainstream news editors (Chadwick, 2013; Kreiss, 2016). On the other hand, the algorithms that govern news feeds on platforms such as Facebook and X (formerly Twitter) act as new agenda setters, determining which stories users encounter most prominently (Tufekci, 2015; Napoli, 2019). Research indicates that social media can still reflect traditional media agendas to a large extent—for example, a major newspaper's story might trend on Twitter, illustrating intermedia agenda-setting

(Vargo et al., 2014). However, new patterns emerge: personalized algorithms mean each user may effectively receive a different agenda, tailored to their past behaviour—a phenomenon linked to "filter bubbles" and "algorithmic gatekeeping" (Pariser, 2011; Zuiderveen Borgesius et al., 2016). This raises the question of whether there is still a coherent public agenda or a multiplicity of fragmented micro-agendas. Scholars such as Meraz (2011) and Vargo and Guo (2017) have explored network agenda-setting, demonstrating that networks of users and media co-create issue salience in increasingly complex and decentralized ways.

The Metric Media phenomenon has implications for agenda-setting at the local level. In communities where these pseudo-local sites proliferate, the issues they highlight (often aligned with a partisan agenda, such as crime, taxes, or particular ideological talking points) may dominate the local conversation in the absence of competing news sources. If residents encounter Metric Media stories on, say, Facebook or Google search results when seeking local news, those issues could become top-of-mind. This is agenda-setting through information voids: by filling a gap left by shrinking local newspapers, Metric Media attempts to define the local news agenda. Whether this succeeds is an empirical question–initial evidence suggests many such "pink slime" sites have relatively low engagement or visibility on their own (Bengani, 2020). However, even low-visibility outlets can have outsized influence if their content is amplified through social sharing or political campaigns. For example, local political actors or interest groups could circulate a Metric Media article with a sensational claim, thus reaching audiences indirectly. This dynamic underscores that agenda-setting is a joint product of media content and platform-mediated distribution in the digital age, requiring us to consider algorithmic and network factors alongside classical media effects.

It is also important to note that while agenda-setting focuses on issue salience, propaganda often goes further to influence interpretation and action. However, controlling the agenda is a powerful first step for any propagandist: if one can insert certain topics into public debate (e.g., exaggerated claims of voter fraud in local elections via a network of sites), one has already guided public attention and possibly prepared the ground for further persuasion. Thus, agenda-setting theory provides a bridge between broad media influence and the more pointed manipulations studied in propaganda research.

Methodologically, agenda-setting studies in the digital era have expanded to include big data analytics (e.g., analyzing millions of tweets or posts to see which issues trend) and experiments (testing how exposure to certain topics affects perceived importance). These newer methods complement the traditional content analysis and survey approach. They also reveal nuances, such as the speed of agenda-setting effects (digital media can set agendas in hours rather than days) and the role of algorithmic biases (platforms may boost certain topics for engagement, inadvertently promoting some agendas over others).

Misinformation, Disinformation, and "Fake News" in the Digital Sphere

The digital age has seen an explosion of concerns around "fake news," misinformation, and disinformation, which intersect with propaganda but are distinct concepts that warrant clear definitions. Misinformation generally refers to false or misleading information shared without the intent to deceive, often spread by people who believe it to be true (Wardle & Derakhshan, 2017). Disinformation, by contrast, is false information deliberately created or distributed with the intent to deceive or mislead for a strategic purpose (Wardle & Derakhshan, 2017). In short, disinformation is a subset of misinformation distinguished by the element of intent and harm. Wardle and Derakhshan

(2017) propose using these terms as part of an "information disorder" framework. They define disinformation as content that is intentionally false and designed to cause harm, motivated by factors like political power, financial gain, or malice. When such deceptive content circulates organically, people who unwittingly share it generate misinformation, which describes the same false content but without the sharer's intent to mislead (Wardle & Derakhshan, 2017). In practice, a piece of fake news may originate as disinformation (e.g., a fabricated story by a propagandist) and then be propagated by well-meaning individuals as misinformation.

It is crucial to delineate these terms in a study of digital propaganda. Propaganda, as discussed, often involves a mix of truth and lies crafted to persuade. Disinformation is closely related – many propaganda campaigns rely on deliberate falsehoods (disinformation) to achieve their goals. However, propaganda could also involve selectively true information presented in biased ways (which might not qualify as disinformation if strictly true but could still be misleading). Thus, the misinformation/disinformation definition is used to discuss falsity and accuracy, and propaganda to emphasize intent and strategy (which may or may not involve falsehoods).

The term “fake news” gained prominence in public discourse around 2016–2017, especially during and after the U.S. presidential election. However, scholars have since debated its usefulness as an analytical concept. Tandoc et al. (2017) demonstrated that “fake news” has been employed to describe a broad range of phenomena, including satire, parody, fabrication, manipulation, and propaganda. Because of this conceptual ambiguity—as well as its frequent politicization, such as when political actors use the term to discredit unfavourable but factually accurate reporting—many scholars advocate for the use of more precise terms like misinformation and disinformation (Bakir & McStay, 2017;

Egelhofer & Lecheler, 2019). Wardle (2018) argues that much of what is labelled "fake news" is not entirely fabricated. This reinforces the notion that misinformation (false or misleading information shared without harmful intent) and disinformation (deliberately false information intended to deceive) are more robust conceptual tools for understanding the complexities of information disorder in the digital age (Wardle & Derakhshan, 2017).

The Proliferation and Impact of Digital Misinformation

Online platforms have dramatically lowered the barriers to spreading false or misleading content. A viral falsehood can reach millions on social media in hours. Studies show that false news can spread "farther, faster, deeper" than truthful news on social platforms (Kunda, 1990, p. 482; Pennycook & Rand, 2020, p. 4945). Vosoughi et al. (2018), analyzing thousands of Twitter cascades, found that false stories (especially on political topics) diffused significantly more broadly than true ones, often because they evoke strong reactions such as surprise or disgust (Vosoughi et al., 2018, p. 1148). Psychological factors like confirmation bias amplify this: people are inclined to believe and share information that aligns with their pre-existing beliefs or partisan leanings. Pennycook and Rand (2020) likewise note that individuals often fail to critically evaluate content if it fits their worldview, leading to the spread of misinformation even by regular users who do not intend harm (Pennycook & Rand, 2020, p. 4945).

This has profound societal implications. Widespread misinformation can erode trust in institutions, create false public perceptions, and even influence democratic processes. During the 2016 U.S. presidential election, many fake or misleading news stories circulated online. While Allcott and Gentzkow (2017) concluded that fake news likely was not the singular determinant of the election's outcome, they acknowledged that it shaped voter

perceptions and was one of many factors impacting decision-making. In public health, misinformation about vaccines or COVID-19 has led to lower trust in medical guidance and harmful behaviours. These examples underline that misinformation/disinformation are not trivial matters of people being "mistaken"–at scale, they represent a collective risk to informed public discourse and even public safety.

However, not all scholars agree on the severity or mechanisms of phenomena like echo chambers and filter bubbles. Sunstein (2018) warns that echo chambers–environments where one's existing views are reinforced by homogeneous exposure–exacerbate polarization and misinformation (Napoli, 2019, p. 73). Others argue that the effect of echo chambers may be overstated, noting that many users actually encounter diverse views online (the so-called "cross-cutting exposures"). The consensus is that some echo chamber effect exists, but its extent can vary by platform and individual. Regardless, a clearly documented trend is that today's media environment makes it easy for false information to travel quickly and for like-minded communities to insulate themselves with preferred narratives (Vosoughi et al., 2018; Lewandowsky et al., 2012).

Digital platforms themselves play a role in this. Algorithmic curation on sites like Facebook, YouTube, or Twitter (now X) often prioritizes content that maximizes engagement—which can inadvertently favour sensational or emotionally charged misinformation. Sun (2023) describes algorithmic curation as the process by which platforms filter and recommend content based on users' past behaviour. If algorithms learn that users click on outrageous headlines, they may show more of those, creating a feedback loop. In this sense, platforms can amplify disinformation because it is often engineered to be catchy or shareable. As Tufekci (2017) noted, algorithms designed to capture attention—for

advertising profits—can be weaponized to spread misinformation and extremist content, posing threats to democratic discourse.

Metric Media's strategy intersects with these dynamics. By creating a large volume of localized "news" stories, the network increases the chance that its content will be indexed by Google or shared on Facebook by unwary readers, thus entering algorithmic recommendation systems. Indeed, according to a Columbia Journalism Review analysis, one of the goals of such pseudo-news networks is to influence opinions and collect user data (Begani, 2020). When users visit these sites or sign up for newsletters, their data can be harvested–email addresses, ZIP codes, reading habits–which can later be used to micro-target them with political advertisements or more content. In this way, the spread of disinformation and the exploitation of surveillance capitalism converge.

Combatting Digital Disinformation: Media Literacy and Platform Governance

Addressing the tide of misinformation and disinformation requires a multi-pronged approach. One essential strategy is media and digital literacy education. Scholars argue that cultivating critical thinking skills in audiences–teaching people how to verify sources, recognize biased or fake content, and understand journalistic standards–can reduce susceptibility to false information (Tandoc et al., 2017; Usher, 2013). Hobbs (2011) emphasizes that media literacy is a civic necessity in an era of information overload; audiences must be empowered to navigate content critically to maintain a healthy public sphere. This includes understanding terms like misinformation/disinformation and propagandist tactics (for example, recognizing when a "local news site" might not be what it seems, as with Metric Media).

Another approach is the reinforcement of journalistic standards and fact-checking. Traditional news organizations have increasingly adopted fact-checking desks and correction transparency, and independent fact-checkers (e.g., PolitiFact, Snopes) work to debunk viral falsehoods. Tandoc et al. (2017) note that many newsrooms now see fact-checking as part of their social responsibility, though it requires resources and may have limited reach. The challenge is scaling these efforts to match the speed of misinformation online. Some have proposed technical solutions, such as algorithms automatically detecting likely false content or trust indicators attached to news sources. Experimental tools using AI can flag dubious claims, but they raise concerns about accuracy (false positives) and censorship if implemented bluntly.

This brings us to the role of platform governance in curbing disinformation. Platform governance refers to the policies and actions by tech companies (often in conjunction with regulators) that manage content and behaviour on their services (Gorwa, 2019). In recent years, platforms like Facebook, Twitter, and YouTube have introduced measures such as removing or flagging demonstrably false content (especially related to elections or health), downranking false news in feeds, banning repeat disinformation "superspreaders," and partnering with fact-checkers to provide corrective info. There have also been external governance efforts–for example, Europe's Code of Practice on Disinformation and laws like Germany's NetzDG, which impose fines if platforms do not promptly remove illegal fake news or hate content (Gorwa, 2021). Gorwa (2019) conceptualizes platform governance as a complex interplay of legal, political, and economic relationships that structure how online content is regulated. In essence, it is a mix of self-regulation by companies, state regulation, and input from civil society.

The efficacy and ethics of platform governance measures are hotly debated. On one side, without platform intervention, the online space can become a free-for-all of disinformation that undermines democracy. On the other side, aggressive moderation runs into issues of free expression and potential bias–who decides what is "false" or "harmful" enough to remove (Shu et al., 2017)? For instance, automated content moderation algorithms can erroneously take down satire or legitimate debate, and governments may abuse "fake news" laws to silence dissent (Wardle, 2020). Scholars like Roberts (2019) and Gillespie (2018) highlight the enormous power platforms hold as information gatekeepers and the need for transparency and accountability in how they wield this power. Algorithmic governance adds another layer: platforms increasingly rely on AI to sort content (for scale), but algorithmic moderation is prone to both mistakes and lack of explainability (Gorwa, Binns, & Katzenbach, 2020).

Metric Media's existence raises specific governance questions: how should search engines or social networks treat a network of sites masquerading as local news funded by political groups? Some might argue for labelling them as partisan or lower credibility sources in news indexes (as Google and Facebook have begun to do with state-controlled media, for example). Others would warn that singling out specific outlets is a slippery slope. Currently, these "pink slime" sites largely operate freely (Bengani, 2020). The Columbia Journalism Review findings suggest that many Metric Media sites had low visibility in search and social media (Bengani, 2020), possibly because the platforms' algorithms detected low engagement or certain spam-like patterns. However, that is an incidental outcome: no formal policy prevents a well-funded disinformation local site from becoming highly visible. This gap in platform governance, dealing with structured disinformation

operations disguised as news organizations, remains an area where both researchers and policymakers are exploring solutions.

Surveillance Capitalism and Data Colonialism in Digital Media

Modern journalism and information distribution are deeply entwined with the economic logic of the Internet. Surveillance capitalism, a term coined by Shoshana Zuboff (2019), describes the dominant economic model of the tech industry: the extraction and monetization of personal data through surveillance-like practices (Zuboff, 2019, p. 61). Companies like Google, Facebook, and many others offer "free" services but track user behaviour extensively, converting it into data that can predict and influence future behaviour (usually for advertising) (Zuboff, 2019, p. 45). Zuboff's work articulates how this marks a departure from traditional industrial capitalism–here, the human experience becomes a raw material, and behavioural predictions are the product sold in "behavioral futures markets." (Zuboff, 2019, p. 77). In Zuboff's analysis, the drive for more data creates a sweeping commodification of privacy and personal life. For journalism, this has been double-edged: on one side, news organizations rely on tech platforms to reach audiences and increasingly adopt data-driven strategies; on the other, the advertising revenues that sustained journalism have been primarily usurped by these surveillance capitalists (Google and Facebook command the bulk of digital ad spend, leaving media outlets struggling) (Chyi, 2012, p. 506).

One consequence is the pressure on news outlets to maximize engagement. Bell (2016) observed that when outlets depend on online ad revenue and algorithmic distribution, they face incentives to produce clickbait or sensational content to garner clicks and shares. Petre (2018) documents the "quantification of the newsroom," where metrics

like clicks, shares, and time-on-page begin to influence editorial decisions, potentially at the expense of traditional news values. The ethical concern here is that truth and civic importance might be sacrificed for virality and ad dollars–effectively aligning journalism closer to propaganda-like output in pursuit of attention. This dynamic can indirectly facilitate misinformation: if outrageous claims get more clicks, a financially desperate outlet might be tempted to publish them or frame news in alarmist ways.

Zuboff's theory primarily critiques corporate behaviour, but as some critics note, it does not cover everything – governments also engage in data collection, and not all data use is commercial. To broaden the lens, Nick Couldry and Ulises Mejias's concept of "data colonialism" (2019) is instructive. Data colonialism asserts that the massive appropriation of social data by corporations (and others) is akin to a new form of colonialism that annexes human life as a territory for extraction (Couldry & Mejias, 2019). Just as historical colonialism seized land and resources, data colonialism seizes personal information and human interactions, turning them into assets to be mined (Couldry & Mejias, 2019). The underlying drive is the same: acquire resources (here, data) from which capital can extract economic value. Couldry and Mejias argue that this establishes a new social order in which continuous tracking and datafication of life become normalized (Couldry & Mejias, 2019). They see surveillance capitalism not as an isolated aberration but as part of the broader evolution of capitalism entangled with colonial logics (Couldry & Mejias, 2019).

These concepts help explain structural changes and power shifts in journalism and media. Platform companies (the surveillers/colonizers) like Facebook and Google have become gatekeepers for news, both technologically and economically. Napoli and Caplan (2017) note that algorithms increasingly dictate the news people consume, raising concerns about the loss of editorial autonomy and transparency. Chakravartty et al. (2018) further

argue that surveillance capitalism undermines journalism's democratic function by treating audience attention as a commodity and prioritizing data extraction over informed public deliberation.

For propaganda and disinformation actors, this environment is fertile ground. They can exploit the algorithmic targeting capabilities (a product of surveillance capitalism) to micro-target messages, and they can exploit the infrastructures of data colonialism to gather detailed profiles on populations for tailored influence. For instance, the Cambridge Analytica scandal illustrated how political operatives harvested Facebook data on millions of Americans to craft psychographic targeted propaganda during elections–a literal intersection of propaganda aims with surveillance capitalist tools (Bakir, 2020).

Metric Media's operations also intersect with these concepts. The CJR investigation noted that these "local" sites push partisan content and aim to collect user data (Bengani, 2020). By luring users under the pretense of local news, they can accumulate information (emails, demographics, reading habits), presumably fed back into political campaigning machinery (Bengani, 2020). This reflects a data colonial mindset: local communities are treated as data sources to be mined and influenced rather than as publics to be genuinely informed. Moreover, Metric Media's expansion was financed by political actors, likely because digital advertising (a product of surveillance capitalism) made it cheap to deploy hundreds of sites and automatically generate content (Begani, 2020). In earlier eras, sustaining a propaganda newspaper in many towns would be cost-prohibitive; now, one can algorithmically churn out thousands of localized articles (e.g., "High School X sports roundup" or "County Y property values up 5%") to create the appearance of a local news presence (Bengani, 2020). This data-driven approach to propaganda leverages free public

data sets (like school scores, crime stats, etc.) – essentially appropriating those data (we might say "colonizing" them) to serve a political agenda.

From a theoretical standpoint, integrating surveillance capitalism and data colonialism into media studies highlights the macro-level forces that shape information environments. They remind us that propaganda and misinformation do not exist in a vacuum; they are enabled or constrained by the economic and regulatory context. If the attention economy rewards the outrageous, propaganda will get more outrageous. If data flows are unregulated, propagandists will eagerly collect and use data just as corporations do. There is, then, a convergence of interest between platform companies that want engagement and propagandists who want influence–both benefit from content that grabs attention, often sensational or divisive.

One emerging gap is the need for more research on how these economic structures facilitate pseudo-journalism networks like Metric Media. We understand how they hurt traditional journalism (loss of revenue, etc.), but understanding how they fertilize disinformation networks is a newer area. Some scholars (e.g., Marwick & Lewis, 2017; Benkler et al., 2018) have begun mapping the political economy of disinformation, noting that an entire ecosystem of clickbait sites and partisan outlets has arisen to monetize polarization and propaganda (Sunstein, 2018; Marwick & Lewis, 2017). However, more work is needed to connect the dots between, say, programmatic advertising systems and the funding of sites like Metric Media. Additionally, data colonialism raises questions of justice and rights – if local populations are being data-mined and manipulated, is this a form of digital colonialism that requires intervention akin to anti-colonial resistance, but in data form?

Pseudo-Local Journalism and the Case of Metric Media

The decline of traditional local news in the United States over the past two decades has created what researchers call "news deserts" – communities with little to no original local journalism (Abernathy, 2018). This void has serious implications: lower civic engagement, less government accountability, and a populace starved of community-specific information (Meyer et al., 2019). Unfortunately, it also creates an opening for what appears to be a perversion of local journalism (McChesney & Pickard, 2011). Pseudo-local journalism refers to outlets that pose as local news providers but are not authentically engaged in journalistic service to those communities. Metric Media and affiliated networks (e.g., Locality Labs, Franklin Archer) are the foremost examples, earning the moniker "pink slime" news sites – a term likening them to processed filler material masquerading as real meat (Bengani, 2019).

Priyanjana Bengani's investigations (2019, 2020, 2021) at the Tow Center for Digital Journalism provide the foundational understanding of Metric Media. In late 2019, Bengani reported on an "intricately linked network of 450 sites" that pretended to be local or business news (Bengani, 2019). By mid-2020, research showed that the network had ballooned to over 1,200 sites across all 50 states (Bengani, 2020). These sites have innocuous-sounding names, often mimicking small-town newspapers (e.g., Springfield Times or County Gazette), and they produce vast volumes of content. The production model is what stands out: approximately 90% of stories are algorithmically generated–taking publicly available datasets (like school rankings, police blotters, and health inspection scores) or repurposing wire service articles–and publishing them with minimal human oversight (Bengani, 2020). This creates a facade of a busy news outlet covering local minutiae. The remaining content (around 10%) carries human bylines and tends to have a

strong political bias, skewing conservative. Examples include raising the alarm about voter fraud using partisan sources, pieces attacking Democratic politicians, or op-eds lauding conservative policies (Bengani, 2020).

Two particularly troubling practices were uncovered: first, some of these operations produced physical print newspapers during election seasons that were essentially campaign flyers but formatted like local papers–and mailed them to households without clearly disclosing their political origin (Bengani, 2019). Second, the content choices often correlate with election cycles and political talking points, suggesting coordination. Bengani (2019) noted a pattern: these outlets would surge activity around elections, focusing on issues advantageous to certain candidates (e.g., crime spikes when a "tough on crime" platform was being pushed). By 2021, Bengani and colleagues, through analysis of IRS filings and other records (Bengani, 2021, p. 12), traced funding and found links to a conservative Roman Catholic political advocacy group and other dark-money entities that bankrolled Metric Media's expansion (Bengani, 2021, p. 17). One group launched a campaign against Joe Biden weeks before the 2020 election, aligning with the type of content observed on the sites (Bengani, 2021, p. 20).

From a theoretical perspective, pseudo-local journalism, like Metric Media, blurs boundaries between traditional categories. It essentially uses the form of journalism (local news format, headlines, site design) as political propaganda delivery. It also leverages algorithmic automation, a trait of digital media, for cost-efficiency and scale. This hybridity means scholars must borrow from multiple frameworks to analyze it: propaganda theory (to assess intent and persuasion tactics), journalism studies (to see how it imitates news and exploits the erosion of journalistic institutions), and information science (to follow the data and algorithm aspects). Methodologically, Bengani's work combined investigative

journalism methods (open-source intelligence (OSINT), FOIA for tax records) with content analysis. Going forward, more computational approaches could map these sites' reach and network connections (e.g., tracking their articles' social media presence or search engine optimization (SEO) tactics).

One gap in existing research is the effect of these pseudo-local sites on audiences. Do residents realize these are not genuine local news outlets? Do the sites manage to change minds or voting behaviour, or are they largely ignored? A recent study by Darr (2024) in The ANNALS of the American Academy of Political and Social Science examined the credibility and readership of such "pink slime" outlets. Initial findings suggest that while the quantity of sites is high, the engagement per site is often low (many have sparse web traffic and few local readers) (Darr, 2024). However, even limited engagement might matter in tight electoral contexts, or the content might be repackaged via campaign messaging. Another study could investigate whether people can distinguish pink slime content from authentic news in blind tests to gauge the potential for deception.

Furthermore, there is the question of platform response: search engines and social media companies could tweak algorithms to downrank sites with pink slime characteristics (e.g., mass-produced content, no genuine local reportage). However, there is a slippery slope concern and technical challenge in doing so accurately. Again, this area needs research–essentially, how to algorithmically detect and mitigate pseudo-news networks without harming authentic local news or engaging in partisan discrimination.

Another gap is exploring analogous phenomena in other countries or contexts. The U.S. case of Metric Media might be the most extensive documented example, but the concept of partisan networks posing as local media could spread (or may already exist in

other democracies). For instance, are similar European networks exploiting local news decline, perhaps tied to different political factions? Comparative research could tell us if this is a broader trend or primarily a U.S. phenomenon catalyzed by the particular decline of local newspapers and the campaign finance environment.

Finally, potential solutions: some scholars propose bolstering genuine local journalism through policy (grants, non-profit models, public funding) to fill the void that pink slime sites exploit (Meyer et al., 2019). If every community had a trustworthy local news source, the market for imposters would shrink. Media literacy efforts specifically targeted at local communities could also help residents vet new "news" sources that pop up. Organizations like the NewsGuard project attempt to rate news websites for reliability; integrating such ratings into browsers or feeds could alert users that, for example, "Springfield Times is a low-trust site with political backing" – akin to a nutrition label.

