

JOÃO CARLOS CORREIA
PEDRO JERÓNIMO
EDS / INÊS AMARAL

DISIN FOR MATION STUDIES



LABCOM
COMMUNICATION
& ARTS

DISINFORMATION STUDIES

PERSPECTIVES FROM AN EMERGING FIELD

JOÃO CARLOS CORREIA

PEDRO JERÓNIMO

INÊS AMARAL

/EDS



LABCOM
COMMUNICATION
& ARTS

**Technical
Specification**

Title

Disinformation Studies:
Perspectives from An Emerging Field

Editors

João Carlos Correia
Pedro Jerónimo
Inês Amaral

LabCom Books

www.labcom.ubi.pt

Collection

Communication Books

Direction

Gisela Gonçalves

Graphic Design

Cristina Lopes

ISBN

978-989-654-861-2 (print)

978-989-654-863-6 (pdf)

978-989-654-862-9 (epub)

Legal Deposit

507398/22

Print

Print-on-demand

Universidade da Beira Interior
Rua Marquês D'Ávila e Bolama.
6201-001 Covilhã. Portugal
www.ubi.pt

Covilhã, 2022



© 2022, João Carlos Correia, Pedro Jerónimo & Inês Amaral

© 2022, Beira Interior University

*Publishing copyright authorizations from both articles and images
are exclusively the author's responsibility.*

Contents

Preface	9
Abstracts	13
Chapter 1 – Capturing and dissecting the complexity of production and dissemination of conspiracy theories, hate-based rhetoric, and mis- and disinformation online	23
Arđian Shajkovci, Ramón Ruti, Asli Altinbay, Matteo Gregori, Amanda Garry & Allison Mcdowell-Smith	
Chapter 2 – The twilight zone: case studies in misinformation & mass media	53
Alyssa M. Brumis	
Chapter 3 – When communication meets international relations perspectives: understanding disinformation in a multicentric political environment	87
Alessandra Massa & Giuseppe Anzera	
Chapter 4 – Fake news, post-truth, and journalism: weaknesses and strategies in 2018 brazilian elections	115
Luísa Guimarães Torre	
Chapter 5 – Opinion-oriented news as a source of polarized disinformation on the EU: a case study analysis during the 2019 EP elections	151
Rubén Rivas-De-Roca & Mar García Gordillo	
Chapter 6 – Lies are all around but who are the liars? Addressing online disinformation platforms in the Czech Republic and Slovakia	171
Ján Hacek & Lucia Virostková	
Chapter 7 – Technologies and fact-checking: a sociotechnical mapping	193
Oscar Westlund, Rebekah Larsen, Lucas Graves, Lasha Kavtaradze & Steen Steensen	

Chapter 8 – The role of fact-checking in fighting the ‘infodemic’ of disinformation on Covid-19: a case study of Polígrafo	237
Marina Ferreira & Inês Amaral	
Chapter 9 – Framing Covid-19: how fact-checking circulate on the Facebook far-right	265
Raquel Recuero, Taiane Volcan, Felipe Soares, Otávio Vinhas & Luiz Ricardo Huttner	
Authors	289

Preface

Disinformation is a keyword of our societies. Would it be so? Rephrasing the sentence, one could say that disinformation in all his extension it is a keyword and a condition of all human societies. According to Kahneman, Sibony and Sunstein (2021)¹, wherever there is human judgment there is noise. Opinions are not an evil source of error. Only in an ideal world, in a situation of ideal speech and communication each message would works as a kind of magic bullet, and the sender as someone who shares a rifle, firing a single shot in the bull's eye (Kahneman et al, 2021, p.3).

Communication needs facts but also needs an effort of interpretative hermeneutics. Journalism needs hard news oriented to the facts but also needs good context, thoughtful editorials, and efficient storytelling that help to give cultural meaning to chaotic experiences of everyday life.

The journalistic activity is action-oriented towards the construction of social reality, which is objectified through everyday practices of representation of what happened.

From this perspective, social construction, in the media research sector, is the production of meaning through the action of productive practices and routines that organize the journalistic profession. The news making and statements produced in its realization are not limited to reproducing reality but intervene in the social construction of it. In this perspective, more than simple mirrors of a pre-existing reality, journalists and,

1. Kahneman, Daniel; Sibony, Oliver & Sunstein, Cass (2021). *Noise*. London, Harper Collins.

consequently, the statements produced by them, intervene actively in the construction of the conditions and the way in which reality is perceived. Thus, “it is impossible to establish a radical distinction between reality and the media that must reflect that reality, because the news helps to build reality itself” (Traquina, 2001, p. 28)². The social and political world is not a predetermined and rigid reality at journalists reflect. Journalists don’t are passive observers but active participants in the construction of reality. Facts do not exist in themselves, endowed with evidence and self-sufficient ontological thickness that journalistic utterances would be pure reflexes. They are the product of an encounter between the facts and their reports, without which, indeed, they would not exist as journalistic facts. Reality is not something completely autonomous and distinct from the way the actors interpret it, internalize it, re-elaborate it, and redefine historically and culturally. The world reported in the news is the result of categorization activities and not a simple act of naming reality as if it were ready to be classified. The journalistic report is not an act of describing or saying in a direct, determined, and precise way an empirical fact that happened in the outside world, but it is an act of presenting a reality that is constituted even with the active participation of the reader.