Gaps in Literature and Future Directions

This review surveyed extensive literature on propaganda, media effects, digital disinformation, and socio-technical structures underpinning contemporary journalism. Several research gaps and open questions emerged, underlining the contribution that further study – including this thesis – can make:

- **Pseudo-Local Propaganda Impact:** While networks like Metric Media are documented, there is a lack of empirical research on their impact on audiences and local democratic processes. Do these sites persuade or mislead local readers, or are they preaching to the choir (or to hardly anyone at all)? The literature has identified the problem but not fully measured its effects. This gap is crucial for understanding

the real-world significance of digital propaganda at the community level (Meyer et al., 2019). Future studies could employ surveys or experiments in affected communities to see if exposure to such sites correlates with misperceptions or changes in political behaviour.

- **Cross-Platform Dynamics of Disinformation:** Much research has been U.S.-centric and focused on single platforms (e.g., Twitter or Facebook studies of misinformation spread). Scholars need a more holistic picture of how disinformation flows across multiple platforms and channels–for instance, how a story might start on a pink slime site, get picked up on Facebook groups, and then be amplified by talk radio or messaging apps. Understanding these trajectories could inform more effective interventions. Additionally, comparative research in non-U.S. contexts (where messaging apps like WhatsApp or Telegram might play a more significant role than open websites) is needed to generalize theories of digital propaganda (Sunstein, 2017).
- **Algorithmic Transparency and Effects:** The literature points to algorithmic curation as a key factor in modern news consumption, but there is a gap in independent data on algorithms. Platforms often guard details of how their news feed or search ranking algorithms work. Researchers like Napoli (2019) express concern that metrics-driven systems can distort journalism, but to engage critically, scholars need more transparency or creative methods (audits, simulations) to study these algorithms. This is a methodological gap and a policy issue (advocating algorithmic transparency or access for researchers). How, for example, does Google Search treat a Metric Media site vs. a legitimate local paper in results? Initial evidence suggests low SEO for pink slime sites, but systematic analysis is lacking.

- **Data Colonialism and Journalism:** Data colonialism concepts are still new and mostly theoretical. Empirical research applying it to journalism or propaganda could yield insights. For instance, is there evidence that communities of colour or marginalized groups are disproportionately targeted in data-driven propaganda campaigns (a form of digital colonial targeting)? How do audiences feel about the trade-off of free news content for data extraction–are they aware of it, and does it matter to them? These sociological questions intersect with media studies and could deepen our understanding of power imbalances in the digital public sphere.
- **Methodological Innovation in Propaganda Research:** Many classic studies relied on content analysis, case studies, or surveys. Newer challenges might benefit from network analysis (mapping how propaganda narratives spread through networks), machine learning (to detect coordinated inauthentic behaviour or clusters of pseudo-news sites), and ethnography (interviewing practitioners on both journalism and propaganda sides). There is a gap in interdisciplinary approaches–combining communication theory with data science techniques–to tackle phenomena like coordinated disinformation campaigns. For example, detecting a network like Metric Media initially required investigative journalism; in the future, algorithms might be trained to flag such networks early by recognizing patterns (e.g., identical site templates, standard Google Analytics codes, etc.).
- **Normative and Ethical Questions:** As a final gap, there is relatively little written on the ethical implications for journalists themselves operating in this contested space. Journalists are now often aware that their work can be co-opted or mimicked by propagandists. How has this affected journalistic practice and identity? Likewise, what are the ethical lines for governments or civil society in countering propaganda? For instance, if the government funded local journalism to combat Metric Media's

influence, would that raise its own independence issues? These questions go beyond empirical description to normative considerations and are an important frontier as societies decide how to respond to digital propaganda.

- **Transitions in scholarship:** The review also reveals how theories are being updated. Propaganda theory incorporates digital platform analysis, agenda-setting examines social media, and surveillance capitalism is supplemented by critiques like data colonialism. This thesis builds on that by explicitly bridging legacy theories of media influence with contemporary digital realities.

The case of Metric Media is a stark reminder of the power of media to shape public opinion and the fragility of our information ecosystems when economic, technological, and political factors align in troubling ways. Yet, it also highlights why scholarship and informed practice matter: by rigorously studying such cases, clearly defining terms, and critiquing methods and theories, the research can better equip journalists, policymakers, and the public to address the challenges.

3. Methodology

This study employs a case study research design centred on Metric Media, integrating computational journalism and data science techniques to investigate the network's content and online dissemination. The methodology is deliberately multi-method, combining large-scale text analysis, network analysis, and visualization. This design enables a rigorous examination of Metric Media's tactics from multiple angles–content, sentiment, thematic structure, and social media engagement–in order to answer the research questions with both depth and breadth. In alignment with critical media theories, the methods revealed subtle patterns of agenda-setting and propaganda in the data while leveraging concepts from surveillance capitalism to understand how digital platforms are utilized. The following sections detail the research design, data acquisition and preparation, analytical techniques, integration of findings, and the study's limitations and ethical considerations.

Research Design

This investigation is structured as an in-depth case study of Metric Media's pseudo-local news network, using a mixed-methods approach grounded in digital content analysis. Adopting a case study design is appropriate because the research focuses on a single complex phenomenon–a coordinated network of news-lookalike websites–and seeks to understand its mechanisms in context (Yin et al., 2018). The research design blends quantitative computational analysis with qualitative interpretation. This allows the study to capture measurable patterns in the content and dissemination (e.g. frequency of

particular topics or sentiment trends) while also interpreting their significance in light of journalism and communication theories.

Multiple Methods Aligned with Research Questions: The methodology was carefully mapped to the study's research questions (RQs) and objectives, ensuring that each method contributes directly to answering specific aspects of these questions:

- RQ1 (Creation and Dissemination): To examine how Metric Media creates, hosts, and disseminates its pseudo-news platforms, the study uses digital data scraping and content mapping. The research captures the structure and output of the network's platforms by collecting a comprehensive dataset of articles (via RSS feeds) and cataloguing associated social media pages. This addresses the mechanisms of content production and distribution (e.g. how widely content is spread across sites and social channels). The case study design enables detailed observation of these processes in the real-world context of a known propaganda network.

- RQ2 (Efficacy of Attack Vectors): To evaluate the effectiveness of Metric Media's attack vectors (such as automated posting and micro-targeted content), the methodology combines text analytics (to identify themes and emotional tone of content) with engagement analysis of social media data. The study gauges how well Metric Media's strategies capture audience attention by analyzing content patterns (e.g., repetitive automated story templates or region-specific topics) and measuring user interactions on platforms like Facebook. This approach reveals whether tactics like high-volume automated posting translate into meaningful engagement or if micro-targeted local stories trigger community response.

- RQ3 (Detection and Mitigation Strategies): The research design is inherently reflexive and action-oriented in exploring practical methods for journalists and the public to detect and counteract toxic messages. It demonstrates the use of computational tools (e.g. clustering, sentiment detection, network mapping) to uncover patterns of coordinated propaganda, effectively modelling a detection framework. By documenting the analytical process and its findings, the study provides a template to help others identify similar networks. For instance, the prominence of specific entities or unusual uniformity in sentiment across supposed local news sites could serve as red flags for media observers. The multi-method approach–combining content analysis, cross-platform data, and triangulation with external knowledge–is a replicable mitigation strategy (enabling verification of suspicious networks through data). In designing this aspect, the research remained reflexive, recognizing that the investigative process is part of the contribution: it shows how propagandists' technologies can be repurposed to expose misinformation (directly addressing RQ4).

- RQ4 (Repurposing Technologies to Expose Misinformation): This question is woven throughout the methodology by illustrating how advanced data science techniques, often leveraged in targeted advertising or propaganda dissemination, can be repurposed for transparency and accountability. The research design employs natural language processing (NLP) tools and machine learning not to spread content but to dissect it. The study shows these technologies in a new light by using methods like network analysis and clustering – similar in spirit to how social media platforms segment audiences or how propagandists might algorithmically generate content. For example, clustering algorithms that could be used to personalize content

delivery are here used to group and reveal the thematic focus of Metric Media's articles. This alignment is conceptually grounded in Zuboff's (2019) ideas of surveillance capitalism: data-driven techniques typically used to influence behaviour can instead be harnessed to investigate and counter that influence. Thus, the research design proves that sophisticated computational approaches can support journalistic inquiry and expose manipulative communication tactics.

Theoretical Alignment: At its core, the research design is informed by foundational theories in media and communication, ensuring that each methodological choice resonates with a conceptual framework:

- Agenda-Setting Theory: The study probes how this network may be shaping the public agenda by analyzing which topics are emphasized across thousands of Metric Media articles (McCombs & Shaw, 1972). The selection of topic modelling/clustering as a method reflects this theoretical lens–it lets the data reveal key issue themes being promoted. This is critical because agenda-setting research suggests that the frequency and prominence of topics in news can influence what audiences perceive as important. The design intentionally captures these frequencies and patterns, enabling the discussion in later chapters on how Metric Media's content choices might subtly influence public perception.

- Propaganda Model: Herman and Chomsky's (1988) propaganda model posits that media controlled by powerful interests will systematically bias content in favour of those interests. The research design addresses this by using content analysis techniques (e.g. sentiment analysis, entity frequency counts) to detect biases or

anomalies in coverage. The emphasis on measuring sentiment and tone is inspired by the idea that effective propaganda often presents itself as balanced or neutral to avoid detection (Jowett & O'Donnell, 2012). By quantifying sentiment and identifying recurring entities, the study can uncover patterns such as consistently neutral tone or selective emphasis on certain figures/topics that align with propaganda strategies. The methodology operationalizes the propaganda model: it looks for evidence of filtering and framing by systematically analyzing what and how Metric Media communicates.

- Surveillance Capitalism: Zuboff's (2019) concept of surveillance capitalism underscores how data collection and targeted messaging are used for influence and profit. This theoretical perspective influenced the design's inclusion of social media and network analysis. Recognizing that Metric Media's strategy might involve leveraging digital platforms and data analytics (either to micro-target content or to monitor engagement), the study collects platform data (Facebook pages, Twitter accounts) to see how the network uses these channels. The decision to map out Metric Media's online presence and engagement patterns acknowledges that any modern propaganda effort is intertwined with platform algorithms and user data flows. Additionally, by reflecting on the platform constraints (e.g. API limitations), the design implicitly speaks to how corporate control of data (a theme of surveillance capitalism) can hinder independent analysis–a point further addressed in the limitations section.

Data Acquisition

Data Sources: The primary data for this study consists of digital textual content and related metadata drawn from Metric Media's network of websites and associated social media pages. Given Metric Media's extensive reach (with over a thousand pseudo-local news sites across the U.S.), a comprehensive data collection strategy was essential. The study leveraged an existing dataset curated by an activist-driven project known as MassMove (MassMove et al., 2024). This group of researchers and civic activists has been tracking Metric Media's operations since 2020, and they made available a repository on GitHub containing information on Metric Media sites and content. Utilizing this source ensured that the data acquisition started from the most up-to-date and wide-ranging list of Metric Media outlets, which would have been difficult to compile independently. Notably, the MassMove dataset provided a list of domains, initial content feeds for Metric Media sites, and records of any known social media accounts (Facebook pages, Twitter handles) linked to each site.

RSS Feed Collection: Most textual data was collected via Really Simple Syndication (RSS) feeds from Metric Media websites. RSS (Really Simple Syndication) feeds are XML-based web feeds that facilitate the distribution of frequently updated digital content, such as news articles and blog posts, in a standardized, machine-readable format. This form of content syndication became integral to early web infrastructure by enabling users to aggregate updates from multiple sources in one place, improving accessibility and information flow (O'Reilly, 2007). As digital communication evolved, RSS played a pivotal role in shaping how information was disseminated online, particularly in the context of the

transition to Web 2.0 (Kilgour, 1998). Each Metric Media site offers an RSS feed of its articles, making it an efficient entry point for data scraping. By tapping into RSS feeds, the study captured article titles, descriptions (or summaries), publication timestamps, and source site identifiers for many posts across the entire network. RSS was a strategic choice for data acquisition because it provides a continuous, real-time stream of content from the sites, allowing the collection of posts as they appeared over time. Moreover, prior research has demonstrated the value of RSS feeds for monitoring news trends and media output at scale (Guo et al., 2016), aligning with the research's goal of observing Metric Media's content patterns on a large scale.

A Python-based web scraping pipeline was implemented to systematically gather the RSS feed data. This pipeline used well-established libraries: feedparser was used to parse RSS XML feeds directly, and the requests library was used to fetch raw HTML content when needed. Using these libraries in tandem allowed for robust data acquisition: feedparser efficiently handled the structured feed data (avoiding the need to decode XML manually). At the same time, requests enabled the retrieval of additional page content in cases where the feed provided only a snippet and a link to the full article. Over several weeks of collection (covering the latter part of 2023 into early 2024), the scraper iterated through the list of known Metric Media site feeds (obtained via MassMove) and aggregated the posts into a unified dataset. The outcome was a comprehensive corpus of RSS feed entries spanning numerous sites and dates, effectively capturing Metric Media's published content. Each entry in this corpus includes the textual description of the article (often the first few paragraphs or summary), the date/time of publication, and the source website's identifier.

Social Media Data Collection: In addition to on-site content, the study gathered data on Metric Media's social media presence to address the dissemination vectors. Based on the MassMove repository information, nearly every Metric Media site had a corresponding Facebook page. However, many had little to no activity, and very few had a Twitter (X) account. Recognizing that social media could be a key channel (or at least a claimed space) for Metric Media to push content to local audiences, the study attempted to collect Facebook post data from these pages. A custom Facebook scraper (an open-source tool by Kevin Zúñiga, cited as GitHub, 2023) was employed to extract public posts, timestamps, and engagement metrics (likes, shares, comments) from each identified Facebook page. This method did not require platform API access, which was crucial given the increasing restrictions on official APIs – for example, Twitter's API became paywalled in 2023, severely limiting researchers' ability to collect Twitter data (Gotfredsen, 2023). Using a scraper that mimics ordinary web access to Facebook's public pages, the study collected whatever content was available up to early 2024.

The Facebook data acquisition resulted in a secondary dataset comprising 7,510 Facebook posts made by Metric Media–affiliated pages between April 27, 2017 and February 16, 2024. For each post, the dataset includes the posting page (which corresponds to a Metric Media outlet), the content (often a link back to an article and a short caption or excerpt), and engagement counts. It became apparent during collection that many pages had posted only a handful of times (some only at the page's creation and then abandoned), while a minority of pages were relatively more active. No Twitter data were directly collected (due to the aforementioned API limitations and the general absence of active Twitter accounts for these sites). However, the existence or non-existence of Twitter handles for each site was noted from the MassMove records for context.

Data Storage and Management: All collected data were stored in structured formats conducive to analysis. The RSS feed data were compiled into a master CSV (Comma-Separated Values) file containing fields for the site domain, publication date, and the text of the feed description (article summary). CSV was chosen for its simplicity and compatibility with analysis tools – it serves as a raw data exchange format that can be easily read into data frames for processing. The decision to use CSV and cloud storage (Google Drive) for intermediate data storage was also influenced by practical considerations of collaboration and backup (Bisong, 2019). The Facebook posts and engagement data were likewise stored in CSV form, with separate columns for each engagement metric. A consistent format allowed these datasets to be later merged or compared on common keys (such as site name or date).

Before proceeding to analysis, initial examinations of the datasets were conducted. This included counting the number of entries (posts) gathered per site, checking for any missing data or anomalies, and verifying that the time ranges covered by the data aligned with expectations (for instance, ensuring that the RSS data spanned the intended period and that Facebook data collection did not miss active pages). This exploratory check confirmed that the dataset was robust: the study had content from hundreds of Metric Media sites, and the Facebook dataset, while sparse in places, captured the key period of interest when Metric Media's operations were known to be active. Having acquired a rich dataset of textual content and social media interactions, the next step was to prepare this data for analysis through cleaning and integration.

Data Preparation

Raw data, especially textual data from web sources, require thorough preprocessing to ensure that subsequent analyses are accurate and meaningful. In this study, data preparation involved cleaning the text of noise and formatting issues, transforming data into suitable structures (e.g. converting dates to standard formats), and integrating multiple data sources into a cohesive analytic dataset. These steps were critical for enhancing data quality and aligning the data with the research objectives, as the insights we can draw are only as reliable as the data we feed into our algorithms (Kitchin, 2014).

Text Cleaning: The RSS feed descriptions collected often contained unwanted artifacts such as HTML tags, special characters, or extraneous whitespace. Additionally, the text included common words ("the", "is", "in", etc.) that carry little semantic weight. Cleaning proceeded in several stages:

- All text was standardized to lowercase to ensure uniformity (so that, for example, "News" and "news" are not treated as separate tokens).
- HTML tags were stripped out using the BeautifulSoup library (Richardson, 2007). Removing HTML ensured that only the human-readable content remained, since tags or HTML entities (e.g.) would otherwise introduce noise.
- Non-alphabetic characters such as numbers and punctuation were removed or replaced as appropriate. A series of regular expression patterns (Friedl, 2002) was applied to eliminate these characters, except where they were part of meaningful context (for instance, a year "2023" might be relevant, so numeric removal was done carefully after considering context).

- A stopword filter was applied to remove common function words that do not contribute to topical meaning. Using established stopword lists (Silge & Robinson, 2017), words like “and”, “the”, “by” were dropped from the text. This step reduces dimensionality in text analysis and highlights more substantive words (e.g. nouns, verbs, proper nouns), which are more indicative of content themes or sentiments.
- Extra whitespace and line breaks were normalized (collapsed to single spaces) to tidy the text.

After these operations, each article description was reduced to a cleaned textual snippet, stored in a new field (e.g. final_description). This preprocessed text was human-readable and analysis-ready, devoid of formatting distractions and trivial words, thus focusing subsequent analysis on meaningful linguistic content.

Data Transformation: Besides text cleaning, the preparation phase transformed specific fields to facilitate analysis. One important transformation was handling the publication date of each post. Dates (initially strings from the RSS feed) were converted into a standardized datetime format and then various temporal features were extracted (McKinney, 2010). For example, each post’s date was parsed into year, month, and day, enabling content aggregation by month or year. This was useful for any time-series analysis, such as checking if certain content spiked around elections or other events. The dates were also used to sort and merge the data with Facebook data by time. Indexing the primary dataset by datetime allowed efficient slicing of the data for different periods, which could be relevant if analyzing trends over the network’s operation timeline.

Another transformation involved focusing the dataset on the columns most relevant to the research questions. The raw RSS data included various fields (some feeds include

author names, categories, etc.), but not all were pertinent. The dataset was pruned to keep key variables: the cleaned text, the site/domain source, the publication timestamp, and eventually analytical results like sentiment score or cluster label. This feature selection ensured that analysis remained targeted and efficient, consistent with recommendations by Blei (2012) on focusing on relevant features in text analysis. Similarly, the key fields retained in the Facebook dataset were the page name (which maps to a Metric Media site), post content or caption, and engagement metrics – extraneous metadata was dropped.

Dataset Integration: With multiple data sources (RSS content data, site metadata, and Facebook engagement data), integration was a crucial preparation step for cross-analysis. Each Metric Media post in the RSS dataset was linked to its origin site (e.g., an entry from "PennsylvaniaChronicle.com" feed is labelled with that site). A separate reference table from the MassMove data provided additional site information, such as the state or region the site purported to serve and whether a given site had known political or organizational affiliations. By merging the RSS content dataset with this site metadata (using the site domain as a key), the study added contextual columns to each post (e.g., a post could now be associated with "State = Pennsylvania" or tagged as part of a subgroup of sites). This contextualization is important for understanding whether certain entities or topics are concentrated in specific regions, reflecting micro-targeting strategies. Merging data in this way follows best practices in computational social science for enriching datasets (Kumar & Shah, 2018), allowing analysis not just of content in isolation but also of content as it relates to source characteristics.

In addition, the results of text analysis procedures (detailed in the next section) were integrated back into the dataset. For example, once Named Entity Recognition was performed on each post (to extract persons, organizations, etc. mentioned), the identified

entities or counts of entities were added as new data columns linked to each post. Likewise, each post received a sentiment score from sentiment analysis, which was attached to that entry. By merging these, the study constructed a rich table where each row (post) has: source site info, cleaned text, list of named entities (or a count/frequency of entities by type), a sentiment label or numeric score, and a timestamp. This unified dataset enabled multifaceted queries – for instance, one could filter posts to see all mentions of a particular politician and check if those mentions were in a positive or negative context (which addresses potential bias in coverage).

Quality Checks: Throughout the preparation phase, the researcher maintained a reflexive stance and careful checking to ensure data integrity. Random samples of cleaned text were inspected to verify that cleaning did not remove meaningful content (e.g., ensuring "U.S." wasn't stripped to just "US" or removed entirely due to the period). Similarly, after merging, spot checks confirmed that the matches between datasets were correct (i.e. the right Facebook page matched the right site). Any anomalies (such as stray entries with null values or mis-parsed dates) were addressed manually or omitted if irreparable. By the end of data preparation, the study had a consolidated master dataset of Metric Media content and related attributes, fully ready for in-depth analysis.

Analytical Techniques

Once the data were cleaned and merged, analytical techniques were applied to extract insights. Each technique was selected to align with the research objectives and to illuminate the data from a particular angle–from identifying who and what the content is about, to how it is written, to the themes that emerge across the corpus, and finally to how the content potentially engages (or fails to engage) audiences on social platforms. The

analytical methods used include Named Entity Recognition (NER), sentiment analysis, topic clustering, and social network and engagement analysis, complemented by data visualization. By combining these, the study addresses both content and context, quantitatively measuring patterns and qualitatively interpreting their significance.

Named Entity Recognition (NER): To understand the focus of Metric Media's content, NER was employed to identify and categorize key entities in the text. NER is a core Natural Language Processing (NLP) technique that tags words or phrases in text as persons, organizations, locations, dates, etc., providing a structured way to catalogue who and what the news is talking about (Nadeau & Sekine, 2007). In this study, NER was crucial for addressing aspects of RQ1 (dissemination content) and RQ3 (detection). The research could reveal patterns, such as which public figures or institutions are most frequently mentioned, by extracting named entities from thousands of posts. This has direct implications for identifying possible bias or agenda. For example, if certain political leaders are referenced far more often or only negatively/positively, that might indicate a targeted campaign. Identifying location entities also helped map the purported local focus of each site's content. A high frequency of local place names would be consistent with Metric Media's strategy to appear as local news, whereas frequent mentions of national organizations or figures might expose a centralized agenda behind the local facade. The spaCy library (Honnibal et al., 2020) and its pre-trained English model were used to perform NER at scale, chosen for its accuracy and efficiency in entity extraction. Every article's cleaned text was processed to pull out entities, which were then counted and analyzed. The results from NER allowed the study to compile lists of the most mentioned people, organizations, and places in the Metric Media corpus. These lists were analyzed relative to the site information: for instance, did certain state-focused sites all frequently

mention the same politician or advocacy group? Such findings would support discussions about coordinated messaging. Moreover, NER output was later combined with sentiment analysis, enabling the study to gauge how those entities were portrayed (e.g., predominantly in neutral or negative contexts). This combination is powerful for detecting propaganda signals–e.g., a pattern where mentions of a particular issue (say, "climate change") are consistently coupled with negative tone could reflect an editorial slant. By structuring unstructured text into entity data, NER thus laid a foundation for quantitative content analysis of Metric Media's narratives.

Sentiment Analysis: While NER answers "who/what is being talked about," sentiment analysis addresses "how is it being talked about." Sentiment analysis involves classifying text along an emotional spectrum–typically positive, negative, or neutral–or assigning a sentiment intensity score (Liu, 2012). This technique was directly applied to each post's text to determine the overall tone of Metric Media's content. The rationale for sentiment analysis ties especially into RQ2 (effectiveness of attack vectors) and the theoretical framework of propaganda. An "attack vector" in propaganda could be an emotional appeal or fear-mongering in content. However, an alternative strategy (as suggested by propaganda theory) is to maintain a neutral or factual tone to build credibility, deploying overt bias sparingly. By measuring sentiment across the corpus, the study can detect whether Metric Media's content skews negative (which might indicate a consistent alarmist or adversarial approach) or remains largely neutral/positive (which might indicate subtlety in persuasion, aligning with propaganda models that stress hidden bias).

For this analysis, a lexicon-based sentiment tool (TextBlob, based on the pattern analyzer) was used to assign each post a sentiment polarity score and categorical label (positive/neutral/negative). The results showed a distribution of sentiment: as reported in

the findings, roughly 62% of posts were neutral, 27% positive, and only 11% negative. This insight–obtainable only through systematic sentiment coding–is central to understanding Metric Media's approach. It suggests that the network predominantly publishes neutral-toned information, which could be interpreted as an attempt to mimic standard local news reporting and avoid overt partisanship. The sentiment analysis thus provided empirical evidence for a key claim: Metric Media's persuasive strategy relies on subtlety (lots of neutral news to build trust, with perhaps occasional strategically placed negative pieces). From a methodological standpoint, sentiment analysis added an emotional dimension to the content analysis, complementing the topic-centric methods.

Additionally, combining sentiment results with the NER data enabled more nuanced analyses, such as checking if certain entities (like "climate change" or a politician's name) tend to appear in negative versus neutral contexts. This cross-analysis enriches the understanding of bias and framing. It is worth noting that the accuracy of automated sentiment tools can vary–idiomatic or sarcastic content can be misclassified. However, for the broad strokes and large volume here, it served as a valuable indicator of tone. The large sample size mitigated limitations in accuracy (patterns still emerge robustly) and are acknowledged in the limitations section.

Topic Modelling and Clustering: The study implemented an unsupervised topic clustering technique to uncover latent themes in the Metric Media content without preconceived notions. Rather than applying a predefined dictionary of topics, a data-driven approach was chosen to let the prominent themes emerge from the text. This was accomplished by transforming the text data into numerical features and then applying clustering algorithms. Specifically, a Term Frequency–Inverse Document Frequency (TF–IDF) vectorization was performed on the cleaned text (Ramos, 2003). TF–IDF converts

each document (here, each post) into a vector of weighted terms, where each term's weight reflects how important that word is in that document relative to the whole corpus. This emphasizes distinctive keywords in each post. Using TF–IDF is appropriate for news content analysis because it down-weights very common words (like "said" or "today") and highlights words that might signal specific topics (like "election" or "church").

With a high-dimensional TF–IDF matrix representing the corpus, a K-Means clustering algorithm was applied to group posts into clusters based on similarity in their word usage (MacQueen, 1967; Jain, 2010). K-Means was selected for its simplicity, scalability, and interpretability in text clustering tasks, as Jain (2010) recommended. The optimal number of clusters (k) was not fixed arbitrarily; instead, an iterative process and validation techniques, including examining cluster cohesion and using the silhouette coefficient method by Rousseeuw (1987), were used to determine a suitable k. Ultimately, the analysis identified five significant clusters (topics) as the most meaningful segmentation of the content, as reported in the Findings. Each cluster represents a set of posts with a common thematic focus derived purely from the text patterns.

These clusters corresponded to distinct themes in the Metric Media content: for example, one cluster was dominated by demographic and census-related terms (indicating a theme of demographic statistics and population changes), while another cluster centred on religious terms (indicating a theme of local religious events and parishes). Additional clusters highlighted topics like crime and civic information (offender registries, voter data), economics (fuel prices and local business), and education (school enrollments and district data). The clustering results directly inform RQ2 (micro-targeted content effectiveness) by revealing what those micro-targeted topics actually are. Metric Media appears to generate content on community-specific interests–demographics, religion, local governance,

etc.–which likely serve as "wedge issues" or points of engagement for targeted audiences. Identifying these five thematic clusters was a key methodological outcome, as it allowed the research to delve deeper into each theme, interpret its significance (with help from literature and external reports), and see how they align with known ideological or strategic goals of the network. Notably, by comparing these data-driven clusters to investigations by journalists (e.g., Bengani's work in Columbia Journalism Review, 2023), the research found a striking correspondence–for instance, the cluster of religious content aligns with reports that Metric Media worked with Catholic organizations to push certain narratives. This demonstrates the value of clustering as a method to unearth meaningful patterns that can then be corroborated with external knowledge, strengthening the validity of the findings.

Social Network and Engagement Analysis: Beyond analyzing the content of Metric Media's articles, the study examined the social media footprint and engagement of the network to gauge how content dissemination efforts fared in practice. This aspect of the analysis responds to RQ2 (effectiveness of attack vectors) and provides context for RQ1's dissemination question by focusing on the channels (Facebook and nominally Twitter) used by Metric Media. Using the earlier Facebook posts dataset, the study performed temporal and user (page) analyses of posting activity and engagement metrics. The posting frequency over time was analyzed by aggregating the number of Facebook posts per month and year across the network. This timeline revealed higher or lower activity periods, which could correspond to external events (e.g., surges during election seasons) or internal changes in Metric Media's strategy. Indeed, the data showed fluctuations and a general increase in engagement in 2022 and 2023, which might reflect ramped-up efforts ahead of elections or improved strategies to draw interactions.

Engagement metrics (likes, shares, comments) were averaged and compared year-over-year to identify trends. The analysis found that while overall engagement was relatively modest (compared to mainstream news pages), there were distinct peaks in specific years, suggesting that a few stories or pages managed to capture attention briefly – possibly due to viral local issues or coordinated promotion. However, one telling finding was that no Metric Media Facebook page had more than 100 posts over the nearly 7-year span. This is an extremely low posting volume for what purports to be a news outlet, confirming that many of these pages were largely inactive placeholders. This insight speaks to the nature of Metric Media's dissemination: the network's strategy on Facebook might have been to create a facade of local news pages without heavily investing in building an organic social audience. The data also allowed for identifying the "most active" pages (those few that did approach 100 posts), which could indicate either test cases where Metric Media tried to generate engagement or specific locales where they focused more effort.

By analyzing engagement per post and page, the study also assessed how effective the content was in prompting interaction. The engagement analysis showed generally low average likes and shares, hinting that Metric Media's Facebook presence did not gain significant traction. This might imply that their primary impact was through the websites themselves (and their appearance in search results or other channels) rather than through widespread social sharing. Such a finding is valuable for RQ2: it suggests that, at least on Facebook, the "attack vector" of pushing content did not translate to high engagement, raising questions about the actual influence of these posts. It also underscores the limitations of reach–a theme in the limitations section–that simply creating content does not guarantee audience impact, especially if the network's social strategy is underdeveloped or if audiences are skeptical.

Lastly, the study considered the network mapping aspect by noting the connectivity (or lack thereof) between Metric Media's sites and social accounts. Since most sites had a Facebook page but little content, and few had Twitter, the "network" in social media terms was sparse. There was no dense web of cross-posting or interlinking accounts–instead, it was a hub-and-spoke model with each site's page existing in isolation. While more descriptive, this observation provided a baseline understanding that Metric Media's influence tactics might rely more on appearances (having many sites and pages as a form of astroturfing local presence) than on genuine network effects or viral spread.

Data Visualization: Throughout the analytical process, visualization was used to interpret and present complex results. For example, once content clusters were formed, geographical mapping was employed to see if specific topics were prevalent in certain regions. Using the Folium library, the study created maps displaying the distribution of the five content clusters across the United States. Each Metric Media site (or groups of sites by state) was visualized with pie charts indicating the proportion of that site's posts belonging to each cluster. This allowed a quick visual assessment of whether, for example, the "religious events" topic was concentrated in particular states or if the "fuel prices" economic topic was nationwide. The maps and charts facilitated triangulation with external knowledge – for instance if a state showed a high proportion of religious content, one might cross-reference if that corresponds to known funding or partnerships (as was the case with some states where Catholic-aligned funding was reported). Visualization was also used for the Facebook data, plotting engagement trends over time and highlighting the top contributors (pages) to give an intuitive sense of scale and variation. All these visual tools served to integrate the analytical techniques by providing a common framework (e.g. the map) to overlay different data aspects (topics and geography).

Integration and Triangulation

A key strength of this study's methodology is integrating multiple methods and data sources to achieve a more profound, cross-validated understanding of Metric Media's operations. Rather than relying on a single analytical lens, the research triangulates findings from text content analysis, thematic clustering, and social media data and even cross-references these with external investigative reports. This multi-pronged approach enhances the credibility and richness of the results, as convergent evidence from different angles reduces the likelihood that observations are artifacts of one method (Jick, 1979).

Content Triangulation: Within the content analysis, the study cross-compared the outputs of NER, sentiment, and clustering to paint a cohesive picture. For instance, consider the theme of religious content that emerged as one of the clusters. Triangulation was achieved by examining the named entities in those cluster posts (did they frequently mention church-related entities?), the sentiment of those posts (were they neutral, positive celebratory, or perhaps negative about controversies?), and the context from literature. Indeed, the cluster analysis flagged a prominent "parish/mass" theme, NER confirmed frequent mentions of church-related proper nouns, and sentiment analysis showed these posts were usually neutral or positive in tone (appropriate for community religious news). When these internal findings are consistent, it strengthens the interpretation that Metric Media deliberately includes benign community religious news – potentially to build goodwill or target faith-based communities. Similarly, for the cluster about "offender registries and voter registration" (a more politically charged theme), integration meant looking at entity data (e.g., mentions of crime-related terms or officials) and sentiment

(possibly more negative or fear-inducing) to conclude that this theme might serve a law-and-order agenda appealing to particular political sentiments. By correlating entity frequencies with cluster membership, the study checked whether specific names (like particular politicians or organizations) were largely confined to specific thematic clusters, which could indicate coordinated messaging (Conover et al., 2011). These internal consistency checks ensured that each method's results were not considered in isolation but as part of an interlocking evidence set.

Cross-Source Triangulation: The methodology also triangulated between the content and social media data. Although one is textual and the other behavioural, their integration addresses how content supply meets audience interaction. For example, one finding was that Metric Media sites heavily emphasize local demographic and economic stories. The research then examined Facebook engagement around those types of stories: did posts about local statistics or gas prices garner any likes or shares? The general observation was that engagement remained low across the board, which suggests that even content tailored to local issues did not organically catch fire on social media. This cross-source comparison tempers the interpretation of Metric Media's effect. While the content analysis shows a sophisticated alignment with local interests (agenda-setting through local proxies), the engagement analysis indicates a limited realization of that influence on Facebook. Such a nuance is only apparent when triangulating content themes with engagement outcomes.

Another cross-source integration involved time: the study compared the volume of content produced (from RSS data) with the volume of Facebook posts over the same period. If Metric Media ramped up content output in a certain year, did the Facebook activity also rise? The integrated timeline suggested some synchronization (in 2022–2023, both content

and engagement saw upticks), hinting that the network perhaps increased efforts on all fronts during a politically significant time. Conversely, the near absence of Twitter activity, corroborated by both the content data (no evidence of Twitter-specific content like very short news bits or hashtags in text) and the MassMove list, triangulates that Twitter was not a utilized vector–an interesting strategic note, possibly due to the network's target demographic being more active on Facebook than Twitter, or due to the difficulties of automating content on Twitter in recent years.

Theory Triangulation: The study also performed an implicit form of triangulation by interpreting results through multiple theoretical frameworks (as seen in the Discussion chapter). Methodologically, this means that when analyzing the data, the researcher considered which theoretical explanation best fits the pattern and sought evidence for that. For example, the prevalence of neutral sentiment was viewed through a propaganda model lens (as subtle indoctrination), but one could also question if it is simply standard journalism practice. To reinforce the propaganda interpretation, the methodology combined neutral sentiment findings with the fact that certain topics aligned with partisan talking points (agenda-setting) and the knowledge of who funds Metric Media (surveillance capitalism context of targeted messaging). It is in the methodology stage that the groundwork for such theory triangulation was laid: the data collection and analysis were broad enough to allow testing different explanations. If, hypothetically, the data had shown a very different pattern (like mostly negative, fear-based content), the interpretation might lean more on direct propaganda via a fear appeals framework. By not tailoring the method to only one expected outcome, the study ensured that it could capture whatever the data showed and then relate it to theory. In practice, the findings aligned with the subtle

propaganda hypothesis, which was confirmed by seeing multiple pieces (topics, tone, external evidence of funding) line up.

External Validation: Triangulation was further enhanced by comparing the study's data-driven findings with external investigative reports and literature. Throughout the methodology, the research remained aware of known information about Metric Media from journalism sources (e.g., articles by Bengani (2019, 2020) and others). After performing, for example, the clustering and seeing a religious cluster, the research recalled that Columbia Journalism Review (Bengani, 2024) reported on Catholic organizations partnering with Metric Media. This prompted a check: indeed, the data showed exactly that kind of content presence. Similarly, knowledge of Metric Media's founder (Brian Timpone) and affiliated companies informed the analysis of the network's structure–while not directly a part of the computational method, this reflexive incorporation of context ensured that the methodology's results were interpreted correctly and not overextended (Bartholomew, 2022). This research considers this a form of data triangulation with expert/secondary sources. It adds credibility: when an algorithm finds pattern X and a journalist independently documented X, we gain confidence in that result. Conversely, if the data had shown something that contradicted all prior reports, it would have signalled the need for caution or further investigation (none did in this case).

Reflexivity in Integration: The research approached the integration of methods with a reflexive understanding of the researcher's own perspective. Because this project inherently challenges a network accused of propaganda, it was important that findings were integrated objectively and not skewed by an intent to demonize Metric Media. The question was asked at each step: "Is there an observable pattern because it is truly in the data or because there was an expectation of wrongdoing?" By letting the data speak first

(through algorithms like clustering) and then cross-verifying with theory and external reports, the methodology guarded against confirmation bias. This reflexivity ensured that, for example, if the data had primarily shown innocuous local news with no partisan bent, the research would report that, even if it defied initial expectations. In effect, the triangulation process served as a check on biases – only when multiple sources of evidence pointed the same way did the study draw firm conclusions.

Limitations and Ethical Considerations

No research methodology is without limitations, and it is important to acknowledge the constraints and ethical considerations that come with a digital investigation of this nature. This study faced challenges related to data completeness, methodological scope, and the evolving landscape of platform policies. Using data curated by activist groups and employing web scraping techniques raised specific ethical questions. Being transparent about these issues contextualizes the findings and demonstrates the rigour with which the research approached potential pitfalls.

Data Scope and Source Limitations: A primary limitation lies in the scope of the dataset and its sources. The study relied heavily on the MassMove activist-curated dataset to identify Metric Media sites and gather RSS feeds. While this source was invaluable (and likely the most comprehensive list available), it may not capture every single outlet or piece of content associated with Metric Media. Additional sites or content streams that are unknown to the activists might exist. If Metric Media launched new sites that were not tracked by the repository, those would be absent from the analysis, potentially biasing results toward the older or more prominent sites. Moreover, as an activist-curated source, MassMove's data might reflect certain biases–for instance, activists might focus on

egregious examples of misinformation in the network, meaning more mundane content could be underrepresented. This risk was mitigated by treating the dataset as a broad starting point and not assuming it to be perfectly exhaustive or balanced; findings were cross-checked with other sources (like news coverage) to ensure no glaring omissions. Nonetheless, the completeness of the data cannot be guaranteed, which is a common issue in studies of covert networks. The alternative–trying to discover sites independently – was not feasible and likely less complete than using MassMove. Thus, dependency on the activist data was required, with caution in generalizing beyond what was captured.

Platform Access and Data Restrictions: The dynamic nature of social media platforms presented limitations, particularly regarding Twitter (now X) data. As noted, Twitter's shift to a paid API in 2023 (Gotfredsen, 2023) essentially cut off the researcher's ability to programmatically collect tweet data related to Metric Media. This means our analysis of dissemination vectors is confined mainly to Facebook (and the content sites themselves). The research would have missed it if Metric Media had significant activity on Twitter (or other platforms like Instagram or YouTube). However, initial exploration suggested they had minimal Twitter presence, which softens this limitation. Still, the inability to incorporate Twitter data highlights a broader issue: platform limitations can shape research outcomes. The study's focus on Facebook engagement was as much a product of research interest as it was of the practical reality that Facebook data were accessible (via scraping), whereas Twitter's were not. This could skew the understanding of "attack vectors"–if, hypothetically, Metric Media engaged in email newsletters or targeted ads (channels not captured), those strategies would be absent from the analysis. Future research could address this by exploring data from email lists or using qualitative

interviews, but for this study, the limitation is that the offered view of dissemination is incomplete, skewed toward platforms that allow researcher access.

Temporal Limitations: The dataset spans up to early 2024, and content was gathered mainly from 2020 onward (when the activist tracking began). Metric Media's operations could evolve, and the methodology provides a historical snapshot rather than continuous monitoring beyond the collection period. Any changes in strategy after the data collection (e.g. new content directions in late 2024 or 2025) would not be reflected. Thus, the conclusions are most applicable to the period studied and may require updating for future developments. Additionally, some of the older Facebook pages had posts back to 2017, introducing a long timeframe. While this is a strength in observing evolution, it also means that earlier content (2017–2018) might not be directly comparable with later content if the network's strategy shifted. The methodology treated the dataset mainly as a whole (with some time-based analyses). However, the temporal drifting of themes is a possibility that could be explored further with time-segmented analysis – which was outside the research's scope.

Analytical Method Limitations: Each analytical technique has its limitations. For NER, automated tools can make mistakes–for example, misidentifying an entity or failing to recognize a less common proper noun. If Metric Media used unconventional names or had optical character recognition (OCR) errors in text, entities could be missed. The research aimed to minimize errors using a reputable NER library (spaCy) and manually verifying a sample of extracted entities. However, some errors undoubtedly remain (e.g., classifying an entity with the wrong type or splitting one entity into two). These errors are generally rare and unlikely to skew aggregate patterns significantly. As mentioned, using a tool like TextBlob comes with accuracy limitations for sentiment analysis. The nuanced

context of news–where a neutral tone can mask a biased implication–is not something an elemental sentiment analyzer can detect. The tool might label a statement as neutral even if it is subtly sarcastic or misleading. We addressed this by interpreting sentiment results as broad trends (e.g., a majority neutral tone) rather than focusing on the sentiment of any single article. In other words, the research does not over-claim precision for sentiment on a case-by-case basis; it instead looks at the forest rather than the trees. Topic clustering via K-Means also has limitations: the algorithm will force every post into a cluster, even if some posts are about niche topics not shared by others. The findings might have missed minor themes because the five dominant clusters drowned them out. Also, cluster interpretation required human judgment–labelling clusters as "demographic" or "religious" themes was done by the researcher reviewing top terms. There is a subjective element present, as another interpreter might label them slightly differently. The research mitigated subjectivity by correlating clusters with clear keyword sets and external references (e.g., recognizing the religious cluster via words like "Parish" and "Mass" is relatively unambiguous). Still, cluster analysis in text is an exploratory tool, and its results should be seen as suggestive categories rather than exact topics in a strict sense (Blei et al. discuss how topic definitions can be fluid (Blei et al., 2003)).

Generalizability: As a case study focusing on one network, the findings and the efficacy of methods have limited generalizability to other contexts. Metric Media is a particular phenomenon with known conservative backing and a specific operational model. The methodology used – heavy on computational text analysis – was well-suited to the large scale of data from this network. This methodology would need adaptation for smaller-scale propaganda efforts or those that operate differently (e.g., purely via memes or videos). Therefore, while the approach demonstrates how to tackle one significant instance of digital

propaganda, it may not directly apply to all forms of media manipulation without adjustments.

Ethical Considerations – Data Use and Privacy: All data used in this study were obtained from public sources: publicly available RSS feeds and publicly accessible Facebook pages. No private or personal user data were collected. This ensures compliance with general ethical standards for research that if information is publicly published, analyzing it is generally permissible (Townsend & Wallace, 2017). However, there are still ethical considerations. Web scraping can violate the terms of service of websites or platforms, even if the data scraped are public. In this case, scraping RSS feeds is usually acceptable as they are intended for automated aggregation. For Facebook, the scraper operates without an API key by loading pages like a regular user; while this likely violates Facebook's terms on automated access, it is a common practice in research when official APIs are restrictive (as they have become). The research design weighed the ethical implications: the data sought were posts made by Metric Media itself (not private user content), so the privacy risk was minimal. The potential harm was more toward the researcher (e.g., legal risks or IP blocks) than toward any individual subjects. Given the public-interest nature of the research–exposing how a network spreads information–the use of scraping was deemed ethically justifiable, aligning with a researcher's duty to uncover truth in service of the public good, provided no personal data was harvested. Nonetheless, the study took care not to scrape beyond what was necessary, and it did not attempt to access any data that required login or was not openly visible.

Ethical Use of Activist Data: Using the MassMove dataset raises questions of consent and representation. The activists compiled data with the intention of public use to hold Metric Media accountable, which aligns with our research purpose. This study has

credited their work (MassMove et al., 2024) and, by using their data, effectively amplify their effort through academic analysis. However, the research remained critical and did not take their interpretations for granted–our methodology re-analyzed everything from scratch to maintain objectivity. One ethical aspect is that activists may have stronger language or explicitly normative stances, whereas academic research strives to be impartial in analysis. By focusing on the data itself and letting the results speak, the research avoided simply echoing an activist narrative and instead provided an independent examination. In writing up results, care was taken to frame conclusions in scholarly terms backed by evidence, even if they align with what activists have alleged.

Reflexivity: Finally, the researcher's own position deserves mention. Engaging in this kind of research, one might have an inherent bias against a network like Metric Media (given that it has been portrayed negatively in media reports). The ethical approach employed was to remain reflexively aware of this potential bias and counter it by strict reliance on data-driven methods. Every claim made is supported by data (and cited accordingly), and methods were chosen that reduce subjective interference (e.g., algorithmic clustering instead of manual content coding, which could be more easily influenced by bias). The researcher kept an audit trail of decisions (why choose X algorithm, why exclude Y data point) to ensure that each choice could be justified scientifically rather than ideologically. This reflexivity is part of ethical research conduct, recognizing that researchers are not neutral robots but must constantly check that their interpretations flow from the evidence, not from preconceived notions.

Conclusion

In designing and executing this methodology, the study combined rigorous computational techniques with a deep awareness of journalistic theory and ethical research practice. The multi-section structure – covering research design, data acquisition, preparation, analysis, integration, and limitations – demonstrates a systematic approach where each component builds toward a comprehensive investigation of Metric Media.

Crucially, this methodological approach proved capable of uncovering subtle patterns in the data that casual observation would likely miss. The integration of content analysis with engagement data and theory exemplifies methodological triangulation, enhancing confidence in the findings. For instance, the convergence of evidence (neutral tone prevalence, thematic clusters around certain civic issues, and low social engagement) consistently pointed to a strategic but perhaps limited-impact propaganda model – a conclusion that is robust precisely because it emerged from multiple analytical angles.

4. Findings

Using the dataset of online posts, data was then processed using machine learning that categorized text sentiments into three primary categories: positive, negative, and neutral. The distribution of these sentiments provides insights into the overall tone and emotional content of the online communication within the dataset. This analysis revealed a notable distribution among the three sentiment categories. Most posts were found to have a neutral sentiment, accounting for about 62 percent of the dataset. This suggests that most of the content was factual or informational, lacking a solid emotional tone that would categorize them as positive or negative.

The positive posts comprised about 27 percent of the dataset, indicating a significant presence of content with a favourable or optimistic tone. This subset of data highlights positive news, uplifting stories, or positively framed information within the online posts. Conversely, negative sentiments were the least prevalent, constituting about eleven percent of the dataset. These posts likely include critical news, adverse events, or content eliciting readers' negative emotions.

The distribution of sentiments within the dataset underscores the complex nature of online communication. The prevalence of neutral sentiment indicates a tendency towards disseminating information intended to inform rather than evoke an emotional response. This is a common characteristic of news outlets and informational websites, which aim to provide readers with facts and data. This may indicate Metric Media's long-term strategy: to build trust, just as local journals would with their intended audience while reserving the possibility of weaponizing information for clients to change the trajectory of public opinion in their favour.