However, in spite of those well-known assumptions from the epistemology and sociology well established since Lippmann to Tuchman, Saperas, Traquina, Singer , and many other scholars, reality has some level of resistance. That’s to say one may discuss many interpretations about the Ukrania – Russia conflict but that’s difficult to say that The Ukrainian Army Invaded Moscow.

Information /Misinformation has many levels

One level of disinformation has to do with facts and appears mostly related with the deliberated twist of numbers, nature of things, and the furious denial of some common sense and taken-for-granted social facts, or even lies.

2. Traquina, Nelson, (2001), “Teorias das Notícias: o estudo do jornalismo no século xx”. In Nelson Traquina, Ana Cabrera, Cristina Ponte & Rogério Santos, O Jornalismo Português em Análise de Casos, Lisboa, Caminho.

It is boosted by the emotional factors and by the well-known affection law: people exposed to, understand better, memorize deeply, and share heavily, message, with they agree to, disregarding the fact of being true or false, depending mostly on the gratification they obtain with that message. It is a kind of information that are articulated with the so-called fake news as a particular repertoire of propaganda in my point of view.

Another level of misinformation has to do with ideological frames, media rhetoric, interpretation and context and it happens in all kinds of media.

It's because of that that I find this book very important.

We thank all authors, and we reassure our commitment to the research on this important issue.

João Carlos Correia

Chapter 1. Capturing and Dissecting The Complexity of Production and Dissemination of Conspiracy Theories, Hate-Based Rhetoric, and Mis-and Disinformation Online

Ardian Shajkovci, Ramón Ruti, Asli Altinbay, Matteo Gregori, Amanda Garry & Allison McDowell-Smith

Fringe independent websites and alternative social media platforms such as Telegram, Parler, and Bitchute have increasingly become major spreaders of mis-and disinformation, conspiracy theories, and hate-based rhetoric. These spaces have also become echo chambers that can largely influence online users seeking propaganda related to specific movements. The purpose of this chapter is to facilitate an initial discussion on the role of smaller and alternative social media platforms, as part of a wider media ecosystem engaged in constructing and communicating conspiracy theories, hate-based rhetoric, and mis-and disinformation. The authors will discuss and introduce innovative tools and methodologies in tracking and analyzing content online, as demonstrated through the case of QAnon conspiracy theory. Furthermore, given that the current political, popular and media conversation about regulating social media primarily focuses upon the role of Twitter, Facebook and to a lesser extent YouTube, the arguments presented by the authors herein suggest that such line of contention may be somewhat misplaced and misleading.

Keywords: Mis-and disinformation, alternative platforms, QAnon, social media, Saas Tools

Chapter 2. The Twilight Zone: Case Studies in Misinformation & Mass Media

Alyssa M. Brumis

Misinformation and disinformation have arguably always existed in American society. However, over the last three decades, with advancements across mass communication, the spread of misinformation has become much more prevalent. As the landscape of mass media advances, so too, does the problem of misinformation. Misinformation spreads across mass media platforms, such as mainstream news media and social media, reaching larger audiences than ever before. The ways in which we consume (mis) information and the speed at which we do so has radically changed. The effects of misinformation leave us relatively defenseless, due to our human inclination to engage in motivated reasoning, or emotion-based reasoning, and due to phenomenon such as the illusory truth effect, or our proclivity to believe false information after repeated exposure. As a surplus of misinformation is disseminated across mass media, we move closer to a “post-truth” society, which comes with dangerous consequences. This collection of case studies explores the disseminators of disinformation across mass media and the effects of mis/disinformation, illustrating its deadly consequences in America. The overarching theme across each case is the notion that mis/disinformation disrupts our connection to facts, logic, and reality, leaving us incapable of making rational and informed decisions. Utilizing interdisciplinary theories from communication studies and the social sciences, I examine the top disseminators of mis/disinformation across three key cases: a 2015 disinformation campaign waged against Planned Parenthood, disseminated across mainstream news media; 2020 election-related misinformation disseminated on Facebook, in then-President Donald Trump’s political advertisements; and 2020/2021 coronavirus misinformation disseminated by Fox News and its pundits. These cases expose mis/disinformation as a threat to public safety, a threat to democracy, and a threat to our public health. When we live in a society that spreads misinformation

like wildfire, we can quickly begin living in what resembles an episode of The Twilight Zone.