While less prevalent than neutral, the positive sentiment suggests that a significant portion of the content is dedicated to sharing good news, achievements, or other favourable content. This reflects an intentional effort by Metric Media to balance informational content with positive stories that may engage readers emotionally and foster a sense of trust. The relatively low proportion of negative sentiment highlights that while adverse events and stories are undoubtedly present, they do not dominate the online discourse in this dataset. This perhaps reflects editorial choices by Metric Media to minimize the focus on harmful content, instead reserving those moments for high impact.

While the study provides valuable insights, limitations emerged early in the analysis. The accuracy of sentiment analysis tools, including TextBlob, can vary depending on the complexity of the language and the context of the content, potentially affecting the reliability of sentiment categorization. The dataset's composition, including the selection of domains and topics, can influence the sentiment distribution, raising questions about the generalizability of the findings. Future studies could explore sentiment analysis across a broader range of online platforms and content types to better understand online sentiment dynamics. Additionally, employing advanced machine learning models and considering contextual factors could enhance the accuracy and depth of sentiment analysis.

Topic cluster modelling was used to scrutinize the data set, with five key themes emerging. The study of tokenized text data revealed the first distinct cluster of topics centred around demographic insights, characterized by frequent mentions of age, residence, and statistical references. This theme underscores the prominence of demographic discussions within the dataset, highlighting interest or concern by Metric Media's publishers in the changing dynamics of population characteristics. For example, the frequent mention of age and median age suggests a focus on the age distribution within the

population. This could indicate discussions about generational shifts, aging populations, or youth demographics, which are crucial for understanding societal changes and planning for future needs. The emphasis on residence suggests an interest in geographic distribution and mobility patterns. This could relate to urbanization trends, migration, or changes in residential preferences, offering insights into how populations evolve.

Understanding demographic trends is vital for policymakers and urban planners. Insights into age distributions can inform health care, education, and social services planning, while knowledge of residential patterns can guide infrastructure development and housing policies. Demographic insights can be used to tailor products and services to meet the needs of specific age groups or geographic areas. This segmentation strategy by Metric Media could enhance targeting and increase the effectiveness of marketing/outreach campaigns for the site's content, whether that is lawmakers and other influential individuals and groups, the general public, or both.

The second topic identified in the topic cluster analysis revolves around religious events, specifically mentioning parishes, masses, and coverage areas. This focus indicates that a significant portion of the dataset's discourse is dedicated to spiritual activities, suggesting a solid communal interest in spirituality and religious gatherings. Specific keywords and phrases included: "Parish," which implies a geographical and community-oriented approach to discussing religious activities; **"**Mass," which directly points to the Catholic tradition, focusing on specific religious rites and gatherings; and "Coverage Area," wherein the discussion extends across different regions, indicating the widespread importance of religious events across a broader community.

The emphasis on parishes and masses highlights the role of religious events in fostering community bonding and social cohesion. These gatherings are not only spiritual or religious but also serve as important social events that bring the community together. The mention of coverage areas suggests an interest in ensuring religious services are accessible to a broad audience. This could indicate efforts by religious organizations to reach out to their congregants, possibly through online services or by establishing new places of worship in underserved areas. The form of content and the outreach this provides would ideally serve as a data harvesting operation, providing analysis that would provide insights for religious organizations into how they can better engage with their communities by identifying areas with a high demand for religious services or tailoring their outreach efforts to meet the specific needs of different demographics. This is in keeping with Bengani's previous research into Metric Media, which uncovered connections to various funding sources, including entities with ties to the Roman Catholic Church. For instance, the collaboration between advocacy groups like CatholicVote and Metric Media on local 'community news' platforms showcases a concerted effort to disseminate stories and advertisements that align with the church's values, particularly on issues like school choice and abortion (Bengani, 2024).

Metric Media's operations have been linked to key figures like Brian Timpone, who has a history of creating low-cost automated journalism. This approach to news production raises questions about the authenticity and bias of content distributed across these platforms. The network's structure involves several entities, including Newsinator and Local Government Information Services (LGIS), which have engaged in activities blurring the lines between journalism and political advocacy. (Bengani, 2024) Timpone and LGIS's involvement in distributing conservative-aligned content through news-like platforms

reflects a broader strategy of influencing public opinion under the guise of local news (Bengani, 2024).

The Catholic Church's indirect involvement becomes more precise when examining funding flows. Organizations such as American Independent Media have provided financial support to the Metric Media Foundation, indicating a shared alignment of principles (Bengani, 2021). This financial backing is part of a larger ecosystem of funding that supports conservative media ventures, including those connected to the Metric Media network. The strategic placement of content that mirrors the Church's positions on various policy issues, alongside utilizing local news formats to reach wider audiences, underscores the intricate relationships between political advocacy, funding, and media influence (Bengani, 2022).

Moving from religious topics, the third emerging theme centred around legal and regulatory matters, explicitly focusing on offender registries, voter registration, and data privacy or ethnicity comparisons. Most discourse in this category concerns governance, legal frameworks, and civic responsibilities. Keywords and phrases included: Offender**”**, which implies discussions around criminal justice and public safety; “Sex Live Voter Registr” [sic], which suggests topics related to voter registration processes and potentially the inclusion of or concerns about individuals with criminal backgrounds in civic activities**;** and **“**Direct Channel Data Ethnic Compar**”** [sic], which points to discussions on data privacy, ethnicity, and possibly the analysis or comparison of demographic data in legal or civic contexts.

The focus on offenders, particularly in the context of sex offender registries, highlights public concern for safety and the mechanisms in place for community protection. It reflects the societal efforts to balance rehabilitation and public awareness. The mention

of voter registration in proximity to discussions about offenders suggests an underlying discourse on the rights and responsibilities of citizens, including those with criminal records. This may indicate debates on inclusivity and the conditions under which civic rights are extended or restricted, but it also forms a key concern for law-and-order conservatives. These posts could inform and influence policymakers and legal experts in developing or refining laws that address public safety, data privacy, and civic rights. Addressing the fears highlighted through these discussions can help build public confidence in legal institutions and encourage broader compliance with legal standards.

The fourth topic focused on economic discussions, particularly regarding fuel prices, gas stations, and various fuel types like diesel, premium, and mid-grade. This topic reflects the public's interest in economic trends, specifically energy consumption and costs. In this case, keywords and phrases included: “City, Lowest, Station, Price,” indicating a localized discussion around fuel prices, possibly comparing prices across different regions or stations**,** and “Regular, Diesel, Premium, Midgrad” [sic] referring to the available fuel types, highlighting consumer choices and preferences in fuel economy and budgeting. It also perhaps speaks to American consumption patterns, with the choice of fossil fuels over electric addressing more conservative voters.

The focus on fuel prices reflects broader concerns about the cost of living and household budgeting. Fluctuations in fuel prices can significantly impact individuals' disposable income and daily expenses, making this a topic of widespread interest. The mention of cities and the search for the lowest prices suggest that consumers are not only sensitive to changes in fuel prices but also actively seek out the best deals, indicative of regional variations in the pricing and the economic strategies consumers employ to manage expenses. The discussion around different fuel types (e.g., diesel, premium) points to

consumer preferences and the trade-offs between cost, vehicle performance, and environmental considerations.

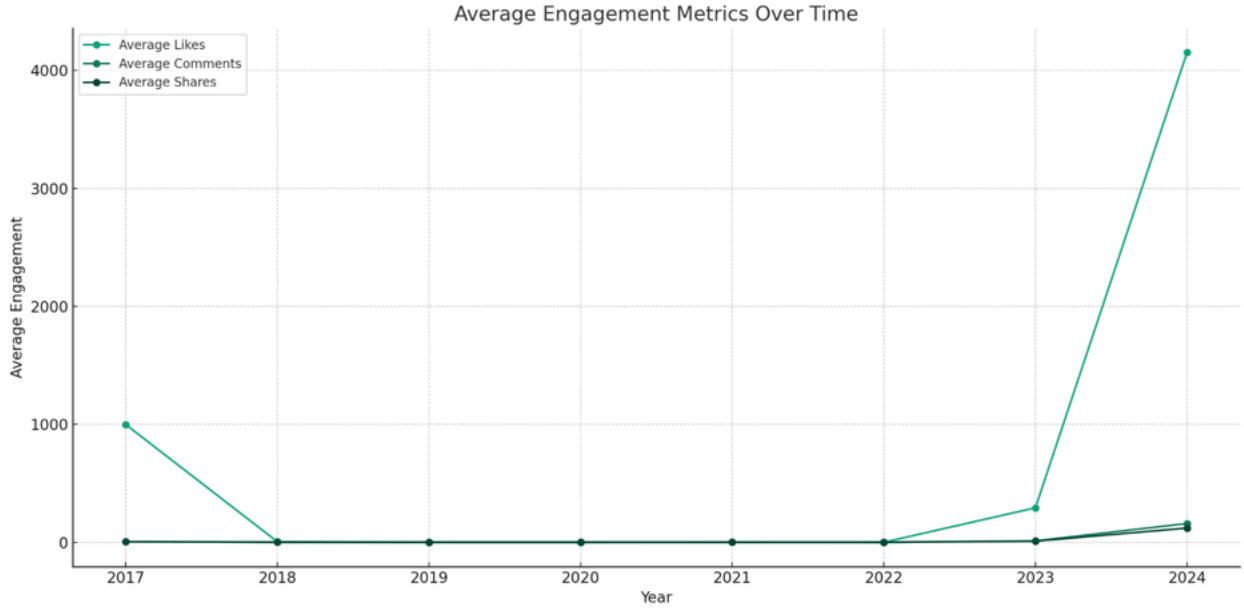
The fifth topic focuses on education, highlighting student enrollment, school demographics, and district planning. It encompasses discussions on school identity, licensure, and the demographic breakdown of educational bodies. This topic underscores the importance of education in public discourse, reflecting concerns, trends, and policies affecting the academic landscape. Keywords and phrases included: "Student Enrollment," indicating discussions around the number of students attending schools, possibly touching on trends, growth, or decline in student populations**; "**School Accord," suggesting an interest in agreements or policies within educational institutions or between schools and other entities**;** and "Identifi District Peopl" [sic], which may refer to identifying specific educational districts and their populations, focusing on demographic data and community makeup.

The emphasis on student enrollment and school accords suggests a keen interest in how educational trends shift, whether due to policy changes, population movements, or other factors. Understanding these trends is crucial for academic planning and resource allocation. The mention of districts and demographics indicates how educational experiences and needs vary across different communities. This can influence policies on educational equity, district rezoning, and targeted educational programs. The reference to licensure and racial demographics hints at discussions around the qualifications of educators and the importance of diversity among teaching staff, an aspect vital for creating inclusive educational environments that reflect the diversity of the student population.

All these topics' insights are likely related to Metric Media's funding sources. The network's conservative funding sources include the Liberty Principles PAC, which paid Newsinator for advertising, and American Independent Media, which provided grants to the Metric Media Foundation due to shared principles (Begani, 2024). Moreover, Think Freely Media and the Franklin News Foundation, tied to John Tillman, have paid entities within the Metric Media network for services such as newswire placement and advertising (Begani, 2021). For example, in 2019, DonorsTrust was the most significant individual contributor to the Metric Media Foundation, emphasizing the foundation's connection to conservative financial networks (Begani, 2024). DonorsTrust is known for funding various conservative causes without disclosing the identities of its contributors, thus acting as a channel for anonymous donations to conservative-aligned initiatives (Bengani, 2021). The network has also been linked to Timothy Dunn, a West Texas oil and gas magnate and executive at Pipeline Media, who has supported the Defend Texas Liberty PAC (Begani, 2024). This PAC, among others, has used the Metric Media network to promote candidates and policies aligned with conservative interests, including direct advocacy for policies favoured by the Texas Public Policy Foundation and attacks on renewable energy policies (Bengani, 2021).

The sentiment analysis and topic review followed the exploration of engagement via social media. X/Twitter yielded next to no presence online. The number of platforms operated by Metric Media with an X presence was firmly in the single digits. This was the case before Elon Musk purchased Twitter, which significantly reduced data scraping efforts. Considering this impediment, the analysis focused exclusively on Metric Media's efforts to engage audiences via Meta's Facebook platform.

The dataset comprises 7,510 Facebook posts from April 27, 2017, to February 16, 2024. Engagement metrics—likes, comments, and shares—offer insights into how users interact with the content. The average likes per post are approximately 193, with a maximum reaching 250,369, indicating significant variance in content reception. Comments average around 7.6 per post, with a single post garnering up to 10,461 comments, showcasing the potential for extensive user interaction. Shares have an average of nearly 6, with the most shared post reaching 7,606 shares, highlighting the capacity for posts to be disseminated widely across the network.



The analysis of engagement trends over time reveals a fluctuating pattern in user interaction with Facebook content. By examining the average likes, comments, and shares per year, some distinct peaks and troughs suggest varying levels of user engagement across different periods. Notably, 2022 and 2023 exhibit a noticeable increase in all forms of engagement, pointing to a heightened interest or activity on the platform during these times. These fluctuations may be attributed to a range of factors, including global events, changes in Facebook's algorithm, or shifts in user behaviour. The observed volatility may

indicate the dynamic nature of social media engagement. Peaks in engagement metrics may correlate with viral content, significant news events, or influential social media campaigns, while troughs could reflect periods of reduced activity or interest.

High-engagement posts are characterized by significantly higher likes, comments, and shares than the dataset average. Examining these posts reveals vital insights into the content that resonates most with users. High-engagement posts are defined as surpassing the 90th percentile in at least one of the engagement metrics. This threshold identifies posts that have achieved exceptional levels of user interaction. The analysis uncovers that many high-engagement posts originated from the "Humans of New York" page, known for its compelling storytelling and human-interest content. “Humans of New York” is not affiliated with Metric Media.

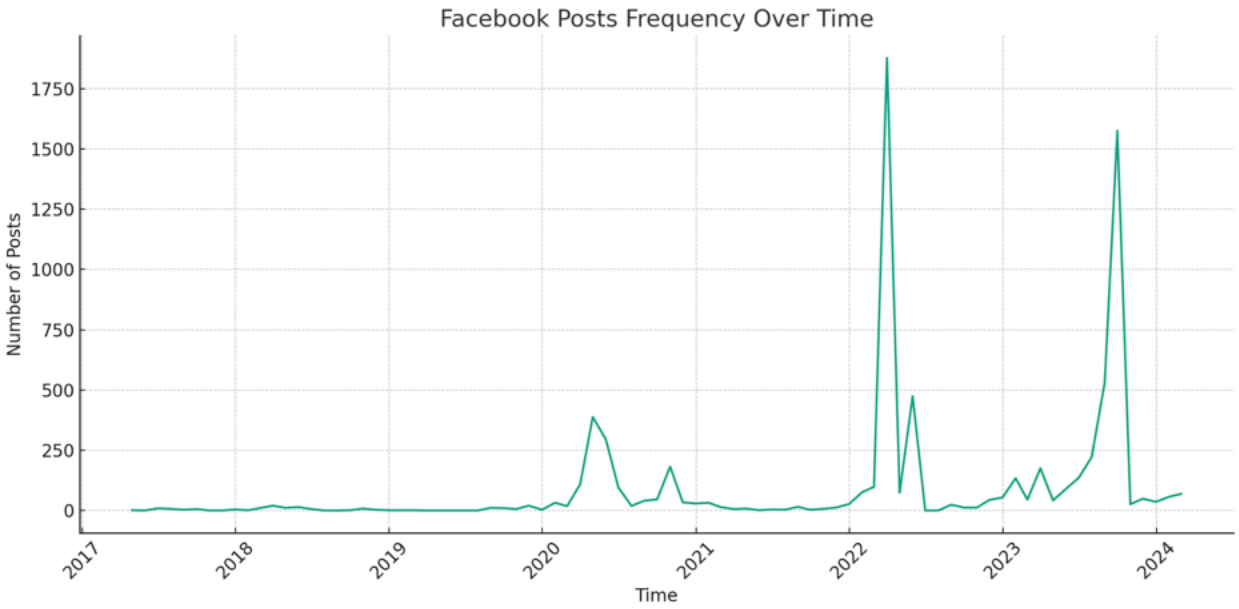
For example, one post on January 5th, 2024, at the time of scraping, had over 84,919 likes, 2547 comments, and 1293 shares (Humans of New York, 2022). It featured three photos of a father recounting an emotional story of seeing an old video of his daughter (Human of New York, 2022). He retold sentimental anecdotes of her growing up in the text, providing a feel-good post for inspiration (Humans of New York, 2022).

The analysis of activity levels is based on the total number of posts made by each user or page. This metric indicates content output and the user or page's commitment to engaging with their audience on the platform. Activity level is quantified by counting the number of posts attributed to each user or page within the dataset. This approach provides a straightforward measure of who is most prolific in generating content.

Metric Media’s properties included The Silver State Times, the most active entity, contributing 60 posts to the dataset. Similar properties in The Dakota Times and East

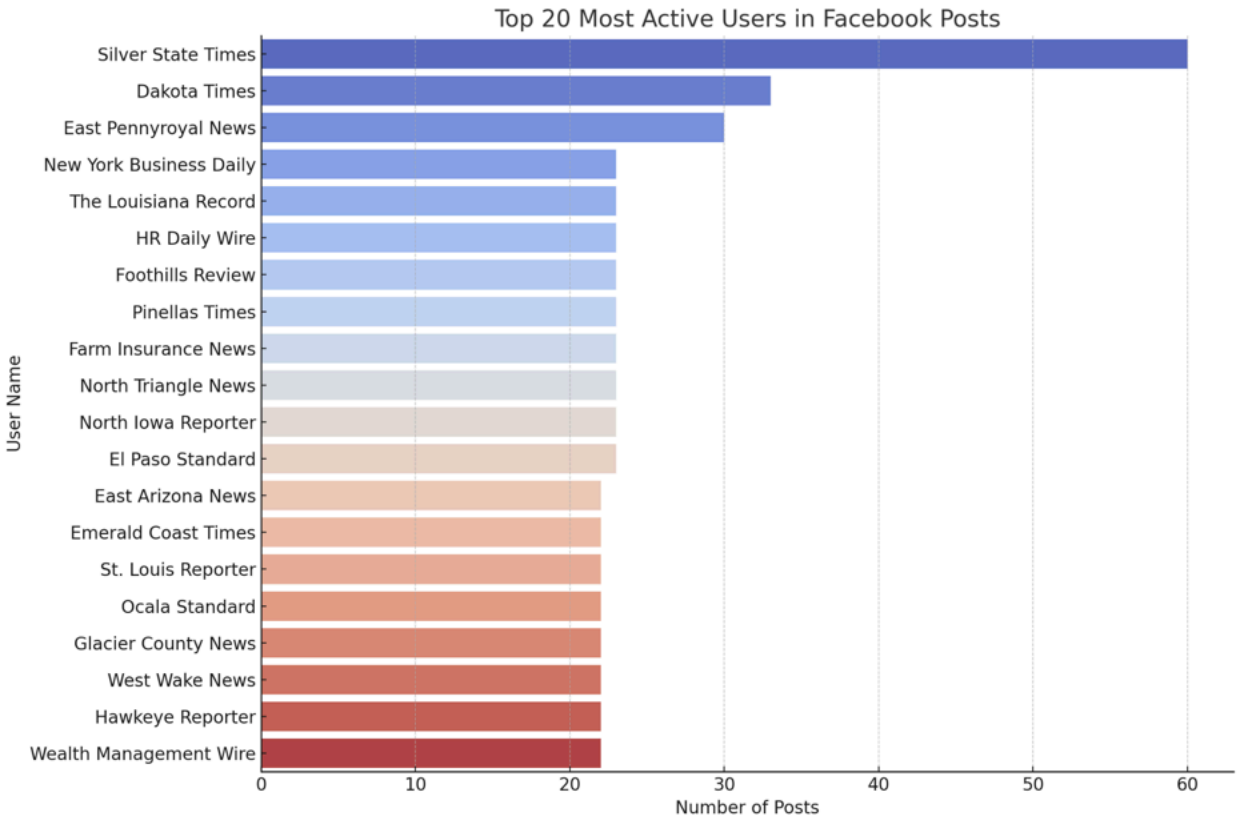
Pennyroyal News followed closely, with 33 and 30 posts highlighting their roles as consistent content creators within their niches. Other notable mentions include The Foothills Review, North Triangle News, and The Louisiana Record, each demonstrating significant activity and maintaining a steady presence on the platform.

Influence is measured by the average engagement (likes, comments, shares) a user or page receives per post. This metric reflects the ability to produce not only content but also content that resonates with and engages the audience. As stated above, "Humans of New York" stands out for its unparalleled influence, with posts averaging 67,365 likes, 2,566 comments, and 2,054 shares. This exceptional level of engagement highlights the profound impact of storytelling and human connection in digital spaces. Although not as well known, Maghreb News Wire demonstrates significant influence, averaging 3,324 likes per post. Other influential entities include Northern California Record and SE La Times, each showcasing the ability to engage audiences through local news and storytelling. The thematic focus, such as human-interest stories by "Humans of New York," perhaps aligns closely with audience interests, driving higher engagement. Regular activity and content relevance, as seen with Maghreb News Wire, cater to niche audiences effectively. Engagement strategies amplify influence, including compelling visuals, narratives, and interactive content.



The "Facebook Posts Frequency Over Time" graph illustrates the monthly posts made across the dataset's timeframe. This visualization allows us to observe trends in posting behaviour over time, potentially highlighting periods of increased or decreased activity.

Key insights drawn from this graph include seasonal or Event-Driven Peaks, with specific months showing higher posting activity, which could correlate with significant events, holidays, or seasonal topics of interest. These peaks reflect the public's heightened interest in sharing or seeking information during these times. Another aspect is Trends Over Time, where overall trends in the volume of posts are plotted, either increasing or decreasing over the dataset's timeframe. This could indicate the growing popularity of Metric Media's content, changes in Facebook's usage, or shifts in the user base's engagement with the platform.



The "Top 20 Most Active Users in Facebook Posts" chart showcases the users (or, more accurately, pages) with the highest number of posts within the dataset. This visualization helps to identify the most prolific contributors, potentially highlighting key influencers, active community members, or news sources. Notably, not a single Facebook page has more than 100 posts, far lower than any news organization would have published in the last four years.

The fluctuating engagement trends observed in the dataset, with noticeable peaks in specific years, reflect the evolving nature of user interaction on Facebook. These trends suggest that many factors, including global events, algorithmic changes, and shifts in content preferences, influence user engagement. The implications for social media strategists and content creators are clear: staying attuned to the broader socio-cultural

context and adapting content strategies accordingly can significantly impact engagement levels. High-engagement posts, particularly those from "Humans of New York," demonstrate the power of storytelling and emotionally resonant content in driving user interaction. This underscores the importance of content quality and relevance in capturing audience attention and fostering engagement. For content creators and publishers, this highlights the value of investing in content that genuinely resonates with their audience's interests and experiences, suggesting a shift towards more local and human-centric content strategies.

This diversity indicates multiple paths to achieving visibility and engagement on the platform, aligning with Zuboff's (2019) concept of surveillance capitalism, which emphasizes monetizing data acquired through surveillance. For marketers and social media professionals, a one-size-fits-all approach may not be practical, perhaps explaining the lack of engagement with most of Metric Media's Facebook pages. Instead, understanding the unique characteristics of their audience, informed by data analysis, and tailoring their strategies accordingly can lead to better engagement outcomes (Smith & Chaffey, 2022). Practically, these insights inform the development of targeted content strategies that leverage the types of content and engagement patterns most likely to resonate with users, enhancing the effectiveness of social media campaigns.

Creating a social media presence is a foundational element of a brand's online identity, contributing to increased brand awareness, improved customer engagement, enhanced brand loyalty, and competitive advantage. The digital age necessitates a strong social media presence as a marker of professionalism and currency. Particularly for small businesses, social media branding is crucial for establishing a solid online presence, connecting with customers, and fostering business growth (Smith & Chaffey, 2022).

The strategy of signing up for social media profiles for brands, influencers, businesses, and media entities without the immediate intent to foster growth through those pages is a nuanced approach in the digital marketing landscape. This tactic serves several purposes, such as protecting against copycat impersonations, preserving the opportunity to leverage these profiles in future marketing strategies, and enhancing a business's asset portfolio for potential sale scenarios.

Moreover, optimizing social media bios and profiles encapsulates a brand's essence and main messages, acting as an essential pitch to visitors. A study by Nazarudin and Handayani (2024) underscores the role of consistent brand presentation in building a positive brand image and increasing sales of exported products. This optimization includes summarizing the benefits of products or services, incorporating relevant keywords for better discoverability, and ensuring brand assets like logos and themes are consistent across platforms to aid brand recognition (Nazarudin & Handayani, 2024).

Selecting appropriate social media platforms is crucial for aligning with a brand's target audience and content strategy, as each platform attracts distinct user demographics and content preferences. For instance, TikTok has gained significant popularity among younger audiences, particularly teenagers. A December 2024 report by the Pew Research Center indicates that 63% of U.S. teens use TikTok, with 16% engaging with the platform almost constantly (Faverio & Sidoti, 2024). In contrast, Instagram maintains a strong presence among young adults, with 71% of individuals aged 18 to 29 utilizing the platform, according to a 2021 Pew Research Center study (Auxier & Anderson, 2021). Understanding where the target audience spends time and their content preferences can significantly influence a brand's social media strategy. For example, a study published in *Computers in Human Behavior* found that personality traits such as extraversion and openness are

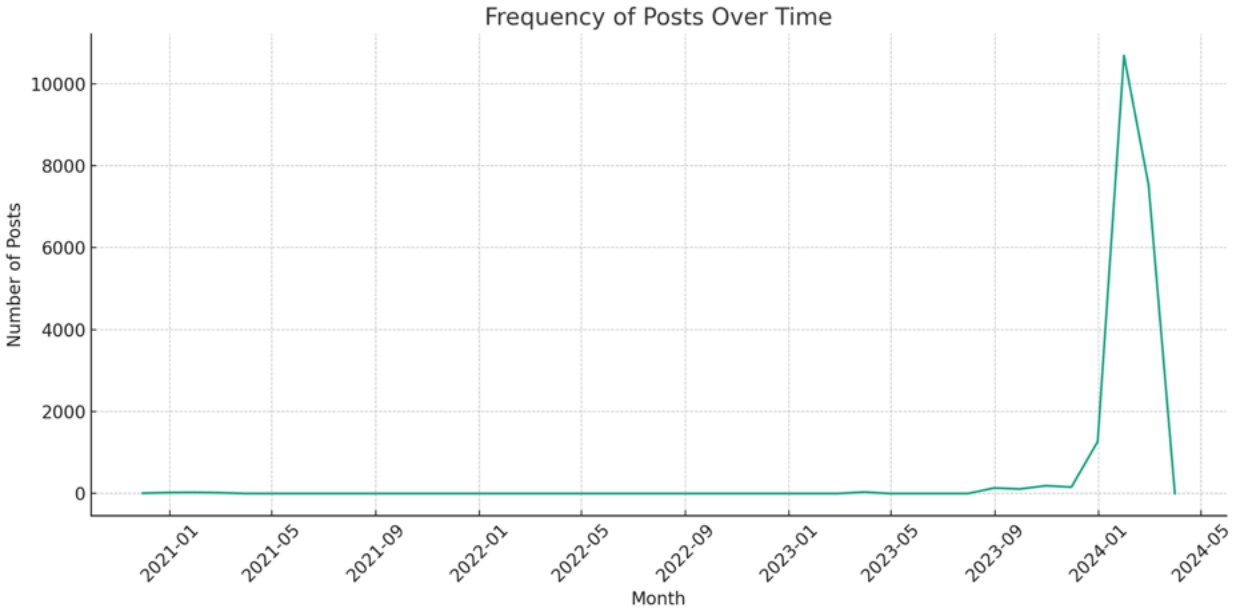
positively correlated with social media usage, suggesting that user engagement varies based on individual characteristics (Correa et al., 2010).

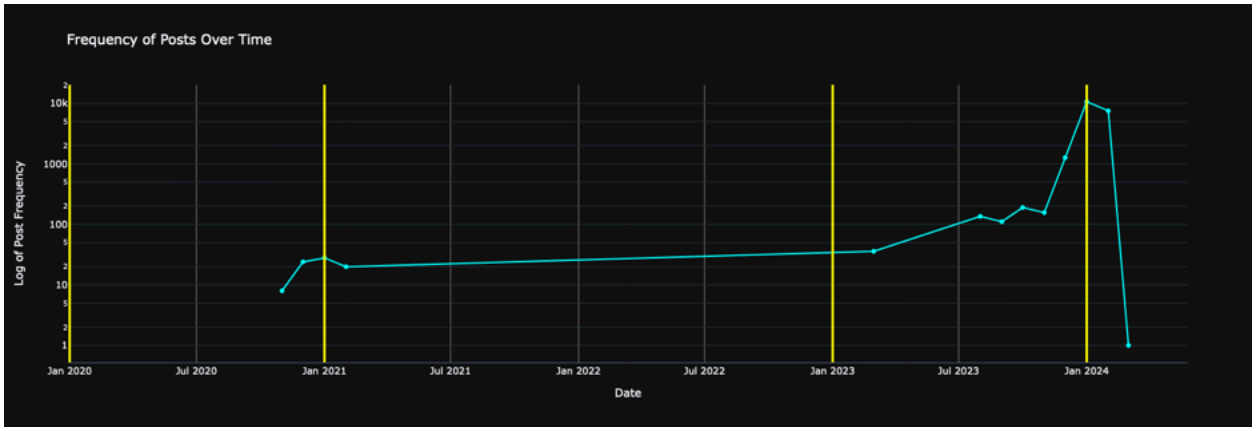
While the dataset provides a rich source of insights, its coverage is limited to a specific set of users/pages and periods. This may affect the ability of the findings to draw generalized conclusions applicable to other contexts or platforms. The analysis is primarily quantitative, focusing on engagement metrics and post-activity and may come up short when measuring the qualitative aspects of content that contribute to engagement, such as narrative quality, emotional resonance, and visual appeal. Incorporating qualitative analyses or mixed-methods approaches could provide a more nuanced understanding of the factors driving user engagement on social media platforms.

The analysis results, employing data scraping and visualization techniques on Metric Media's content, offer significant insights into the organization's operational strategies, demographic targeting, and thematic focuses. Two primary analytical approaches were employed: Time Series Analysis and Word Clouds.

The time series analysis illuminated the ebbs and flows of online post frequency within the specified period. A discernible pattern emerged, showcasing the rhythmic pulsations of digital engagement across months. This temporal mapping highlighted peak periods of content proliferation and shed light on quieter intervals, suggesting a nuanced

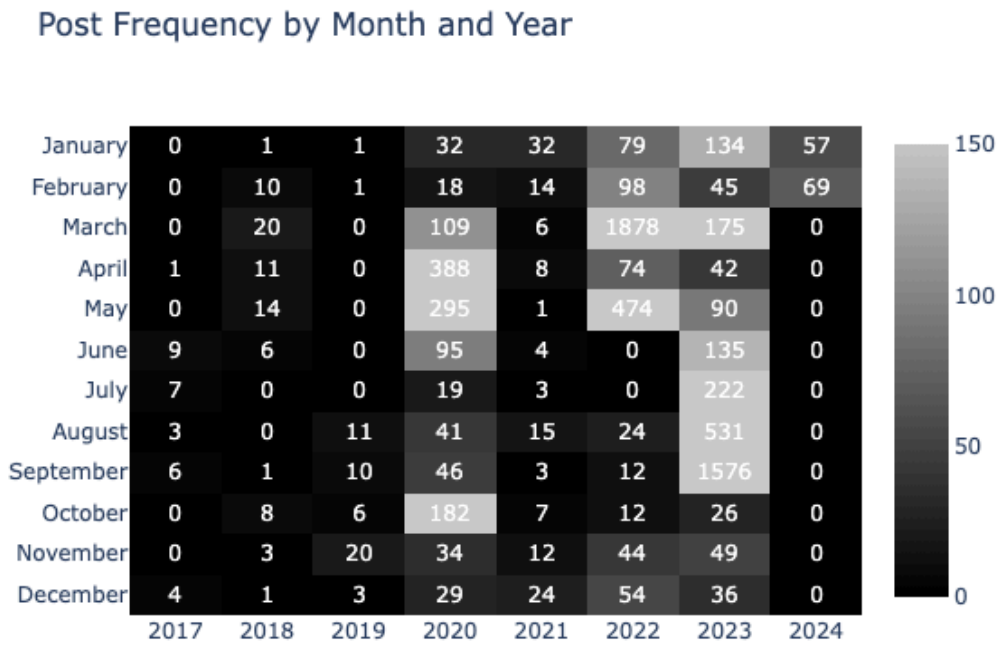
interplay between global events and online activity.





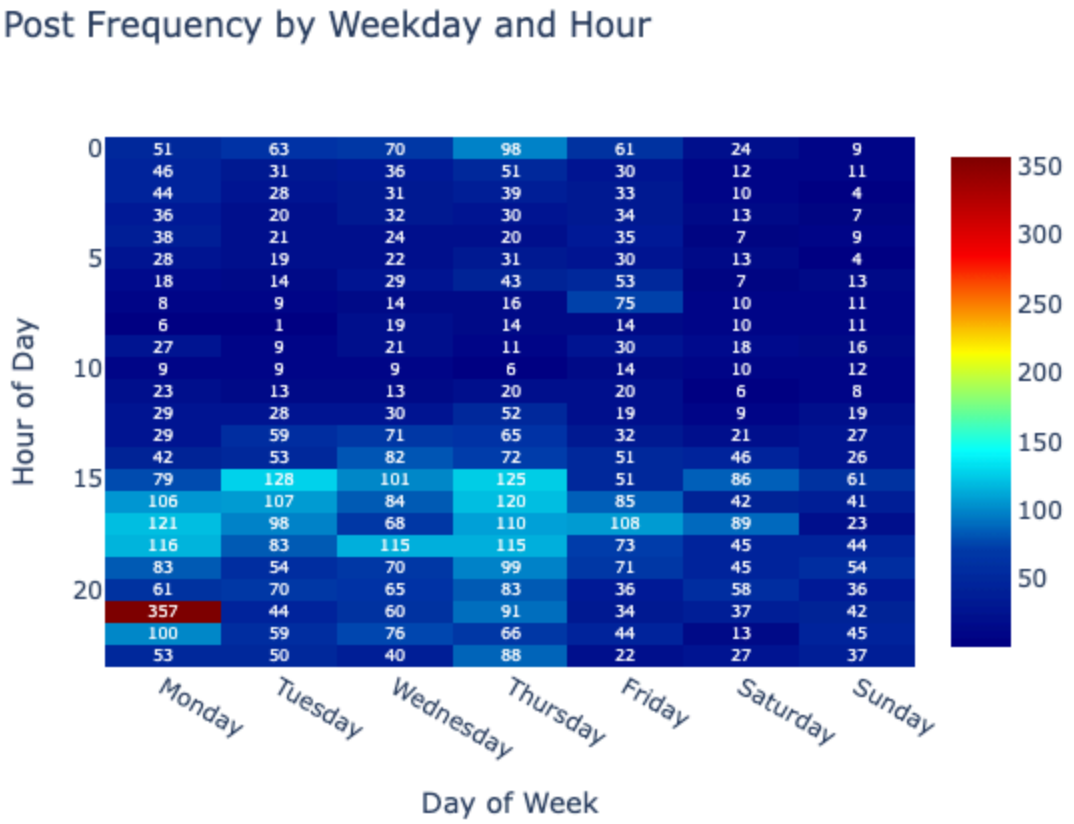
From the image, we can see that the frequency of posts has remained mostly stable over time, with an increase beginning in August 2023 and continuing until the end of the time frame that was applied.

In addition, “heat maps” give us a closer understanding of the posting schedule by month, demonstrating much more effectively how posts have been increasing across

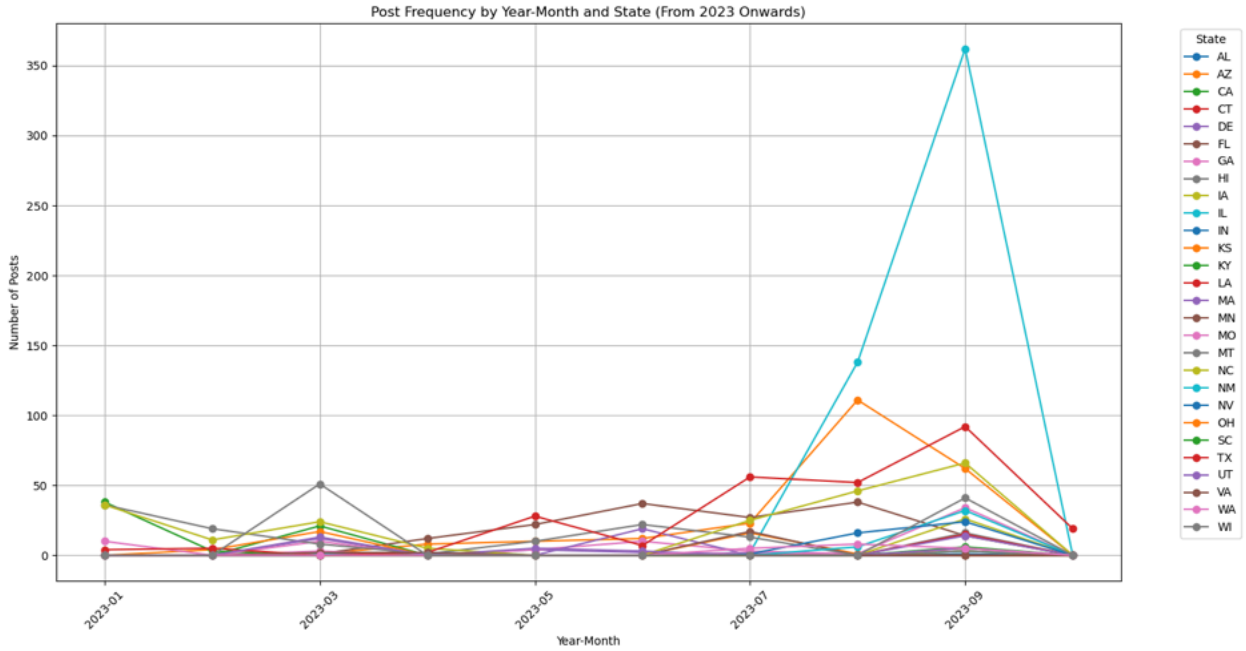


platforms:

When further analysis is conducted, showcasing post frequency by day of the week, it becomes clear that the publishers at Metric Media generally post most frequently during the middle of the business day, with most posts made on Fridays and Saturdays. This is likely to attract the largest audience outside of the workweek. It could also be due to the funding from religious sources, who usually hold services during the weekend:



Finally, by categorizing posts by state, we can see how Metric Media is concentrating its efforts on specific U.S. states:

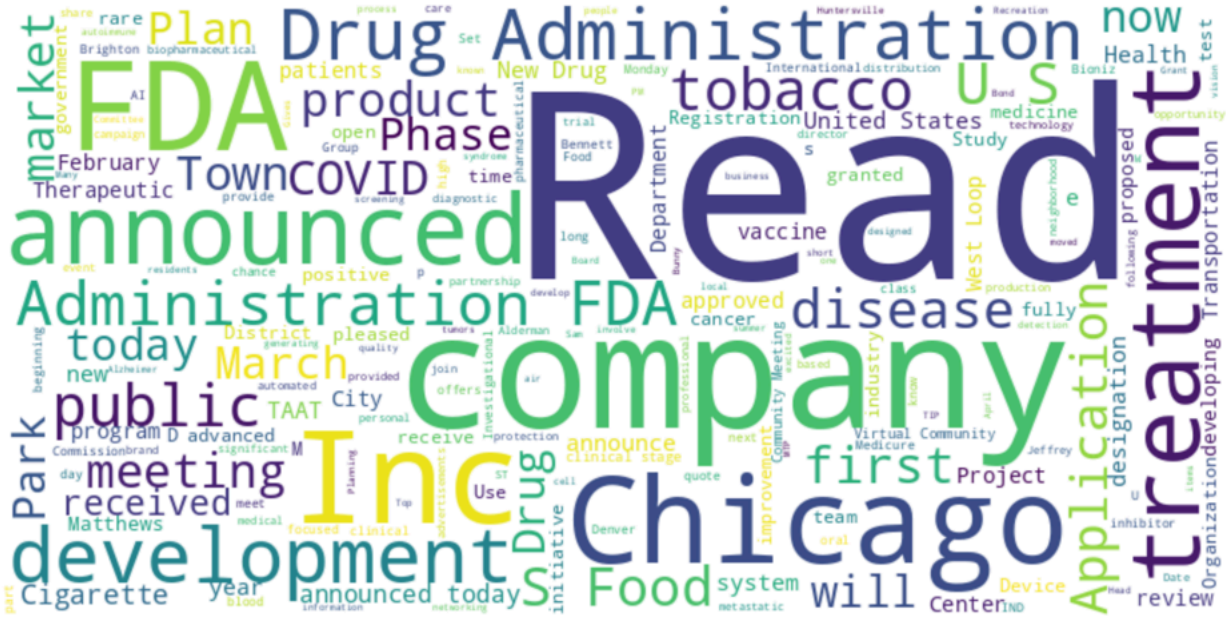


The top states open to Metric Media's information were Illinois, Texas, Arizona, North Carolina, and Florida. This is a crucial discovery, as these states are pivotal in U.S. politics due to their significant electoral influence and evolving political landscapes. These states are often classified as "swing states," meaning they do not consistently favour one political party in presidential elections, becoming central battlegrounds during campaigns (Hernandez, 2024). Their diverse and rapidly changing demographics contribute to this unpredictability. For instance, Arizona and Texas have experienced substantial growth in Latino populations, a demographic with varied political preferences, influencing electoral outcomes (Dionne & Keeney, 2025).

Additionally, these states exhibit pronounced urban-rural divides; urban centers like Chicago (Illinois), Houston (Texas), and Miami (Florida) tend to lean Democratic, whereas rural areas often support Republican candidates. This dichotomy underscores the complex political dynamics within these states. Furthermore, issues such as gerrymandering have

been prominent, with states like North Carolina facing legal challenges over district boundaries, highlighting the contentious nature of redistricting processes (Daley, 2024).

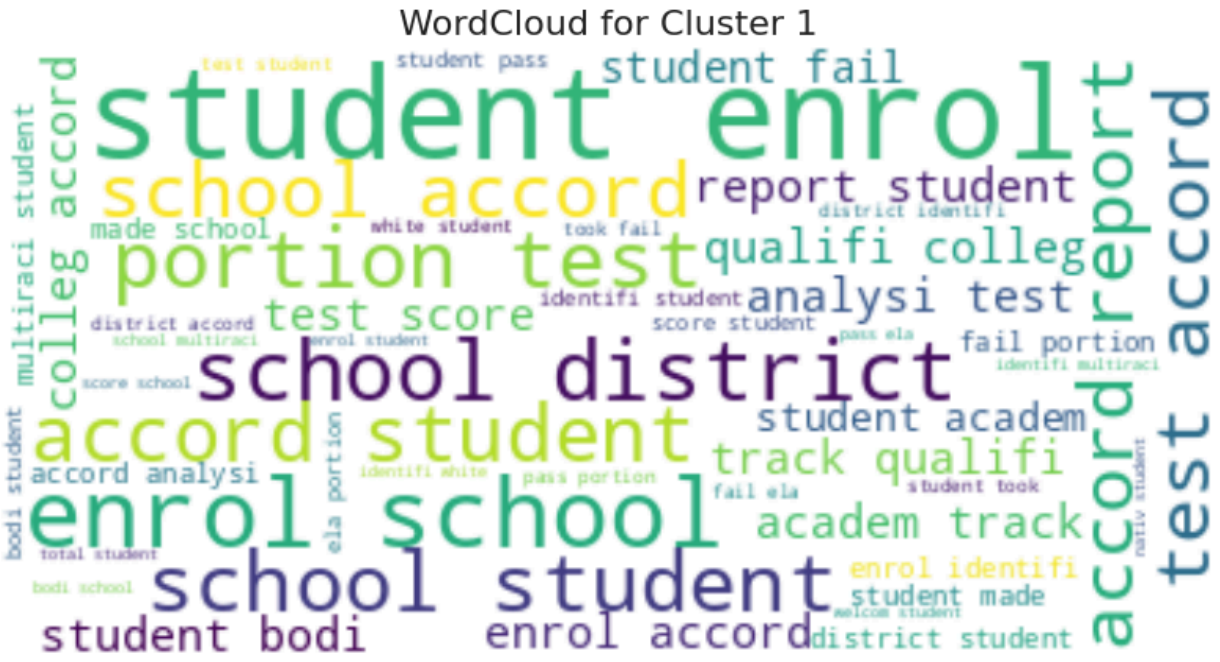
Through the generation of word clouds, the linguistic essence of the posts may be revealed, including the most prominent terms and themes that dominated the digital narrative within our timeframe. This visual exploration revealed a tapestry of topics, reflecting online communities' diverse interests and concerns. The prevalence of specific terms not only highlighted key areas of public interest but also offered insights into the shifting priorities and discourse patterns over time:



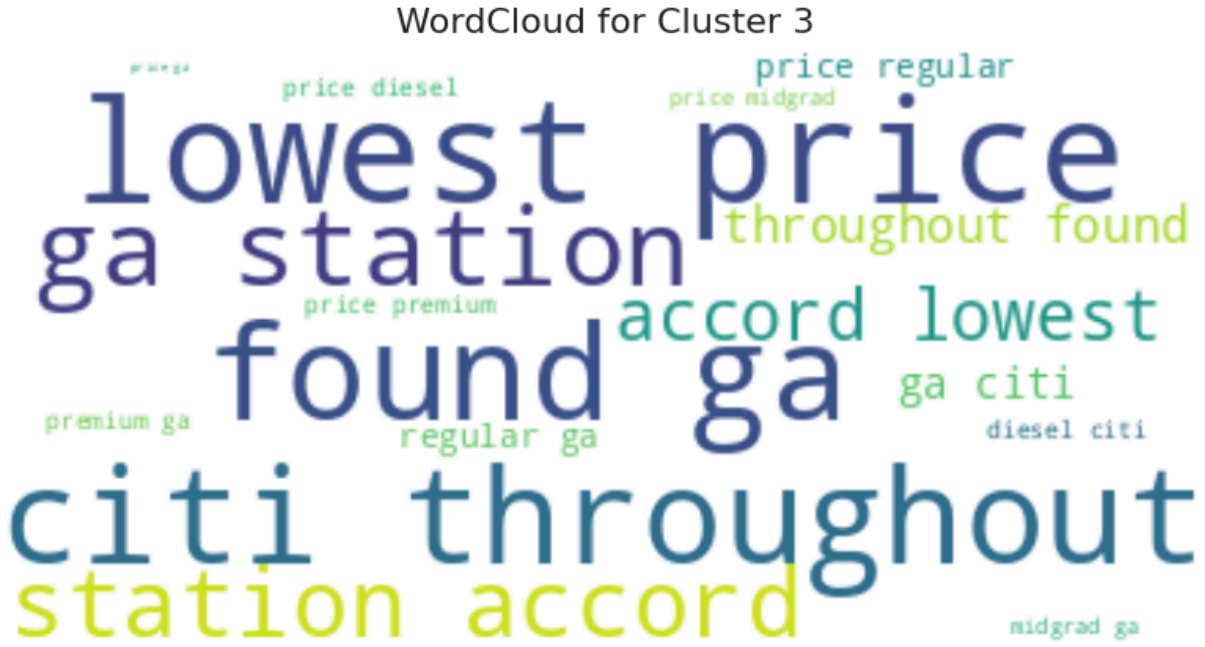
The observed increase in posting frequency during specific periods suggests an alignment with significant news events or targeted campaigns. This temporal pattern underscores the dynamic nature of content strategy, possibly aimed at maximizing audience engagement during peak interest times.