Keywords: misinformation in mass media, disinformation disseminator, consequences of mis/disinformation

Chapter 3. When Communication Meets International Relations Perspectives: Understanding Disinformation in a Multicentric Political Environment

Alessandra Massa & Giuseppe Anzera

This chapter explores how interpretive concepts from the analytical tradition of International Relations (IR) may help explain disinformation in current international politics. As IR interpretative tradition has evolved in recent decades, it has praised openings to practices beyond the traditional power politics focused on tangibles or material resources. Communication and information are becoming competitive assets that states can use to gain citizens' loyalty, shape international public opinion, and position other actors in the international system. To investigate the link between contemporary disinformation (as a communicative phenomenon) and IR, we will analyze a variety of interpretive perspectives. In questioning IR, the multicentric system explains the setting and proliferation of individual skills, whereas complex interdependence enlightens the ongoing information revolution. The concept of soft power, then, describes the changes in non-tangible aspects of power today. Narratives and aesthetic turns are helpful to interpret international reality as a social construct, emphasizing the situated nature of political narratives. A realist approach is then used to explain why political elites deceive, resulting in trust issues affecting popular readings. IR analytical traditions and media studies must be considered in a mixed approach to explain individual and collective interpretations of the current disinformation climate, taking into account the multicentric nature of IR. In addition, narrative interpretations and constructivism highlight the

intersubjective process in global politics. In conclusion, realism suggests that lying is merely another tool for governing anarchy, regardless of its ethical implications.

Keywords: international relations; multicentrism; soft power; constructivism; international narratives.

Chapter 4. Fake news, post-truth, and journalism: weaknesses and strategies in 2018 Brazilian elections

Luísa Torre

In a post-industrial society in which information plays a fundamental role in social structures (Castells, 2012), rumors, the oldest media in the world (Kapferer, 1993), have come to occupy an important space in the public sphere. The rumors, now mediatized, are called fake news. Another phenomenon is added to this: the post-truth (Keyes, 2004), when the truth is closer to an individual's personal beliefs than to the facts.

Now that the enunciation poles have been multiplied (Pinto, 2000), the receiver is informed through innumerable channels, from news portals to social networks, and he himself decides on which to attribute faith (Primo, 2011). Journalism, which has lost centrality in the production, selection and distribution of information (Bentes, 2015), is facing new challenges - and fake news is a fundamental part of this scenario. In part because, instead of the evaluation of newspaper editors, recommendation algorithms used to create personalization are playing a central role in the selection and distribution of content, becoming new information gatekeepers (Heinderyckx and Vos, 2016).

In this society, organized in a network, power is also exercised in a network (Parente, 2004). If there is no power that can be exercised without discursive strategies (Castells, 2015), has this interaction between the issuers of fake news and the traditional news media become a dispute for power?

This work will analyze the speech contained in the denials produced by journalists from Grupo Globo, and in the (so classified) fake news debunked, published in the Fato or Fake section between August 31 and October 6, 2018 - between the beginning of electoral campaign and the first round of Brazilian elections. From the critical discourse analysis, we will investigate the discursive strategies used by fake news broadcasters and professional journalists in 8 articles, looking for elements that reveal if there is a discursive dispute between them.

Keywords: fake news, journalism, power, social media, critical discourse analysis

Chapter 5. Opinion-Oriented News as a Source of Polarized Disinformation on The EU: A Case Study Analysis During The 2019 EP Elections

Rubén Rivas-de-Roca & Mar García Gordillo

The role of truth for society is undermined in an accelerated digital era, since opinions are more important than facts in the shaping of the public opinion. Journalists seem to prefer opinionated stories, especially regarding complex issues such as the EU. This means a crisis of democratic institutions. In the case of the EU, its news coverage is also affected by a distant approach that overlaps with a feeling of remoteness towards the European project. Taking these trends into account, the current research aims to conceptualize the role of opinionated news in disinformation, as this practice takes advantage of a polarized public opinion. Beyond a theoretical approach, we use the multiple-case study as research strategy to assess the degree of opinion-oriented stories about the European Union (EU). The report of this issue suffers from cultural clashes that threaten its journalistic quality. In the multiple-case study here applied, we compare the coverage of EU affairs in local media from Germany, United Kingdom (UK) and Spain. The analysis is performed on a sample of news items on European issues (n=612), collected over a six-month period during the 2019 European Parliament

(EP) elections. The study focuses on two variables (personalization and negativity) through the analysis of headlines, topics and reader's comments. Drawing upon the sample, we argue that the prominence of opinion-oriented news about the EU could boost polarized disinformation. Polarization is more frequent in the UK (polarized liberal), while the German press shows approaches that seek a balance from different sources and the Spanish cases present a low negativity. Our theoretical approach reveals how the current state of play of journalism has influenced the success of polarization in the digital sphere. This disruptive communication around individualization could mean a decline of democratic institutions, as facts are no longer relevant for the audience.

Keywords: Opinionated news; personalization; polarization; disinformation; European Union.