Word clouds generated from the scraped data visually represented the most frequently occurring terms and themes within Metric Media's posts. By drilling down

further using our five key areas of topic cluster analysis, we can see frequent terms used in the copy published to Metric Media's properties. For example, standard school terms emerge using the education cluster of posts.

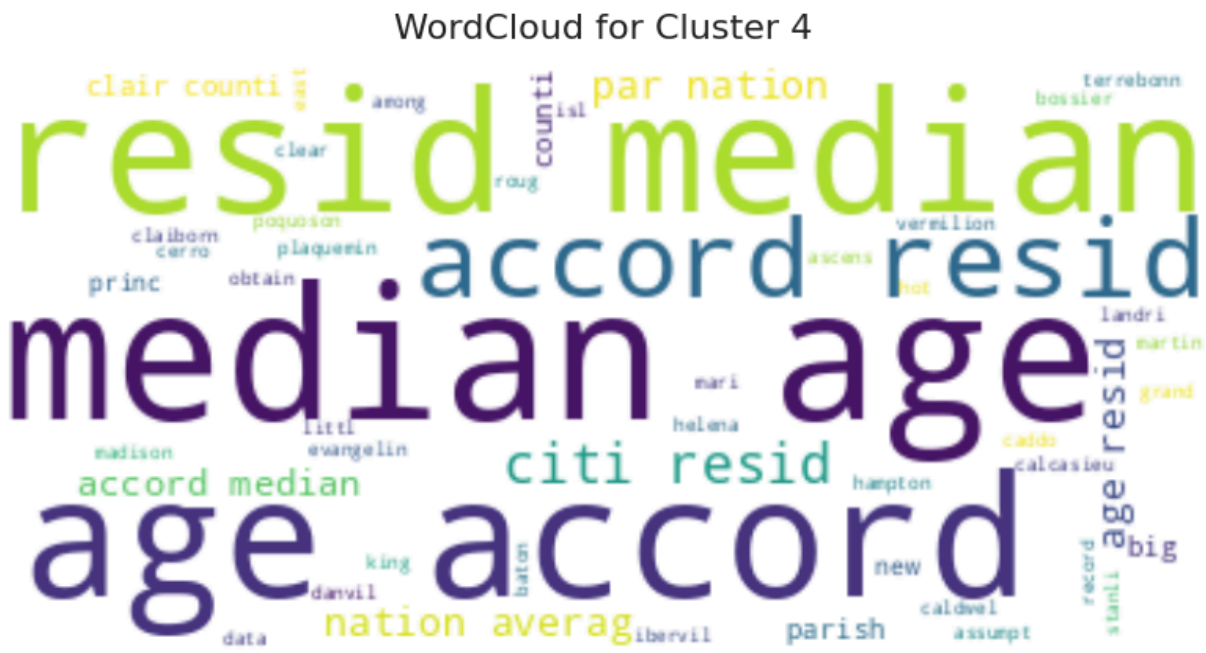


There is a similar pattern with posts related to gas prices:

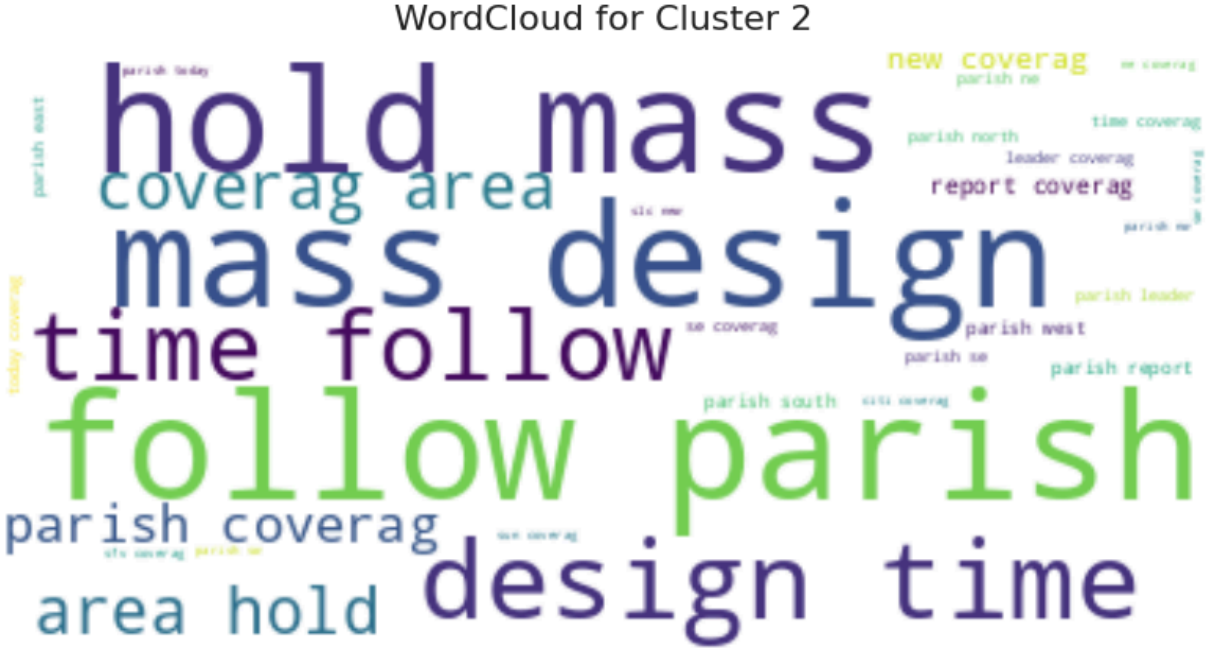


Frequent terms emerge when observing legal discussions:

A similar pattern with posts pertaining to demographic data:



Finally, for posts connected to religious services:



The prominence of specific terms and themes in the word clouds reveals Metric Media's strategic content themes, likely aimed at resonating with its audience's interests and concerns. However, its “audience” might be the wider public and those who fund Metric Media’s operations. The thematic analysis facilitated by word clouds offers a snapshot of the narrative and discourse-shaping efforts undertaken by Metric Media, providing insights into the organization's thematic priorities and content strategy. Notably, the five word clouds represented are themes in the rhetoric of the American right.

In the context of U.S. politics, issues such as demographics, religion, the legal and justice system, oil and gas prices, and education function not merely as policy domains but as ideological battlegrounds that reflect and reinforce broader political divisions. These domains have become central to partisan identity formation and mobilization, particularly as political actors increasingly leverage them to shape narratives around national values and belonging (Sides et al., 2018). For instance, demographic shifts tied to immigration and race have fueled anxieties that are politicized through rhetoric about voter fraud and border security (Hopkins, 2018). Religious affiliations similarly underpin debates on reproductive rights, LGBTQ+ inclusion, and the role of religion in public schools, serving as a primary mobilizer for evangelical political engagement (Whitehead et al., 2018). The justice system is often framed through divergent lenses of “law and order” versus systemic reform, with race and policing at the center of partisan discourse (Clegg & Usmani, 2017). Energy debates, especially around oil and gas, link economic anxiety to environmental policy, with fossil fuel interests often aligning with conservative agendas opposing regulation (Stokes, 2020). Finally, education has become a flashpoint for cultural conflict, particularly surrounding curriculum content, parental control, and equity initiatives, highlighting how schools serve as arenas for ideological contestation (Kretchmar et al., 2014). Collectively,

these topics serve as proxies for deeper ideological commitments, reinforcing polarization and shaping electoral strategies.

The temporal trends, geographical distribution, and thematic focuses uncovered in our study of Metric Media's content offer valuable insights within broader media analysis frameworks. These findings suggest that Metric Media employs a highly adaptive content strategy to maximize audience engagement through targeted and localized content. This strategy aligns with contemporary understandings of digital media operations, where customization and relevance are crucial to capturing audience attention in a highly fragmented media landscape. The observed fluctuations in content production and demographic targeting underscore the dynamic nature of digital media strategies, reflecting broader trends in audience segmentation and targeted marketing. Moreover, the geographical focus highlights the importance of a local focus in media strategy, which is part of a more considerable trend towards hyper-localized content, catering to niche audiences with specific regional interests.

The findings underscore the value of data scraping and visualization techniques in media analysis. Such approaches can uncover nuanced insights into media operations, offering a deeper understanding of content strategies and audience engagement tactics. Researchers are encouraged to leverage these tools to explore other media outlets, facilitating comparative analyses that enrich our understanding of the digital media landscape. Awareness of the patterns and strategies identified through our analysis can inform content creation and distribution approaches. Understanding the significance of localized content and demographic targeting may inspire journalists and media houses to adopt similar strategies, potentially leading to more engaged and diverse audiences. Raising awareness about the sources and nature of their news consumption is crucial in an

era of information overload. Insights from this analysis can contribute to media literacy efforts, helping the public discern the intentions behind content strategies and the importance of consuming a diverse range of media to obtain a well-rounded understanding of local and global issues.

The methodologies and findings presented in this study open avenues for future research, including longitudinal analyses to track changes in media strategies over time and comparative studies with other media outlets. More research could provide a richer context for understanding Metric Media's place within the broader media ecosystem and its impact on public discourse and societal engagement.

5. Discussion

Manufacturing Legitimacy through Local Branding and Neutrality

RQ1: Constructing the Pseudo-Local News Ecosystem. The findings reveal that Metric Media has manufactured an aura of legitimacy by building an expansive network of locally branded news sites and maintaining a strategically neutral tone in its content. In practice, the organization operates over a thousand hyper-local websites across the United States, each mimicking the look and feel of small-town journalism (Benjamin, 2022). This "pink slime" journalism model has filled the void left by collapsing local newspapers, emerging in news deserts under familiar names like Mobile Courant or Suffolk Reporter (Benjamin, 2022). By presenting itself as a constellation of community news outlets, Metric Media constructs credibility by association: audiences tend to trust media that looks local and impartial. The network's neutral branding and generic newspaper aesthetics mask its centrally coordinated, partisan agenda, allowing it to launder its reputation as a legitimate news source while pursuing ideological goals (Napoli, 2019). In effect, Metric Media has created a pseudo-local news ecosystem that appears citizen-oriented on the surface but is orchestrated as a cohesive propaganda machine behind the scenes.

Crucially, Metric Media's content is not generated through traditional local reporting but through automated, mass-produced storytelling. Investigations have found that the vast majority of articles across Metric Media's sites are algorithmically generated, with only a handful bearing human bylines (Benjamin, 2022). This assembly-line approach to news production enables the network to scale rapidly–one study documented that Metric Media

launched hundreds of new outlets in the 2020 U.S. election–and to push out politically curated content with minimal cost or oversight (Benjamin, 2022; Bengani, 2021). Regarding content dissemination, Metric Media leverages programmatic techniques and centralized templates to ensure message consistency across regions. The appearance of localism (e.g., localized titles and minor geographic tweaks in stories) conceals a highly uniform editorial line that flows from a central political strategy. By answering RQ1 (How does Metric Media construct and disseminate its pseudo-news ecosystem?), the research demonstrates that the network's architecture itself–a web of automated local-lookalike sites–is a core part of its propaganda strategy. Metric Media constructs legitimacy by embedding within local trust networks and disseminates its content broadly through algorithm-friendly channels while maintaining a façade of community journalism.

Maintaining a neutral or nonpartisan tone across most stories is central to this legitimacy. The content analysis showed that 61.9% of Metric Media's articles carry neutral sentiment, with only 10.9% negative and the remainder mildly positive. This overwhelming neutrality is a calculated form of strategic ambiguity designed to avoid triggering readers' skepticism. By seldom overtly editorializing or revealing partisan cues, Metric Media's stories read like straightforward local news updates–roadside zoning decisions, church event recaps, fuel price notices–rather than propaganda. Such strategic neutrality serves as "propaganda minimalism," a concept this research proposes to describe propaganda that operates with minimal explicit bias. In propaganda minimalism, ideological content is distilled into subtle choices of topic and framing rather than heavy-handed opinion, making the influence pervasive but hard to pinpoint. This approach draws on a key insight of modern propaganda theory: persuasion is most effective when it does not appear to be persuasion at all. As classic models note, propaganda has evolved from blunt slogans to

sophisticated, low-key tactics that fly under the radar of audience resistance (Ellul, 1965; Herman & Chomsky, 1988). Metric Media's neutral style exemplifies this evolution. It quietly sets agendas by highlighting specific issues (e.g. crime statistics or school board decisions) without accompanying commentary, relying on the audience's assumption that a neutral news outlet is simply reporting important community news. This aligns with Agenda-Setting Theory: selectively emphasizing topics (and omitting others) under a neutral guise, the network influences what readers perceive as salient without appearing biased (McCombs & Shaw, 1972; McCombs, 2008). In short, the network manufactures legitimacy and trustworthiness through tonal restraint. Readers repeatedly exposed to a seemingly impartial local outlet are more likely to accept its information at face value. Psychological studies indicate familiarity breeds credibility–the more one encounters a news source that maintains an objective veneer, the more trust it accrues (Lewandowsky et al., 2017). Metric Media exploits this familiarity effect by saturating local information spaces with routine and benign content, thereby inoculating its audience against suspicion even as it subtly primes them with partisan narratives.

Algorithmic Amplification and Micro-Targeted Reach

A defining feature of Metric Media's operation is its sophisticated use of algorithmic amplification to disseminate content and micro-target specific audiences (further addressing RQ1). In the digital media ecosystem, content distribution is governed mainly by platform algorithms and targeted advertising. Metric Media capitalizes on this by primarily focusing its social media outreach via Facebook, which the engagement analysis showed as the network's dominant social platform. This focus is strategic: Facebook's algorithm tends to favour locally relevant stories and community group sharing, and its

user base skews older and more politically active–precisely the demographics Metric Media aims to influence (Napoli, 2019). The study found that Metric Media's posts garnered significantly more engagement on Facebook than on X/Twitter, indicating a calculated alignment with a platform where misinformation can spread with less scrutiny. By contrast, the near absence of the network on Twitter (a platform heavily monitored by journalists and fact-checkers) potentially suggests an intentional avoidance of watchdog communities that might flag its content. In essence, Metric Media chooses its battleground–a semi-closed environment (Facebook) where algorithmic curation can quietly boost its stories to receptive locals and eschews arenas where it would face pushback.

Beyond organic reach, paid micro-targeting amplifies Metric Media's influence. Evidence indicates that Metric Media and its affiliates have invested heavily in targeted Facebook advertising to push their articles into the feeds of specific demographic and geographic groups (Benjamin, 2022). By mining data on users' location, interests, and political leanings (a practice enabled by the surveillance capitalist infrastructure of social media), the network can zero in on communities most likely to be swayed by certain narratives. For example, a story about rising gas prices or a "crime wave" can be delivered directly to users in suburban or rural counties where those themes resonate politically. According to media watchdog reports, Metric Media and similar partisan networks spent millions of dollars on Facebook and Instagram ads aimed at swing voters in battleground states ahead of the 2020 and 2022 elections (Benjamin, 2022). This combination of localized content and micro-targeted delivery amounts to an algorithmically turbocharged propaganda model: the network supplies a steady flow of tailor-made stories, and platform algorithms (both organic ranking and paid promotion systems) ensure those stories reach the precise audiences that will amplify the network's partisan objectives.

Importantly, algorithmic design biases tend to work in Metric Media's favour. Social media algorithms are known to reward content that drives engagement, which often means emotionally charged or confirmatory information. Even though Metric Media's tone is neutral, the topics it emphasizes (e.g., crime, demographic change, contested civic issues) are chosen for their latent emotional or political charge. When such stories enter Facebook's feed ranking system, they can benefit from users' existing biases. For instance, a local-sounding story on crime may get more shares in communities already anxious about security. Recent research on Twitter's algorithms found that posts linking to low-credibility news sources received disproportionately higher amplification regarding visibility and impressions (Corsi, 2024). Although that study focused on Twitter and specific issues (Corsi., 2024), the broader implication is that recommender systems often amplify "problematic" content inadvertently. In Metric Media's case, a neutral façade might evade content moderation or user skepticism, allowing its pieces to circulate widely before any alarms are raised. As Tufekci (2017) and Pariser (2011) have argued, the architecture of social media creates echo chambers and filter bubbles that can reinforce a group's existing worldview by continually feeding it similar content. Metric Media's micro-targeting exploits this tendency: each community's dedicated "news" site curates stories that reaffirm particular ideological perspectives (e.g., emphasizing traditional values or market-friendly economics), which are then algorithmically fed to that community's members, tightening the feedback loop. The result is a form of audience-specific agenda-setting that operates primarily out of sight. Unlike the broad headlines of national news, these micro-targeted stories seep into local consciousness through Facebook shares, local Google searches, and even direct email newsletters or printed mailers (Benjamin, 2022). Readers encounter them not as propaganda but as serendipitous local news, heightening their persuasive potential.

From the perspective of RQ1 (dissemination), the analysis underscores that Metric Media's reach is amplified not by a broadcast tower or a printing press but by data-driven algorithms and platform dynamics. The network's innovation is marrying old-fashioned local news tropes with cutting-edge micro-targeting. This synergy ensures that ideologically slanted narratives achieve maximum penetration in their intended communities while mainly bypassing the filters that protect the broader public sphere. Metric Media demonstrates how a propaganda operation can weaponize the tools of the digital attention economy–from engagement-optimized algorithms to granular audience data–to boost its signal and quietly shape political discourse at the community level. If traditional agenda-setting was about telling the public what to think about, Metric Media's algorithmic agenda-setting tells each subset of the public what to think about in a customized manner, thereby fragmenting the information space along ideological lines (Benkler et al., 2018).

Ideological Minimalism: Thematic Framing of Content

RQ2: Patterns in Content Strategy. Beneath its neutral tone, Metric Media's content exhibits clear thematic patterns that advance an ideological narrative through carefully curated topics. The topic modelling in this study identified five dominant themes across the network's articles–demographics, religion, civic governance, economics, and education–each corresponding to politically charged domains that are subtly framed to align with conservative or libertarian agendas. This approach can be described as ideological minimalism, wherein each story contains only implicit or minimal signals of bias, yet collectively, the coverage paints a one-sided picture of societal issues. By explicitly addressing RQ2 (identifiable patterns in content strategy), this study finds that Metric

Media's content strategy is to hammer on specific motifs that reinforce its ideological worldview, all while maintaining the appearance of standard local reporting.

One prominent theme is demographic change, often presented through dry statistics about population, age, or migration in local areas. On the surface, these reports seem like neutral community facts. However, in aggregate, they support a narrative of a country in flux–a subtext often exploited in nativist or conservative rhetoric. Readers are continuously reminded of demographic shifts (e.g., influxes of new residents or changes in community makeup), priming concerns about identity and cultural change without a single overt comment. This data-driven framing resonates with what Tufekci (2017) notes about digital propaganda: it can strengthen pre-existing worldviews by feeding audiences selectively curated facts that validate their perceptions. By tailoring demographic reports to specific locales (for instance, highlighting increases in certain ethnic populations in one area or youth population decline in another), Metric Media reinforces the audience's pre-existing concerns or biases. A resident already anxious about immigration might frequently see local "news" of population changes that subtly fan that anxiety, all under the guise of factual reporting.

The religious theme is another striking example. Metric Media sites frequently cover church events, parish activities, and faith-oriented community news. Far from mere community bulletins, these pieces serve a dual purpose: they engage faith-based readers and normalize a traditionalist, religious presence in the public sphere. In the data, references to parishes, masses, and religious schools were recurrent, aligning the network's content with the interests of Christian conservative groups (Bennett & Livingston, 2018). By consistently featuring religious content, Metric Media taps into the trust networks of religious communities and subtly links faith values with political discourse. For example,

an article about a charity drive at a local church might appear apolitical. However, its steady presence alongside political content helps interweave religiosity with civic issues in the reader's mind, echoing the agenda of specific advocacy organizations that fuse Christianity with conservative policy goals.

A third theme concerns legal and civic governance issues, such as voter registration updates, crime reports, and regulatory changes. These topics are presented as civic information–e.g., how to register to vote, new laws about data or offender registries–but often with a framing emphasizing law and order and procedural integrity. By frequently mentioning voter rolls or crime statistics, the sites project an image of being concerned with public safety and electoral transparency. However, the absence of deeper investigative context (for instance, no discussion of voter suppression or wrongful policing) means the coverage tilts towards endorsing a status-quo, security-focused viewpoint (Marwick & Lewis, 2017). Repeated articles on voter registration deadlines, for instance, may seem innocuous, but in certain states, such content ran alongside partisan narratives about voter fraud. The cumulative effect is to legitimize a specific voter integrity discourse without openly making false claims–a form of agenda insinuation that positions Metric Media as an authority on civic matters while hinting at the distrust of new voters or outsiders.

The economic theme in Metric Media's content highlights local business news, especially fuel prices and market trends, but conspicuously omits coverage of progressive economic topics. Stories on gas prices rising or falling are familiar–offering immediate utility to readers–yet discussions of renewable energy initiatives or labour issues are largely absent (Pickard, 2020). This selective economic coverage favours a conservative economic perspective that prizes traditional industries (oil, gas, small businesses) and sidesteps critiques of those industries or exploration of alternatives. By relentlessly

focusing on the cost of fuel and everyday economic concerns, the network reinforces pro-market, anti-regulation sentiments without ever stating them outright. The silence on environmental or progressive economic policies is itself telling; it frames the scope of debate to exclude those perspectives, a classic propaganda technique of omission (Entman, 1993). Readers of these local sites receive a steady diet of economic news that implicitly valorizes free-market conditions and casts doubt on regulatory or green initiatives simply by never mentioning the latter. The resulting worldview is one where the only economic issues that exist are those that align with conservative talking points (taxes, gas prices, local business climate), creating an ideologically slanted public consciousness under the pretense of ordinary business news.

Finally, the education theme frequently appears, with stories about school board decisions, student enrollment numbers, and curricula or school choice policies. These articles often highlight charter schools or privatization efforts (Spring, 2014), echoing libertarian and conservative education reform narratives. For instance, a local site might run a benign-seeming report on a new charter school opening or changes in district zoning, which, taken alone, is standard news. However, across the network, the pattern emphasizes school choice and competition over public school investment. Again, this mirrors broader ideological aims to promote privatization and challenge teachers' unions without explicit editorializing. It is propaganda through framing: the consistent focus on certain types of education stories (and neglect of others, like public school successes or funding challenges) gently nudges readers to view educational issues through a particular lens–one that favours free-market solutions in education and skepticism toward public school systems.

All these thematic patterns illustrate Metric Media's ideological minimalism strategy: each article is a minimalist vessel carrying a hint of ideology. There is rarely an

outright falsehood or a blatant partisan rant. Instead, there is a consistent asymmetry in what is covered and how. By explicitly mapping these patterns, this research answers RQ2–the identifiable content strategy lies in what Metric Media talks about and what it does not, the sentiments it expresses and those it withholds. The neutral or positive tone of most articles (recapping a church fundraiser in uplifting terms, for example) creates a feel-good local vibe. At the same time, selective negativity (the mere 10.9% of negative content) is reserved for moments of political salience. Notably, the network seems to deploy negative or fear-inducing content sparingly but tactically–for example, heightening criticism or alarmism in the lead-up to elections or contentious votes, when such emotions can be most influential. This intermittent negativity, contrasted with generally neutral coverage, can catch readers off guard and have an outsized impact because it comes from an outlet they have come to see as objective. It is as if the mask drops only at critical moments, delivering an ideological punch precisely when it is likely to matter most (Benkler et al., 2018).