Chapter 6. Lies are all around but who are the liars? Addressing online disinformation platforms in the Czech Republic and Slovakia

Websites publishing fake news share similarities across borders that are not yet fully addressed by global or national policies. As in other aspects of policy-making in this area, there is a conflict between the right of free speech and public interest in preventing social and economic damage caused by the deception of vast groups of society. This chapter offers empirical support to the argument that disinformation is adapted to specific vulnerabilities of concrete media systems and their audiences. In post-communist countries of Central Europe, the main structural weaknesses derive from a high concentration of media ownership and dominance of local oligarchs with real political and economic interests in the media ownership structures. The negative association connected with oligarchs who are presumed to exert a strong influence on what mainstream media produce has served as one of the key arguments for creating an alternative media landscape. The other important factor behind the gradual build-up of alternative eco-system in the region is effectively unanimous support of the membership of

the European Union and NATO by mainstream media. The sentiment of non-existing public discourse on Europe's fundamental principles and liberal values has been exploited by actors spreading pro-Russian propaganda and inciting polarization. Based on evidence from the Czech Republic and Slovakia, we argue that providers of false content use a distinguishable set of manipulation methods which ultimately enables the detection and systematic categorization of disinformation outlets. The biggest existing database of untrustworthy websites in the region was established in order to assist advertising companies in selecting secure online marketing placements and prevent negative associations between their clients and controversial online content. Apart from their deliberate use of fake news and misleading context, a lack of transparency on authorship, ownership, and funding is also an important attribute of these communication channels. Attempts at addressing insufficient transparency of numerous untrustworthy outlets have not yet resulted in effective policies against disinformation in the region.

Keywords: Disinformation; Pro-Russian propaganda; Online Eco-system; Czech Republic; Slovakia.

Chapter 7. Technologies and fact-checking: a sociotechnical mapping

Oscar Westlund, Rebekah Larsen, Lucas Graves, Lasha Kavtaradze & Steen Steensen

The fight against misinformation involves both human social actors and digital technologies, as well as a diverse set of institutions. Digital developments both enable and require critical evaluation of sources and information. This chapter assesses and analyzes the sociotechnical infrastructures available to and potentially used in fact-checking, in the form of digital technologies associated with fact-checking. The chapter presents findings from an international and systematic assessment of digital technologies associated with fact-checking. For this assessment, we collected data from primary and secondary sources between fall 2020 and spring 2022. This assessment is guided by a sociotechnical framework that ena-

bles analysis of the interrelationships between humans and technology. It employs a deductive categorization of the three main fact-checking stages of practice. We also then inductively created subcategories of technologies for these three stages. In this chapter, we give a detailed review of each of these categories and subcategories, including descriptions of specific technologies. The chapter offers four main takeaways. First, there is a multitude of technologies associated with fact-checking, many offering affordances that cannot be done by humans alone. Second, most technologies are connected to the identification stage, and are largely owned and controlled by platform and tech companies. Third, the verification stage encompasses a wealth of technological tools by third-party companies that typically require human fact-checkers to manually work with the technology in their fact-checking processes. Fourth, distribution takes different forms and shapes, both analogue and digital, and being proprietary as well as non-proprietary to the fact-checkers. Based on the findings in this chapter, we call for future research around fact-checking technologies and practices, furthering and informed by a sociotechnical lens.

Keywords: Fact-checking, misinformation, disinformation, technology, sociotechnical

Chapter 8. The role of fact-checking in fighting the ‘infodemic’ of disinformation on Covid-19: a case study of Polígrafo

Marina Ferreira & Inês Amaral

The Covid-19 pandemic created an unprecedented scenario, not only in terms of public health, with the determination of unprecedented social isolation measures, but also of massive information sharing on social networks - classified by the WHO as an “infodemic”. However, several mistakes were made early in political communication about the pandemic, mainly due to the need for more information and scientific evidence about the new disease. In addition, several leaders used the pandemic as a weapon for political

combat, disseminating false information according to their governmental needs and personal beliefs.

In response to the alarming amount of misinformation about SARS-CoV-2 and its dangers to public health, widespread mobilisation of fact-checking platforms has been observed to identify and correct false or misleading information. This chapter discusses fact-checkers role in mitigating misinformation about Covid-19 and the limitations their activity presents in this context. The present study aimed to analyse fact-check articles on Covid-19 published between September and December 2020. To ascertain uninformative trends, data were collected regarding the origin of the information studied, the ratings assigned within the Polygraph scale, the topics worked on, the verification methods used and, finally, the countries to which the verified content referred. One of the study's main conclusions is the predominance of verifying facts concerning information classified with a certain degree of falsity. It is also apparent how the evolution of the pandemic and the scientific advances made, for example, through the development of the first vaccine against the disease, affected the amount and type of Covid-19 content checked.

Keywords: Fact-checking, disinformation, Covid-19, Polígrafo.

Chapter 9. Framing Covid-19: How Fact-Checking Circulate On The Facebook Far-Right

Raquel Recuero, Taiane Volcan, Felipe Soares, Otávio Vinhas & Luiz Ricardo Huttner

This research focus on how fact-checking links circulate on Facebook groups/pages that also shared disinformation, particularly, the ones affiliated with the far-right. Through a three-step method that included content analysis, discursive analysis and social network analysis, we analyzed public 860 posts and found out that: (1) while fact-checking does circulate on these groups, they tend to be framed as disinformation through posts on far-right ones, which we call “explicit framing”; (2) the far-right groups tend

to cluster around specific fact-checking links that are mostly shared without a framing text, but whose theme support their own ideological narrative (which we call “silent framing”) and; (3) both explicit and silent framing tend to happen through populist discourse connections.

Keywords: Covid-19; Fact-checking; Far-right; Populism; Facebook.

CAPTURING AND DISSECTING THE COMPLEXITY OF PRODUCTION AND DISSEMINATION OF CONSPIRACY THEORIES, HATE-BASED RHETORIC, AND MIS-AND DISINFORMATION ONLINE

Ardian Shajkovci, Ramón Ruti, Asli Altinbay, Matteo Gregori, Amanda Garry & Allison McDowell-Smith¹

Introduction

The current social media landscape facilitates a direct spread of harmful, manipulative, and misleading information online (Hameleers, Powell, van Der Meer & Bos, 2020; Lewis & Marwick, n.d.) One of the major threats of the current online ecosystem is that mis-and disinformation often has more profound reach and flow than facts themselves (Innes, 2019). At least two-thirds of adult Americans rely on social media for some of their news updates and to react to news stories (Shearer & Gottfried, 2017). However, this can become dangerous when considering how the online environment can manipulate the soft facts of news stories to spread rumors, conspiracies, and social fragmentation (Innes, 2019; Christou 2020; Walther & McCoy, 2021; van Der Vegt, Gill, Macdonald & Kleinberg, 2019). Fringe independent websites and alternative social media platforms such as Telegram, Parler, and Bitchute have increasingly become major spreaders of mis-and disinformation and extremist rhetoric. These spaces have also become

1. The American Counterterrorism Targeting and Resilience Institute (ACTRI) and Storyzy. The authors wish to thank ACTRI Research fellow Samantha Walther for her assistance with data collection and analysis.

echo chambers of hate-based rhetoric and mis-and disinformation that can largely influence online users seeking propaganda related to specific movements. Alternative social media platforms are increasingly being characterized as spaces that “challenge traditional media” and serve as “self-perceptive corrective tool of traditional media” (Walther & McCoy, 2021, p.100).

By primarily utilizing QAnon as a case study, the authors relied on automated data collection processes and a digital disinformation detection platform to gather data surrounding the movement of conspiracy theories and mis-and disinformation online. Given the significant presence of conspiracy theories and mis-and disinformation on smaller and alternative social media platforms that are becoming increasingly popular among many users, the authors also sought to provide insight into innovative methodologies to collecting and analyzing “big data” on social media platforms in general, while also elucidating a larger goal of future collaborations to seek rigorous methodological and technological approaches to address the issue at hand. As also discussed in the ensuing sections, alternative platforms are increasingly becoming aggregates for hate-based and mis-and disinformation related posts, warranting further research to prevent such content from gaining further legitimacy and credibility within the online user base and beyond.

Key Considerations of the Evolving Social Media Landscape

In 2016, Facebook, YouTube, and Twitter started struggling with monitoring and content moderation of their platforms. This was in large due to difficulties managing the presence and distribution of large amounts of information across diverse and polarized user base (Gillespie, 2018). Despite the initial hands-off approach on dealing with the issue, main platforms started addressing hate speech, conspiracy theories, and other harmful content by tightening their Terms of Service (ToS), Community Standards (CS), or implementing more severe actions like censoring content and deplatforming

users. In February of 2022, in large part in response to COVID-19 crisis, Facebook (Meta) introduced its misinformation strategy dubbed as “remove, reduce, inform.” (Meta, 2022). The platform will remove misinformation in those instances when:

- “When misinformation has the potential to cause imminent physical harm.” (Meta 2022, p.1)
- “When misinformation has the potential to interfere with or suppress voting.” (p.1)
- “When content is manipulated in a way that would not be apparent to an average person.” (p.1)

In 2016, Facebook also launched a fact-checking program that “focuses on identifying and addressing viral misinformation, particularly hoaxes with no clear basis in fact.” (Facebook, 2019).

Deplatforming seemed to be an effective intervention to decrease the spread of misinformation, but only in the short term. When users are deplatformed, all their content is removed from the platform, and they can no longer create an account using their real names. Many information technology practitioners and scholars argue that when an account, website, hashtag, or piece of content is taken down, new ones can be created to replace them accordingly (Ali, 2021). Taking down user accounts and censoring their material could be counterproductive and may result in drawing people’s attention to suppressed materials (and hence publicizing extremist’s cause) due to the so-called “Streisand effect” (Tworek, 2020). In fact, while there are ample empirical studies on the efficiency of account takedowns vis-a-vis extremist/terrorist content, little has been done for disinformation networks (Innes, 2021).