Credibility Laundering in a Global Context

While Metric Media is a domestic phenomenon, its tactics fit into a global and historical pattern of propaganda networks that exploit trusted media formats to disseminate disinformation. This study situates Metric Media alongside other propaganda enterprises–from the Sinclair Broadcast Group in the United States to international state-run outlets like Russia's RT (Russia Today) and China Daily–to illustrate how credibility laundering operates across contexts. Credibility laundering is a conceptual insight this research puts forward to describe how propagandists borrow the credibility of trusted institutions (local newspapers, TV news anchors, or respected foreign press) to

legitimize their messaging. In each case, the propagandist dons the trappings of reputable media, thereby cleansing their content of overt partisan markings and making it more palatable to the target audience.

Metric Media's approach of embedding within local news ecosystems is unprecedented in scale but not entirely new in concept. Sinclair Broadcast Group, for example, has been described as using a similar laundering tactic in local television. Sinclair, which owns nearly 200 local TV stations in the U.S., made national headlines in 2018 when it mandated identical script readings by local news anchors, pushing conservative talking points under the guise of a homey news update (Berger, 2024). Viewers across different cities saw their trusted local TV personalities deliver synchronized warnings about "fake news" and partisan narratives, a strategy media analysts flatly labelled propaganda (Berger, 2024). What Sinclair did with the authority of TV anchors, Metric Media does with the familiar brand of local newspapers. Both exploit the notion that "if my local news is saying this, it must be true." Sinclair's ongoing practices–including a new centralized news program that often displaces genuine local reporting–highlight how ideological messaging can be injected into local trust channels (Berger, 2024). The result is the same: audiences find it harder to discern that the message originates from a top-down agenda, not a grassroots concern because it arrives via a source they inherently trust.

On the global stage, state-sponsored media like RT and China Daily have mastered credibility laundering on an international scale. RT (formerly Russia Today) presents itself as a cosmopolitan news network, hiring US journalists and adopting a polished, CNN-like aesthetic, all to mask its role as a Kremlin mouthpiece (Bond, 2024). For years, RT insisted it was independent even as it consistently toed Moscow's political line; it blended legitimate news reporting with subtle propaganda and conspiracy framing. This mirror imaging of

credible media allowed RT to gain a foothold in foreign media markets and influence discourse in democracies by seeming like another cable news option (Bond, 2024). Only later, as geopolitical tensions grew, skepticism of RT increased, and governments recognized it as an information warfare tool.

Similarly, *China Daily*, an English-language newspaper controlled by the Chinese Communist Party's Central Propaganda Department, has used paid supplements such as "China Watch" in U.S. newspapers to disseminate Beijing's narratives (Cook, 2013; Brady, 2009). For years, major U.S. outlets—including *The Washington Post* and *The New York Times*—published these sponsored sections, which were often designed to resemble standard news pages and not always clearly labelled as advertising content (Walker & Ludwig, 2017). By embedding Chinese state messaging within the pages of trusted American newspapers, these inserts served to reach educated readers under a veneer of journalistic legitimacy. This strategy effectively allowed the Chinese government to promote its perspectives on issues such as trade policy, territorial disputes, and human rights, thereby laundering propaganda through respected Western media institutions (Repnikova, 2017).

Metric Media can be seen as a domestic analog to these operations. Instead of foreign governments, domestic political actors and advocacy groups hide behind the front of local journalism. The network's founder, corporate backers, and content collaborators (some of which include partisan political groups (Benjamin, 2022)) remain in the shadows. In contrast, the local "newspaper" brand does the talking. This opacity in funding and ownership is a hallmark of credibility laundering: whether it is the CCP behind China Daily or undisclosed advocacy organizations behind Metric Media, the true source stays concealed so that the message carries more weight with the unsuspecting audience. Indeed,

media watchdog NewsGuard revealed that Metric Media quietly took in funding from conservative advocacy groups and coordinated content strategy with them, all while publishing under innocuous local mastheads (Benjamin, 2022). The message laundering was so effective that by the 2020 election, Metric Media had tripled its footprint and inundated key states with partisan content camouflaged as homegrown news (Benjamin, 2022).

Historically, propaganda has consistently sought channels of trust to disseminate its messages effectively. During the Cold War's early years, Soviet and Western intelligence agencies employed front organizations and sympathetic media outlets to plant stories that local media would subsequently pick up, thereby influencing public opinion. For instance, the British Foreign Office's Information Research Department (IRD) was established in 1948 to publish anti-communist propaganda and support anti-communist politicians, academics, and writers (Lashmar & Oliver, 1998). Similarly, the United States engaged in covert propaganda strategies, such as the Congress for Cultural Freedom, which aimed to influence cultural and intellectual communities worldwide (Coleman, 1989).

In the contemporary digital era, the scale and precision of such propaganda efforts have expanded significantly. The emergence of "pink slime" journalism, characterized by partisan content masquerading as local news, exemplifies this trend. In the case of local journalism, this credibility stems from a longstanding tradition of newspapers being integral to the community's fabric—reporting on school events, city council meetings, and local sports. Metric Media strategically intersperses genuine local news with other content to maintain this facade, reminiscent of earlier propaganda strategies that sought to blend authentic information with disinformation to subtly influence public perception (Benkler et al., 2018).

This strategy is akin to introducing a drop of poison into a large soup pot. While most of the "soup" consists of ordinary news, each serving delivers the toxic effect insidiously. Such methods underscore the importance of media literacy and critical evaluation of news sources in today's information landscape.

By comparing Metric Media to Sinclair, RT, and China Daily, this study underscores a key insight: strategic neutrality and localism are the new face of propaganda, as potent as overt state media broadcasts of the past. The global context shows that authoritarian states and partisan domestic actors converge on similar infiltration strategies. Russia's overt propaganda outlets and China's diplomatic media campaigns both realized that appearing unbiased and credible is more critical than ever in a skeptical information age. Likewise, Metric Media realized that to influence Americans, it should dress up as the most trusted messenger–the hometown paper. This comparative perspective highlights how Metric Media is part of a broader ecosystem of modern propaganda that prizes subtlety, deniability, and trust exploitation. As Benkler et al. (2018) note, the digital media environment has enabled a proliferation of such "networked propaganda" operations, each adapting the credibility-laundering tactic to its own locale and audience.

Countermeasures and the Role of Data Science

RQ3: Detecting and Resisting Micro-Targeted Propaganda. The rise of pseudo-local networks like Metric Media poses urgent questions about how journalists, policymakers, and citizens can detect and counteract such covert propaganda. This study's findings inform several strategies for resistance and accountability. It emphasizes that a

multifaceted response is needed to address the supply side (the networks producing disinformation) and the distribution side (the algorithms and platforms spreading it). In answer to RQ3 (practical strategies to detect, deconstruct, and resist micro-targeted propaganda), the research outlines interventions across the domains of transparency, literacy, journalism, and policy:

- **Transparency in Ownership and Funding:** Sunlight is a powerful disinfectant. Requiring disclosure of a media outlet's owners, sponsors, and affiliations would make hiding behind fake local names harder for operations like Metric Media. For instance, just as the European Union's Digital Services Act now mandates transparency in online content origins (Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and Amending Directive 2000/31/EC (Digital Services Act) (Text with EEA Relevance), 2022), similar regulations could compel any organization purporting to provide news to openly list who is behind it. If every Metric Media site had a clear label indicating its parent network and funding sources, readers could make more informed judgments about credibility. Regulators should consider updating press and campaign finance laws to treat these coordinated networks as political influence operations rather than independent local media. Mandatory funding disclosure would effectively neuter the "local" deception by revealing that what appears to be your neighbourhood outlet is part of a national partisan enterprise.

- **Algorithmic Accountability:** Social media platforms and search engines must be pressed to reveal and adjust how their algorithms amplify pseudo-news. If Facebook's newsfeed algorithm disproportionately boosts content from pink-slime

sites, this is something that should be audited and corrected in the interest of an informed public sphere. Scholars and advocacy groups have begun calling for algorithmic transparency requirements (Vaidhyanathan, 2018), which would force platforms to divulge criteria for news rankings and flag state-controlled or deceptive news sources. Platforms could implement interventions in the feed, such as demoting content from domains known to masquerade as news or providing context labels (e.g., "This source is part of a network that does not disclose ownership"). Moreover, advertising systems should be scrutinized: platforms should enforce due diligence on political ad buyers. In Facebook's case, had there been stricter verification for who is purchasing ads targeting voters (and an obligation to report such spending to users), Metric Media's massive targeted ad campaigns might have been better known and mitigated (Benjamin, 2022). The onus is increasingly on tech companies to balance their engagement-driven algorithms with the democratic necessity of not amplifying fraud. External algorithm audits and stronger regulation (akin to financial audits for banks) could ensure that the "invisible hand" of the algorithm is not invisibly distorting our news diets.

- **Strengthening Investigative Journalism and Fact-Checking:** The traditional media's response to information warfare must be to double down on investigative tactics but updated for the digital age. Collaborative journalism initiatives can be pivotal: investigative reporters, data scientists, and open-source researchers should form coalitions to unmask networks like Metric Media. In this study, by leveraging data science methods, it demonstrated one model of exposing a propaganda network's behaviour. News organizations could adopt similar approaches–for example, creating internal teams that continuously scrape content from suspected networks, using AI to identify patterns or automated text, and rapidly publishing

analyses to alert the public. Partnering with academia can also enhance these efforts (Carlson, 2020); journalism schools and computational social science departments can jointly build tools (e.g., machine learning classifiers for pseudo-news) that newsrooms alone might not have the resources to develop. Additionally, independent fact-checking consortia should specifically target pseudo-local news claims. Just as political fact-checkers debunk candidates' statements, a dedicated network of fact-checkers could monitor Metric Media's and its clones' outputs in real time, flagging false or misleading claims. Importantly, findings from these investigations need broad dissemination, perhaps via the channels propagandists use. For example, if a Metric Media story goes viral on Facebook, platforms could algorithmically insert a fact-check or a warning for users who see or share it (as is now done for some misinformation). Such responsive fact-checking requires that news organizations treat these "local" stories with the same scrutiny as nationally circulated fake news. The challenge is scaling up the attention: tens of thousands of community-specific articles are more complex to police than a single viral fake headline. This is where computational assistance (e.g., automated detection of suspicious language or content patterns) will be invaluable.

- **Media Literacy and Public Education:** Empowering readers to identify and disregard propaganda is a long-term, yet crucial, countermeasure. Media literacy 2.0 extends beyond recognizing biased language; it must include understanding the new forms of disinformation, like pink-slime local news. Educational institutions and public libraries can implement programs that teach citizens to verify whether a local news site is authentic. For instance, checking for transparent "About us" pages, examining the breadth of coverage, and using online tools to trace a site's metadata or affiliations. Curriculum developers should incorporate contemporary case studies

(like Metric Media) into news analysis lessons, demonstrating strategic neutrality tactics and how to spot them. Research suggests that inoculating people by warning them about micro-targeted propaganda can reduce effectiveness (van der Linden et al., 2020). Therefore, public awareness campaigns could be employed, perhaps even by state election boards or civic organizations, to alert communities that not all "local news" they encounter, especially online, is what it seems. The goal is for a populace that approaches such content with healthy skepticism and is equipped with the tools to cross-check it against reputable sources or report it as suspicious. Civil society groups might also crowdsource the monitoring of pseudo-local sites–akin to neighbourhood watch but for information, where volunteers in different states keep tabs on their so-called local outlets and share anomalies on a national platform.

- **Policy and Legal Responses:** Legislators increasingly acknowledge the threat of covert propaganda to democracy. Beyond transparency laws, other policy ideas include updating election advertising rules (so that electioneering via sham news articles is treated as a political ad in kind) and exploring the applicability of truth-in-advertising laws to news content when undisclosed political entities fund it. Another avenue is supporting public interest journalism financially to fill the void that Metric Media exploits–if authentic, local journalism flourishes, the audience for the fake version shrinks. Some experts advocate for direct government support or subsidies for local news (Pickard, 2020) to counteract the market failure that disinformation networks prey upon. Additionally, legal scholars have discussed adapting anti-propaganda provisions (which exist for foreign propaganda, such as the Foreign Agents Registration Act in the U.S.) to domestic actors that mimic news. While thorny from a First Amendment perspective, such measures underscore that society need not treat propagandistic fabrications as equivalent to bona fide

journalism. At a minimum, non-profit watchdogs and public agencies can maintain publicly accessible databases of known pseudo-news sites, which can feed into browser extensions or social media algorithms to warn users automatically.

RQ4: Data-Driven Methods for Accountability. A key contribution of this research is demonstrating that propagandists' tools–data mining, automation, algorithmic targeting–can be repurposed for accountability and analysis, albeit with critical epistemological caveats. In addressing RQ4 (can data-driven methods used for manipulation be mobilized for accountability?), the study answers in the affirmative: yes, but with a need for critical reflection on methodology. The research effectively reverse-engineered aspects of its playbook by employing web scraping, sentiment analysis, and topic modelling on Metric Media's output. This shows that computational journalism methods can illuminate patterns of bias and coordination that would be impossible to discern through traditional reading of news articles alone. The data-driven approach revealed quantifiable biases–e.g., the skew toward neutral sentiment and the defined set of recurrent topics – which provided concrete evidence of intentional strategy. Network visualization and engagement metrics helped trace how the content flows through social platforms, exposing the otherwise opaque dynamics of algorithmic amplification. In essence, the study turned the lens of big data back onto the propagandists, using their digital footprints to map their influence network. This reflects a broader epistemological point: data is a double-edged sword in the digital age. The same data that powers micro-targeting can empower researchers to unmask micro-targeters if collected ethically and analyzed critically.

However, mobilizing data-driven methods for truth-seeking raises essential methodological and epistemological considerations. For one, measuring propaganda through algorithms is fraught with the risk of abstraction and loss of nuance. Using a lexicon-based sentiment tool (TextBlob) provided a useful overall picture but could have missed subtleties like sarcasm or context-dependent meaning. This highlights a caution: quantitative proxies for propaganda (like sentiment scores) have limits. A low negativity percentage, for example, signalled strategic positivity/neutrality, but an automated tool might label something as neutral that a human would recognize as insidiously biased through framing. Future research should incorporate more advanced natural language processing (NLP) models (Devlin et al., 2019) and even human-in-the-loop validation to capture such subtleties. Epistemologically, scholars must ask: what counts as evidence of propaganda? Data science can count and cluster, but interpreting those patterns requires theoretical insight and context. The study approached this by grounding its data analysis in established communication theories (agenda-setting, propaganda model, etc.), which provided a lens to make sense of the numbers. This combination of computational and critical methods is a hallmark of computational social science and, as this study shows, is very potent in propaganda research. However, it also forces a reflection on methodology: the study effectively treats propaganda content as data to be objectified and analyzed, which could risk stripping away the human intent and deception behind it. Researchers must be careful to maintain a critical perspective, not letting the allure of big data's objectivity obscure the fundamentally normative questions at stake (e.g., what is false, what is manipulative, who benefits).

Another implication is the need for interdisciplinary literacy. Journalists and social scientists venturing into data science must ensure they understand the tools (and their

biases) they use; conversely, data scientists analyzing propaganda need political and cultural context. This study's approach exemplifies how bridging these domains can yield powerful insights–for instance, using topic modelling outputs to drive qualitative content analysis and vice versa–but it requires comfort in multiple epistemic cultures. The methodology itself can become part of the resistance to disinformation. If those who spread disinformation weaponize secrecy and obfuscation, researchers can counter it with transparency and reproducibility. Still, they must remain aware of new counter-tactics: propagandists could try to game detection algorithms (e.g., by varying their language to avoid repetition that algorithms catch). This means the data science tools for accountability must continuously evolve, possibly incorporating machine learning to detect coordination at deeper levels (e.g., semantic similarities, network behaviour patterns) rather than just surface text features.

Limitations and Future Research

No study is without limitations, and it is vital to recognize where the analysis of Metric Media might fall short and how future research can build on it.

First, the sentiment analysis tool (TextBlob) provides a constrained view of emotional tone. As noted, it may misclassify sarcasm, satire, or context-dependent language, potentially underestimating the presence of cynical or coded negative sentiment in Metric Media's ostensibly "neutral" articles. Future studies should leverage more nuanced AI language models–for example, transformer-based models like BERT or GPT–with deeper contextual understanding and can detect subtle linguistic cues of bias or persuasion. However, even advanced models should be used judiciously: researchers must validate their outputs against human judgment to ensure propaganda cues (like

dog-whistle phrases or loaded metaphors) are not overlooked. Incorporating cross-model approaches (combining multiple NLP techniques) might improve reliability in detecting the sly tactics of propaganda minimalism.

Second, this thesis focused on a single network (Metric Media) primarily within the United States. The generalizability of findings to other contexts remains an open question. Comparative research is a ripe next step. Scholars could examine similar pseudo-news networks in different countries–for instance, are there Metric Media equivalents in Europe, Asia, or Latin America exploiting local languages and issues? Early signs (Bradshaw & Howard, 2019) suggest that localized disinformation tactics are a global trend, but the specifics may vary by political culture. For example, a comparative study might find that the thematic clusters (like religion or education) differ in countries with different majority religions or social cleavages. It would also be enlightening to see if transnational linkages exist: do networks in different countries learn from or collaborate with each other? Given the influence of state actors like Russia and China in the information space, future work might explore whether foreign governments are covertly supporting "Metric Media-style" operations abroad to sway local populations.

Another limitation is the temporal scope. The data was a snapshot covering about 7,500 posts over a defined period. Propaganda strategies can adapt quickly, especially in response to exposure or countermeasures. The research may have captured Metric Media at a particular stage in its evolution. Longitudinal studies following the network's content over multiple election cycles or major events could reveal how it adjusts its tactics–for instance, does it increase negative partisanship in election years or shift the topic emphasis in response to current events (like a pandemic or social movement)? Tracking these changes over time would deepen understanding of the network's strategy and resilience. It would

also allow researchers to test causality more rigorously: do spikes in certain content precede shifts in public opinion in targeted communities? Do interventions (like a journalistic exposé about Metric Media) measurably impact the network's reach or strategy?

Financial opacity was a challenge that could not be fully unravelled. While this research cited reports of right-wing advocacy funding (Benjamin, 2022; Bengani 2020, 2021), a dedicated forensic investigation into funding streams–possibly by financial journalists or auditors–could greatly enrich the picture. Following the money might uncover connections to political campaigns, PACs, or government funds. Such work often requires access to internal documents or leaks, which was beyond the scope of our data-centric approach. Nonetheless, knowing who profits and who pays for Metric Media could validate the assumption of its ideological alignment and reveal potential pressure points (for instance, if major corporate donors are involved, they could be named and shamed).

Finally, regarding methodology and epistemology, an ongoing area for future research is refining the interdisciplinary toolkit for studying propaganda. This study approach combined computational analysis with theory, but there is room to formalize this integration. One could develop a propaganda detection algorithm grounded in the patterns identified here–for example, an algorithm that flags news sites with abnormally high neutral sentiment combined with particular topic frequencies and lack of original reporting. Testing such an algorithm on known cases (Metric Media, Sinclair content, state media republished pieces) versus genuine local journalism could yield a diagnostic tool for researchers and perhaps tech companies. Moreover, future research should engage with the normative implications more deeply: as we create tools to identify propaganda, we must also wrestle with questions of censorship, free speech, and the delineation between "persuasion" and "propaganda." Interdisciplinary teams, including ethicists, legal scholars,

and computer scientists, will be needed to ensure that interventions do not overreach or infringe on legitimate discourse.

The outstanding challenges underscore that journalism and democracy are at a crossroads in the face of such covert propaganda. If pseudo-local networks continue to grow unabated, public trust in information may erode even further, and communities might find themselves enclosed in carefully manicured realities. The onus is on researchers, supported by vigilant citizens and responsive policymakers, to keep innovating in exposing and countering these operations. Each limitation identified is a roadmap: future scholarship can build a more robust defence of the truth in the digital public square by addressing it.

6. Conclusion

Metric Media's pseudo-local news network is a bellwether of the new propaganda paradigm confronting democracies. This thesis set out to dissect how a nominally local journalism operation could, in fact, function as a sophisticated instrument of micro-targeted disinformation. Through a combination of computational analysis and critical theory, the study provides a comprehensive answer to that puzzle, illuminating the mechanisms by which Metric Media manufactures legitimacy, shapes narratives, and exploits the vulnerabilities of the digital media ecosystem. In explicitly addressing the research questions, the study demonstrated that Metric Media constructs its news ecosystem by scaling out a vast network of automated local-lookalike sites (RQ1), each carefully curated to emphasize themes and sentiments that align with a covert partisan agenda (RQ2). The study showed that the network's power is magnified by modern information infrastructure: social media algorithms and targeted distribution funnel its content to receptive audiences, heightening its influence without overt visibility. It also identified concrete avenues for intervention, suggesting that what has been used to mislead can be turned around to inform–the same data-driven methods can help journalists and citizens detect and counteract pseudo-news incursions (RQ3 & RQ4).

These findings' significance lies in unpacking a single network and illustrating a broader societal challenge. Metric Media represents a convergence of trends–the collapse of traditional local news, the rise of platform-mediated information, and the weaponization of journalistic symbols–to create a perfect storm for democracy. It exploits the trust people place in "local news" and the reach of big tech to even the most minor outlets. In doing so, it has redefined the battleground of public opinion: no longer confined to national debates or

overt propaganda outlets, the contest for truth has moved into the neighbourhood, onto Facebook feeds and community forums. Under a cloak of neutrality, this micro-targeting of ideology is a wake-up call that the defence of an informed citizenry must also operate at the micro level. Communities can no longer take for granted that the friendly local news source is genuinely local or even news–vigilance is needed at all levels of media consumption.

Theoretically, this work contributes to an evolving understanding of propaganda in the digital era. It extends classic frameworks by showing that agenda-setting can now be audience-specific and invisible, that Herman and Chomsky's propaganda filters have mutated into algorithmic forms, and that Zuboff's concept of surveillance capitalism has political as well as commercial ramifications–data is not just used to sell products but to sell ideologies. The study introduced the concepts of propaganda minimalism and credibility laundering to the scholarly lexicon, offering language to describe how modern propagandists succeed by understatement and impersonation rather than overstatement. These concepts sharpen scholars' critical toolkit for analyzing information warfare that is at once subtle and far-reaching. Propaganda minimalism captures the essence of Metric Media's strategy: influence through restraint and repetition, while credibility laundering encapsulates its masquerade method. Both have applicability beyond this case, and future studies might refine these ideas or apply them to other domains (for example, the spread of health misinformation often follows a similar pattern of pseudo-credible sources and minimalistic claims).

Practically, this thesis underscores an urgent need for innovation in safeguarding the information commons. The recommendations, from transparency mandates to media literacy to algorithmic adjustments, are starting points for policy deliberation and civic action. Implementing them will require overcoming political polarization (since any attempt

to curtail a propaganda network may be framed as partisan) and engaging a public that is often unaware of the scope of the problem. However, this study clarified the cost of inaction: allowing Metric Media and its clones to proliferate unchecked will further erode the boundary between reality and manipulation in communities nationwide in the U.S. The stakes are exemplified whenever a significant share of citizens base their voting decisions or civic perceptions on filtered information these shadow media provide. In such a scenario, public discourse and democratic decision-making become, in effect, privatized and hijacked by those who control the false "local" narrative.