Notable cases of misinformation-related deplatformed users are David Icke, influencer of the “New Age” conspiracy theories also tied to Qanon movement, with 800,000 followers on Facebook as of April 2020, and Kate Shemirani, a nurse known for her anti-vaccine and antisemitic rhetoric,

with thousands of followers at the time of her account removal in September of 2020. A study by Innes & Innes (2021) demonstrated that the deplatforming of Shemirani and Icke triggered aided suppressing problematic behavior and content in the short term, but also led to two unintended consequences: (1) the spawning of fake accounts spreading the content produced by the ‘influencer’ off-line or on other platforms and (2) the increasing of the resilience of the network by encouraging them to establish their presence in other platforms. Compared with other problematic extremist content, controlling mis-and disinformation is more challenging given (1) the less clear-cut contested nature of the harm and (2) because its removal may raise concerns about freedom of speech rights (Douek, 2021). Alternative measures to deplatforming have been taken into consideration to tackle the spread of mis-disinformation on online platforms, including content labeling and warnings, fact checking, media literacy campaigns, and limits on advertising revenue. However, current literature lacks empirical evidence on their effectiveness to reduce the spread of mis-and disinformation, and the heterogenous nature of disinformation makes intervention difficult to develop (Mena, 2019; Pennicook et al., 2020).

Given these actions to restrict misinformation, disinformation, and conspiracy theories, many conservative users have and continue to migrate to less regulated and monitored platforms. Gab, for example, after the so-called Twitter “purge” of far-right users in 2016, gained millions of new users in just a few weeks (Rogers, 2020). A study on Twitter, Gab, and Reddit users’ datasets showed that 59% of users who were deplatformed from Twitter created a new account on Gab (Ali et al., 2021). Aliapoulios et. al noted how “Parler experienced large influxes of new users in close temporal proximity with real-world events related to online censorship and deplatforming on mainstream platforms.” (p. 7)

The far-right/conspiracy theorists’ migration from public-facing to close social media platforms took place principally in 2018, and many described it as massive. Past research analyzing QAnon on Telegram has shown the

movement's growth on the platform and observed a 17% increase in group size of QAnon-related groups over 39 days in 2020. (Garry, 2021). The duality of such new platforms allows conspiracy theorists to create open channels, where individuals believing the same theories can spread evidence supporting their movement's thoughts, as well as close, invite-only channels used to further convert more believers. Though the overall numbers of interactions recorded are generally lower than on mainstream platforms, the content shared in alternative social media is likely to be more toxic (Ali, 2021). A study by Zihiri et Al (2022) indicated that among far-right QAnon groups in Telegram, the main topics remain race-related in a negative and discriminatory perspective. Besides, the absence of moderation of harmful content—unlike on public-facing media—coupled with the anonymity offered by alternative platforms, foster a high-potential radicalization environment (Katz, 2020).

QAnon Case Study

QAnon supporters are considered the pioneers of a distributed infrastructure of disinformation and misinformation that spans multiple platforms, including those often described as isolated, peripheral venues of the internet. To better understand the QAnon phenomenon, it is necessary to understand how such platforms work as structural bridges of the wider communication ecology. Previous studies demonstrated significant prevalence of deceptive information on smaller platforms relative to mainstream social media platforms (Zeng & Schäfer, 2021). Although QAnon originated from a relatively unpopular platform like 4chan, the movement quickly emerged in other mainstream media like Reddit, Twitter, Facebook and most importantly YouTube. QAnon's presence in the famed video-sharing platform is still moderate, with 868 active channels and 5,352 videos garnering 122,107,633 views as of December 2019 (de Zeeuw et al., 2020). Starting June 2018, QAnon content on the platform evolved from informal, "personal" vlogging style documentaries to professional cinematic videos, aimed at increasing the reach of the movement in the mainstream. Despite QAnon managing to

push its content through YouTube's search and recommendation systems, the platform is still quite isolated from the whole online disinformation ecosystem. A study by Gallagher et al. (2020) found that 20.4% of QAnon-related posts on Facebook redirected to YouTube videos, while only a few of the YouTube videos shared a link to QAnon's communities on other platforms (Gallagher et al., 2020).

Among QAnon followers on 8Kun and Gab, information shared mainly involves COVID-19 related content, including news updates and government response, as well as medical discussions about the virus. Among these discussions, Zeng & Schäfer (2021) found that 32% of the posts feature conspiracy theories content. Such media platforms are not used to produce disinformation but are rather exploited to repost content shared from other sources. As seen in Zeng & Schäfer (2021), YouTube and Twitter are the two most sourced websites on Gab and 8kun (13% of the overall posts). Other prominent sources cited include controversial far-right news websites like ZeroHedge and, not surprisingly, QAnon-related websites like *wearethene.ws* and *qnotables.com*. Similar results have been found while analyzing the data collected by the American Counterterrorism Targeting and Resilience Institute (ACTRI) "Telegram Monitoring Project," an initiative that monitors and analyzes the activities of groups linked to extremism, violent extremism, conspiracy theories, and mis- and disinformation on Telegram.

As can be seen in Fig. 1 below, links redirecting to video-sharing platforms, namely *youtube.com* (and mobile version *m.youtube.com*), *rumble.com*, *bitchute.com*, and *odysee.com* are among the most popular. The latter three websites are known to have a hands-off approach vis-à-vis content moderation within their domains; *rumble* is known to recommend conspiracy theories and harmful content more often than non-controversial content (House et al., 2021). *Bitchute* has been strongly criticized for allowing politically extreme and hateful content, more so than any other platform (Trujillo et al., 2020). *Rumble* has gained popularity for allowing users to share antisemitic and anti-vaccine content to generate revenue through

cryptocurrencies (Bogle, 2021). In line with what the authors found utilizing Storyzy data platform, as discussed further below, most news links redirect to Russian-based RT websites, an international TV network reportedly promoting anti-Americanism and anti-western sentiment, recently banned by the EU for spreading “toxic and harmful disinformation” following the Ukraine invasion. Interestingly, almost 50% of the overall URLs shared redirect to another Telegram channel or group (t.me). The result is even more telling when one considers that less than 1% of the links redirect to other alternative social media platforms. While more robust empirical research is warranted to demonstrate the flow of links from alternative social media platforms to Telegram channels, this result also suggests that Telegram’s mis- and disinformation communities do not seem to have a need to promote in other places, other than in already existing platforms namely. A snapshot of the partial network can be seen in Fig.2. below. The data collected in the project also revealed several links redirecting to lawful academic and scientific papers. This result is consistent with the hypothesis that alt-right and conspiracy theories groups seek to achieve intellectual legitimacy by actively engaging with scholarly literature—then moving back to simple lifestyle suggestions and populism.

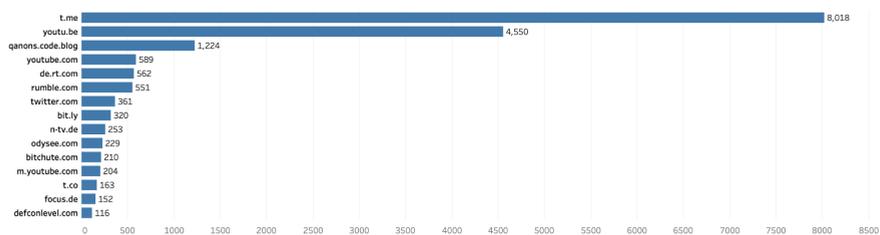


Figure 1—Number of URLs shared within disinformation groups on Telegram by domain. Source: Authors

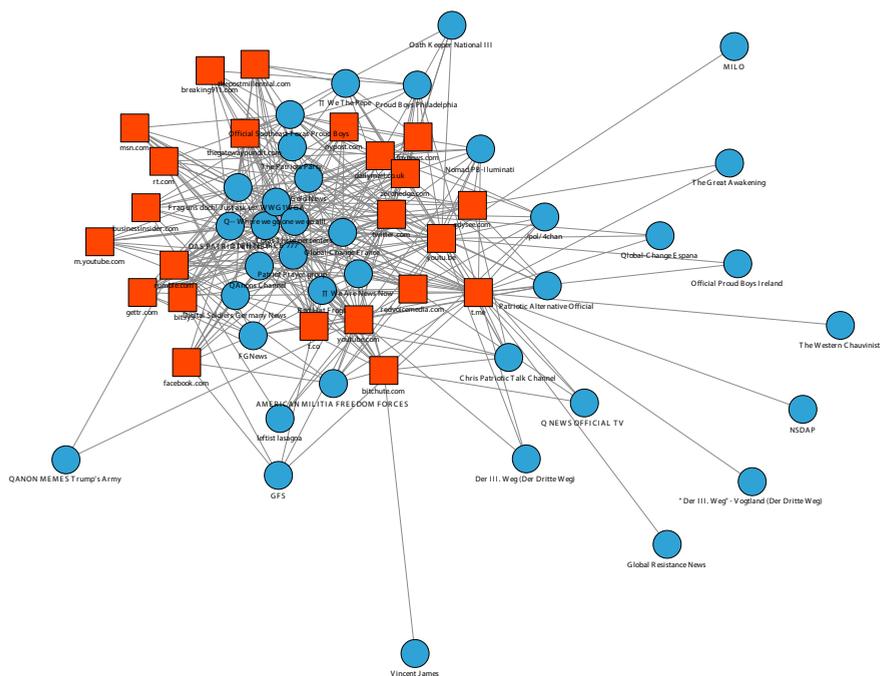


Figure 2—Social Network of Telegram channels and linked sources from the ACTRI Telegram Monitoring Project Nov. 2021 -Feb. 2022. Red nodes represent web domains, blue nodes represent Telegram channels. For aesthetic reasons, the plot displays only domains shared by at least 15 of the 40 most active channels. Source: Authors

The fact that a large part of the content shared comes from authoritative, neutral, or non-politicized sources lends credence to the idea that QAnon supporters tend to construct conspiratorial explanations of news reported somewhere else, rather than deliberately creating fake news from scratch. This process of assembling pieces of information (called crumbs) from breaking news events into “coherent” conspiracy narratives is referred to as “baking” (Munn, 2019).