In closing, the story of Metric Media is not just about one network's deception–it is about the resilience of democratic societies in the face of new forms of falsehood. This thesis has shone a light on a dark corner of the media landscape, revealing how easily the symbols of community journalism can be repurposed for propaganda. The findings serve as a stark reminder that democracy's safeguards must evolve in step with technology. Just as propaganda has taken a new shape in the 21st century, so must strategies for promoting truth and transparency. If there is a silver lining, it is that exposing the machinery of pseudo-news is the first step toward disarming it. By naming and understanding tactics like credibility laundering, policymakers, platforms, and the public are equipped with the terminology and insight to confront them head-on.

The significance of this research is captured in a simple but powerful idea: information integrity is the bedrock of democracy. When that integrity is subverted at the local level, the foundation of the entire democratic project is at risk. Knowledge is power, and citizens can reclaim some of that power by learning exactly how Metric Media and networks like them operate. This research affirms that an informed citizenry–armed with both the facts of the matter and an understanding of the forces at play–can indeed mount a

defence against even the most insidious propaganda. In the battle for hearts and minds now being waged through hometown headlines and Facebook feeds, shining a light on the truth is the most effective weapon. Democracy does not require a perfectly discerning public, but it does require that we not let deception run rampant under false pretenses. This thesis, in its deep dive into Metric Media, ultimately speaks to a larger imperative: to ensure that the news that binds our communities is not what tears our democracy apart. The fight for factual, accountable local journalism is nothing less than a fight for the soul of an informed society–a fight we cannot afford to lose.

7. Works Cited

Abernathy, P. M. (2018). *The expanding news desert*. Center for Innovation and Sustainability in Local Media, School of Media and ….

Adjin-Tettey, T. (2022). Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education. *Cogent Arts & Humanities*, *9*, 2037229. https://doi.org/10.1080/23311983.2022.2037229

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

American Press Institute. (2013, October 9). *The elements of journalism*. https://www.americanpressinstitute.org/journalism-essentials/what-is-journalism/elements-journalism/

Auxier, B., & Anderson, M. (2021, April 7). Social Media Use in 2021. *Pew Research Center*. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/

Bakir, V. (2020). Psychological Operations in Digital Political Campaigns: Assessing Cambridge Analytica's Psychographic Profiling and Targeting. *Frontiers in Communication*, *5*. https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2020.00067

Bakir, V., Herring, E., Miller, D., & Robinson, P. (2018). Lying and deception in politics. In J. Meibauer (Ed.), *The Oxford Handbook of Lying* (pp. 529–540). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198736578.013.42

Bakir, V., & McStay, A. (2017). Fake News and The Economy of Emotions: Problems, causes, solutions. *Digital Journalism*, *6*, 1–22. https://doi.org/10.1080/21670811.2017.1345645

Bartholomew, J. (2022, September 19). The rise and rise of partisan local newsrooms. *Columbia Journalism Review*. https://www.cjr.org/tow_center/the-rise-of-partisan-local-newsrooms.php

Bell, E. (2016, March 7). The End of the News as We Know It: How Facebook Swallowed Journalism. *Medium*. https://medium.com/tow-center/the-end-of-the-news-as-we-know-it-how-facebook-swallowed-journalism-60344fa50962

Bengani, P. (2019, December 18). Hundreds of 'pink slime' local news outlets are distributing algorithmic stories and conservative talking points. *Columbia Journalism Review*. https://www.cjr.org/tow_center_reports/hundreds-of-pink-slime-local-news-outlets-are-distributing-algorithmic-stories-conservative-talking-points.php

Bengani, P. (2020, August 4). As election looms, a network of mysterious 'pink slime' local news outlets nearly triples in size. *Columbia Journalism Review*. https://www.cjr.org/analysis/as-election-looms-a-network-of-mysterious-pink-slime-local-news-outlets-nearly-triples-in-size.php

Bengani, P. (2021a, October 14). Advocacy groups and Metric Media collaborate on local 'community news.' *Columbia Journalism Review*.

https://www.cjr.org/tow_center_reports/community-newsmaker-metric-media-local-news.php

Bengani, P. (2021b, October 14). The Metric Media network runs more than 1,200 local news sites. Here are some of the non-profits funding them. *Columbia Journalism Review*. https://www.cjr.org/tow_center_reports/metric-media-lobbyists-funding.php

Bengani, P. (2024, August 8). The nonprofits and PACs that spent $14 million on the Metric Media network in 2021–22. *Columbia Journalism Review*. https://www.cjr.org/tow_center/the-non-profits-and-pacs-that-spent-14-million-on-the-metric-media-network-in-2021-22.php

Benjamin, J. (2022, December 5). Fake local news sites on pace to outnumber local US newspapers. *The Media Leader*. https://uk.themedialeader.com/fake-local-news-sites-on-pace-to-outnumber-local-us-newspapers/#:~:text=Metric%20Media%20group%2C%20founded%20in,Mobile%20Courant%2C%20and%20Suffolk%20Reporter

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press. https://doi.org/10.1093/oso/9780190923624.001.0001

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Berger, E. (2024, July 2). TV giant known for rightwing disinformation doubles down on its national news agenda. *The Guardian*.

https://www.theguardian.com/us-news/article/2024/jul/02/sinclair-tv-disinformation-conservative-news#:~:text=Sinclair%2C%20one%20of%20the%20largest,Joe%20Biden%2C%20media%20analysts%20say

Bernays, E. L. (1947). The Engineering of Consent. *The ANNALS of the American Academy of Political and Social Science*, *250*(1), 113–120. https://doi.org/10.1177/000271624725000116

Bisong, E. (2019). *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*. https://doi.org/10.1007/978-1-4842-4470-8

Blei, D. M. (2012). Probabilistic topic models. *Commun. ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Bogart, L. (1995). *Commercial Culture: The Media System and the Public Interest*. https://api.semanticscholar.org/CorpusID:153332687

Bond, S. (2024, September 26). Why Russia's broadcaster RT turned to covertly funding American pro-Trump influencers. *NPR*. https://www.npr.org/2024/09/25/nx-s1-5114246/russia-today-rt-election-influence-tenet-indictment#:~:text=RT%20was%20long%20known%20to,cooled%2C%20skepticism%20of%20RT%20grew

Borgesius, F. J. Z., Trilling, D., Möller, J., Bodó, B., Vreese, C. H. de, & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, *5*(1), 1–16. https://doi.org/10.14763/2016.1.401

Bradshaw, S., & Howard, P. (2019). *The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation*.

Brady, A.-M. (2009). *Marketing dictatorship: Propaganda and thought work in contemporary China*. Rowman & Littlefield Publishers.

Carlson, M. (2020). Journalistic critical incidents as boundary making and the making of boundaries around critical incidents. In *Critical Incidents in Journalism* (pp. 28–42). Routledge.

Chadwick, A. (2013). *The Hybrid Media System: Politics and Power*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199759477.001.0001

Chaffey, D., & Smith, P. (2022). *Digital Marketing Excellence: Planning, Optimizing and Integrating Online Marketing*. https://doi.org/10.4324/9781003009498

Chakravartty, P., Kuo, R., Grubbs, V., & McIlwain, C. (2018). #CommunicationSoWhite. *Journal of Communication*, *68*(2), 254–266. https://doi.org/10.1093/joc/jqy003

Chyi, H. (2012). Paying for What? How Much? And Why (Not)? Predictors of Paying Intent for Multiplatform Newspapers. *The International Journal on Media Management*, *14*, 227–250. https://doi.org/10.1080/14241277.2012.657284

Clegg, J., & Usmani, A. (2017). The Racial Politics of Mass Incarceration. In *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3025670

Coleman, P. (1989). *The Liberal Conspiracy: The Congress for Cultural Freedom and the Struggle for the Mind of Postwar Europe*. Free Press. https://books.google.ca/books?id=yaorAAAAMAAJ

Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political Polarization on Twitter. In *Fifth International AAAI Conference on Weblogs and Social Media*.

Cook, S. (2013). *The Long Shadow of Chinese Censorship: How the Communist Party's Media Restrictions Affect News Outlets Around the World*. Center for International Media Assistance.

Correa, T., Hinsley, A. W., & de Zúñiga, H. G. (2010). Who interacts on the Web?: The intersection of users' personality and social media use. *Computers in Human Behavior*, *26*(2), 247–253. https://doi.org/10.1016/j.chb.2009.09.003

Corsi, G. (2024a). Evaluating Twitter's algorithmic amplification of low-credibility content: An observational study. *EPJ Data Science*, *13*(1), 18. https://doi.org/10.1140/epjds/s13688-024-00456-3

Corsi, G. (2024b). Evaluating Twitter's algorithmic amplification of low-credibility content: An observational study. *EPJ Data Science*, *13*(1), 18. https://doi.org/10.1140/epjds/s13688-024-00456-3

Cottle, S. (2006). *Mediatized Conflict: Developments in Media and Conflict Studies*.

Couldry, N., & Mejias, U. (2019). *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. https://doi.org/10.1515/9781503609754

Daley, D. (2024, October 4). Two men have re-engineered the US electoral system in favor of Republicans. *The Guardian*. https://www.theguardian.com/commentisfree/2024/oct/04/electoral-college-map-gerrymandering

Darr, J. (2024). How Sticky Is Pink Slime? Assessing the Credibility of Deceptive Local Media. *The ANNALS of the American Academy of Political and Social Science*, *707*, 109–124. https://doi.org/10.1177/00027162231214696

Dionne, E. J., & Keeney, M. (2025, January 2). What the nation told us in 2024, state by state. *Brookings Institute*. https://www.brookings.edu/articles/what-the-nation-told-us-in-2024-state-by-state/

Egelhofer, J. L., & and Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association*, *43*(2), 97–116. https://doi.org/10.1080/23808985.2019.1602782

Ellul, J. (1965). *Propaganda: The Formation of Men's Attitudes*. Knopf Doubleday Publishing Group. https://books.google.ca/books?id=yKfZAAAAMAAJ

Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, *43*(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Faverio, M., & Sidoti, O. (2024, December 12). Teens, Social Media and Technology 2024. *Pew Research Center*. https://www.pewresearch.org/internet/2024/12/12/teens-social-media-and-technology-2024/

Freedman, D. (2014). *The contradictions of media power*. Bloomsbury Publishing.

French, L., & Manzanarez, M. (2004). *NAFTA & neocolonialism: Comparative criminal, human, & social justice*. University Press of America.

Fuchs, C. (2018). *Digital demagogue: Authoritarian capitalism in the age of Trump and Twitter*. Pluto Press.

Gillespie, T. (2018). Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. In *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. https://doi.org/10.12987/9780300235029

Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, *22*(6), 854–871. https://doi.org/10.1080/1369118X.2019.1573914

Gorwa, R. (2021). *Platform governance: The transnational politics of online content regulation* [PhD Thesis]. University of Oxford.

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, *7*(1), 2053951719897945. https://doi.org/10.1177/2053951719897945

Gotfredsen, S. G. (2023, February 8). Journalists Remain on Twitter, but Tweet Slightly Less. *Columbia Journalism Review*. https://www.cjr.org/tow_center/journalists-remain-on-twitter-but-tweet-slightly-less.php

Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, *117*(27), 15536–15545. https://doi.org/10.1073/pnas.1920498117

Guess, A., Nyhan, B., & Reifler, J. (n.d.). *Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign*. http://www.dartmouth.edu/~nyhan/fake-news-2016.pdf

Guo, M., Zhou, J., Tang, F., & Shen, Y. (2016). *Pervasive computing: Concepts, technologies and applications*. CRC Press.

Harder, R., Sevenans, J., & Aelst, P. (2017). Intermedia Agenda Setting in the Social Media Age: How Traditional Players Dominate the News Agenda in Election Times. *The International Journal of Press/Politics*, *22*, 194016121770496. https://doi.org/10.1177/1940161217704969

Herman, E. S., & Chomsky, N. (1988). *Manufacturing Consent: The Political Economy of the Mass Media*. Pantheon Books. https://books.google.ca/books?id=Up5sAAAAIAAJ

Hernandez, J. (2024, August 29). How swing states came to be critical in U.S. presidential elections. *NPR*. https://www.npr.org/2024/08/29/nx-s1-4936052/swing-states-presidential-elections

Hobbs, R. (2011). *Digital and media literacy: Connecting culture and classroom*. (pp. xi, 214). Corwin Press.

Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). *spaCy: Industrial-strength Natural Language Processing in Python*. https://doi.org/10.5281/zenodo.1212303

Hopkins, D. J. (2018). *The increasingly United States: How and why American political behavior nationalized*. University of Chicago Press.

Humans of New York. (2024, January 5). [Facebook Page]. *Facebook*. https://www.facebook.com/humansofnewyork/posts/pfbid0rEiWWasMjEJ4gnTfsqmq84Bh4Tqtr1PvgoDWtauLs3x9QGEW1LhcSxn8P1m7qgurl

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Award Winning Papers from the 19th International Conference on Pattern Recognition (ICPR)*, *31*(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

Jick, T. D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly*, *24*(4), 602–611. JSTOR. https://doi.org/10.2307/2392366

Jowett, G. S., & O'Donnell, V. (2012). *Propaganda & Persuasion* (5th ed.). SAGE Publications.

Kellner, D. (1992). *The persian gulf TV war*. Routledge.

Kilgour, F. G. (1998). *The evolution of the book* (Vol. 1–1 online resource (179 pages)). Oxford University Press; WorldCat. http://site.ebrary.com/id/10358318

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 2053951714528481. https://doi.org/10.1177/2053951714528481

Kreiss, D. (2016). *Prototype Politics: Technology-Intensive Campaigning and the Data of Democracy*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199350247.001.0001

Kretchmar, K., Sondel ,Beth, & and Ferrare, J. J. (2014). Mapping the terrain: Teach For America, charter school reform, and corporate sponsorship. *Journal of Education Policy*, *29*(6), 742–759. https://doi.org/10.1080/02680939.2014.880812

Kumar, S., & Shah, N. (2018). *False Information on Web and Social Media: A Survey*. https://arxiv.org/abs/1804.08559

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Lashmar, P., & Oliver, J. E. (1998). *Britain's secret propaganda war*. https://api.semanticscholar.org/CorpusID:156693186

Lasswell, H. D. (1927). The Theory of Political Propaganda. *American Political Science Review*, *21*(3), 627–631. https://doi.org/10.2307/1945515

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Lewis, S. C., & Usher, N. (2013). Open source and journalism: Toward new frameworks for imagining news innovation. *Media, Culture & Society*, *35*(5), 602–619. https://doi.org/10.1177/0163443713485494

Liu, B. (2022). *Sentiment analysis and opinion mining*. Springer Nature.

MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations*. https://api.semanticscholar.org/CorpusID:6278891

Marlin, R. (2013). *Propaganda and the Ethics of Persuasion—Second Edition*. Broadview Press. https://books.google.ca/books?id=qDUmAgAAQBAJ

Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online. *New York: Data & Society Research Institute*, *359*, 1146–1151.

MassMove. (2024). *MassMove—Attack Vectors* [Dataset]. GitHub. https://github.com/MassMove/AttackVectors

McChesney, R. D. (2004). *The problem of the media: US communication politics in the twenty-first century*. NYU Press.

McChesney, R. W., & Pickard, V. (2011). *Will the last reporter please turn out the lights: The collapse of journalism and what can be done to fix it*. New Press, The.

Mccombs, M. (1997). Building Consensus: The News Media's Agenda-Setting Roles. *Political Communication - POLIT COMMUN*, *14*, 433–443. https://doi.org/10.1080/105846097199236

McCombs, M. (2005). A Look at Agenda-setting: Past, present and future. *Journalism Studies*, *6*(4), 543–557. https://doi.org/10.1080/14616700500250438

Mccombs, M. (2008). *Setting the agenda: The mass media and public opinion*.

McCombs, M. E., & Shaw, D. L. (1972). The Agenda-Setting Function of Mass Media. *The Public Opinion Quarterly*, *36*(2), 176–187. JSTOR.

McCombs, M., & Reynolds, A. (2002). News influence on our pictures of the world. In *Media effects* (pp. 11–28). Routledge.

McKinney, W. (2010). *Data Structures for Statistical Computing in Python* (p. 61). https://doi.org/10.25080/Majora-92bf1922-00a

Meraz, S. (2009). Is There an Elite Hold? Traditional Media to Social Media Agenda Setting Influence in Blog Networks. *Journal of Computer-Mediated Communication*, *14*(3), 682–707. https://doi.org/10.1111/j.1083-6101.2009.01458.x

Meyer, E. T., Shankar, K., Willis, M., Sharma, S., & Sawyer, S. (2019). The social informatics of knowledge. *Journal of the Association for Information Science and Technology*, *70*(4), 307–312. https://doi.org/10.1002/asi.24205

Nadeau, D., & Sekine, S. (2007). A Survey of Named Entity Recognition and Classification. *Lingvisticae Investigationes*, *30*. https://doi.org/10.1075/li.30.1.03nad

Napoli, P. (2019). *Social media and the public interest: Media regulation in the disinformation age*. Columbia university press.

Napoli, P., & Caplan, R. (2017). Why media companies insist they're not media companies, why they're wrong, and why it matters. *First Monday*, *22*(5). https://doi.org/10.5210/fm.v22i5.7051

Nazarudin, M., & Handayani, W. (2024). Optimizing Social Media as a Global Marketing Tool: Qualitative Analysis of Different Approaches in Enhancing the Image and Sales of Exported Products. *Business and Investment Review*, *2*, 23–29. https://doi.org/10.61292/birev.89

O'Reilly, T. (2007). What Is Web 2.0: Design Patterns And Business Models For The Next Generation Of Software. *University Library of Munich, Germany, MPRA Paper*, *65*.

Pariser, E. (2011). *The Filter Bubble: What The Internet Is Hiding From You*. Penguin Books Limited. https://books.google.ca/books?id=-FWO0puw3nYC

Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived

Accuracy of Headlines Without Warnings. *Management Science*, *66*(11), 4944–4957. https://doi.org/10.1287/mnsc.2019.3478

Petre, C. (2018). Engineering Consent: How the Design and Marketing of Newsroom Analytics Tools Rationalize Journalists' Labor. *Digital Journalism*, *6*(4), 509–527. https://doi.org/10.1080/21670811.2018.1444998

Pickard, V. (2020). *Democracy without Journalism?: Confronting the Misinformation Society*. Oxford University Press. https://doi.org/10.1093/oso/9780190946753.001.0001

Ramos, J. (2003). *Using TF-IDF to determine word relevance in document queries*.

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and Amending Directive 2000/31/EC (Digital Services Act) (Text with EEA Relevance) (2022). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065

Repnikova, M. (2017). *Media Politics in China: Improvising Power under Authoritarianism*. Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781108164474

Roberts, S. T. (2019). *Behind the Screen: Content Moderation in the Shadows of Social Media*. Yale University Press; JSTOR. https://doi.org/10.2307/j.ctvhrcz0v

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Scheufele, D. A., & Tewksbury, D. (2007). Framing, agenda setting, and priming: The evolution of three media effects models. *Journal of Communication*, *57*(1), 9–20.

Shu, K., Wang, S., & Liu, H. (2017). *Exploiting Tri-Relationship for Fake News Detection*. https://doi.org/10.48550/arXiv.1712.07709

Sides, J., Tesler, M., & Vavreck, L. (2018). *Identity Crisis*. Princeton University Press; JSTOR. https://doi.org/10.2307/j.ctvc77mmb

Smith, C. (2015). *Social media and the destruction of world heritage as global propaganda: Inaugural lecture*.

Spring, J. H. (2014). *Political Agendas for Education: From Race to the Top to Saving the Planet*. Routledge. https://books.google.ca/books?id=2xRMzwEACAAJ

Sproule, J. M. (1997). *Propaganda and democracy: The American experience of media and mass persuasion*. Cambridge University Press; WorldCat.

Stokes, L. C. (2020). *Short Circuiting Policy: Interest Groups and the Battle Over Clean Energy and Climate Policy in the American States*. Oxford University Press. https://doi.org/10.1093/oso/9780190074258.001.0001

Sun, H. (2023). Regulating Algorithmic Disinformation. *The Columbia Journal of Law & the Arts*, *46*(4), 367–417. https://doi.org/10.52214/jla.v46i3.11237

Sunstein, C. R. (2018). *Republic: Divided democracy in the age of social media*.

Tandoc, E., Lim, Z., & Ling, R. (2017). Defining "Fake News": A typology of scholarly definitions. *Digital Journalism*, *6*, 1–17. https://doi.org/10.1080/21670811.2017.1360143

Taylor, P. M. (2003). *Munitions of the Mind: A History of Propaganda, Third Edition*. Manchester University Press. https://books.google.ca/books?id=_KS3fBvjkicC

Townsend, L., & Wallace, C. (2017). The ethics of using social media data in research: A new framework. In *The ethics of online research* (pp. 189–207). Emerald Publishing Limited.

Tufekci, Z. (2015). Facebook Said Its Algorithms Do Help Form Echo Chambers, and the Tech Press Missed It. *New Perspectives Quarterly*, *32*(3), 9–12. https://doi.org/10.1111/npqu.11519

Tufekci, Z. (2017). *Twitter and tear gas: The power and fragility of networked protest*. Yale University Press.

UNESCO. (2022). *Journalism is a public good: World trends in freedom of expression and media development; Global report 2021/2022*. https://unesdoc.unesco.org/ark:/48223/pf0000380618.locale=en

Vaidhyanathan, S. (2018). *Antisocial media: How Facebook disconnects us and undermines democracy*. Oxford University Press.

Vargo, C. J., & Guo, L. (2017). Networks, big data, and intermedia agenda setting: An analysis of traditional, partisan, and emerging online US news. *Journalism & Mass Communication Quarterly*, *94*(4), 1031–1055.

Vargo, C. J., Guo, L., McCombs, M., & Shaw, D. L. (2014). Network Issue Agendas on Twitter During the 2012 U.S. Presidential Election. *Journal of Communication*, *64*(2), 296–316. https://doi.org/10.1111/jcom.12089

Vliegenthart, R., & Walgrave, S. (2008). The Contingency of Intermedia Agenda Setting: A Longitudinal Study in Belgium. *Journalism & Mass Communication Quarterly*, *85*(4), 860–877. https://doi.org/10.1177/107769900808500409

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Walker, C., & Ludwig, J. (2017). The meaning of sharp power: How authoritarian states project influence. *Foreign Affairs*, *16*(11).

Wanta, W., & Hu, Y.-W. (1994). The Effects of Credibility, Reliance, and Exposure on Media Agenda-Setting: A Path Analysis Model. *Journalism Quarterly*, *71*(1), 90–98. https://doi.org/10.1177/107769909407100109

Wardle, C. (2018). The Need for Smarter Definitions and Practical, Timely Empirical Research on Information Disorder. *Digital Journalism*, *6*, 951–963. https://doi.org/10.1080/21670811.2018.1502047

Wardle, C., & Derakhshan, H. (2017). *INFORMATION DISORDER : Toward an interdisciplinary framework for research and policy making Information Disorder Toward an interdisciplinary framework for research and policymaking*.

Welch, D. (2014). *Propaganda, power and persuasion: From World War I to Wikileaks*. I.B. Tauris; WorldCat.

Whitehead, A. L., Perry, S. L., & Baker, J. O. (2018). Make America Christian Again: Christian Nationalism and Voting for Donald Trump in the 2016 Presidential Election. *Sociology of Religion*, *79*(2), 147–171. https://doi.org/10.1093/socrel/srx070

Yin, R. K., Campbell, D. T., & Sage Publications. (2018). *Case study research and applications: Design and methods* (Sixth edition). SAGE Publications, Inc.; WorldCat.

Zollmann, F. (2017). *Media, Propaganda and the Politics of Intervention*. Peter Lang Verlag. https://doi.org/10.3726/b11346

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

Zúñiga, K. (2023). *Facebook Scraper* [Dataset]. GitHub. https://github.com/kevinzg/facebook-scraper