Little empirical research has been conducted on the role of smaller social media platforms when it comes to mis- and disinformation, despite some available evidence indicating that the phenomenon is not just a Facebook or a Twitter problem. In 2016, forum-based social media Reddit, 9gag, and

Tumblr removed several hundred accounts used to spread divisive narratives on Russia's 2016 elections and a rhetoric to subvert US democracy (Lapowsky, 2018; Russel & Collins, 2018; Thomsen, 2018). New generation social media platforms leverage their open and anonymous (or even encrypted) nature to attract new users and hence represent an ideal entry point to insert mis-and disinformation into the online discussion. The mis-and disinformation content shared in the platform is likely to proliferate across the same platform, and later to different ones, to build credibility around the topic, amplify its impact and eventually hiding its origin. Melesehevich and Schafer (2018) compared the spread of disinformation to "illicit fund laundering, "specifically, "Just as ill-gotten money needs to be moved from an illegitimate source into an established financial institution, disinformation is most powerful when a façade of legitimacy is created through "information laundering." (p.1) The information laundering process consists of 3 steps: (1) "placement," i.e., the posting of the misleading information on social media; (2) "layering," i.e., the spread of disinformation from its point of origin to more credible sources; and (3) "integration," i.e., the phase where the disinformation is adopted by several sources and disseminated by other users on social media (Melesehevich & Schafer, 2018).

Methodology

The authors have focused several recent articles on the spread of hate-based rhetoric, conspiracy theories, and mis-and disinformation online. The research has particularly focused on the prevalence of the aforementioned on smaller, alternative social media platforms such as Telegram. The authors generated automated data collection processes ("bots") and relied on Storyzy's *Software-as-a-Service (SaaS)* tool. Storyzy developed a SaaS platform that helps to identify and analyze the global spread of online disinformation. This platform allows analysts to analyze disinformation networks and information environments (Storyzy 2022). In the context of this chapter, the authors utilized Storyzy platform to collect data on the spread of QAnon disinformation. The Storyzy platform offers source analysis,

topic analysis, article analysis, image analysis, IP address analysis, country analysis, social media analysis, and helps users to analyze disinformation sites, blogs, articles, YouTube channels, alternative video platforms, Telegram channels and tweets. The authors, for instance, were able to input certain topic keywords, such as “QAnon,” into Storyzy’s platform, which would return information on all cited articles, trusted and disinformation, regarding the topic. Storyzy contains a range of source classifications, including “Trusted, Disinformation, Think Tank, NGO, Academic, Association, Government, IGO, Political Party, Religious and Uncategorized” (Storyzy 2022). The two main categories analyzed in this chapter were “Trusted” and “Disinformation.” Disinformation is broken down further into subcategories which include “Clickbait, Conspiracy, Extreme left, Extreme right, False information, Hate, Propaganda, Pseudoscience, Satire, and Tabloid” content (Storyzy 2022).

Data Collection

The initial study by the authors relied on the Storyzy platform to gather quantitative data regarding the spread of QAnon related conspiracy theories and disinformation (Garry, Walther, Mohamed & Mohamed, 2021). The authors utilized Storyzy’s topic analysis feature to study QAnon sites and blogs that were sharing disinformation between March 1, 2020, and November 1, 2020. Simply by inserting the topic “QAnon” into the database and selecting this date range, the Storyzy platform automatically returned results regarding “Article Volume History about QAnon” broken down by source type (trusted sources versus disinformation sources). The results also included information of the 10 most cited articles about QAnon, 10 most cited sources about QAnon, top 10 source locations, a network analysis of the articles, and information on HTML tags to identify possible influence networks. However, the results focused on the article volume history to identify patterns and trends.

Storyzy platform has expanded analytic capabilities to extract data from smaller, alternative social media spaces, particularly focused on Telegram,

through its “Telegram Analysis” feature. Relying on an initial dataset of 208 Telegram channels classified as spreading hate-based rhetoric on Telegram, Storyzy was able to expand its algorithms to extract data from public channels and is continuously growing its index of channels and messages. To expand on the initial results, the authors also utilized the most recent “Telegram Analysis” feature provided by Storyzy. The Telegram feature is applied the same way as the topic analysis but analyzing the message volume history and links on Telegram channels rather than independent sites and blogs. The Telegram feature returns results regarding message volume history over the selected date range, the content of the messages by category, top 10 Telegram messages and top 10 Telegram accounts over the selected period, top 10 Telegram messages and top 10 Telegram accounts on the day preceding the request, top 10 sources and top 10 articles shared in Telegram messages, and the Telegram accounts network displaying the links with the shared sources which allows to identify Telegram communities that spread disinformation.

With the Telegram feature being novel, the number of channels in the Storyzy database is a work in progress and continues to grow daily. While the number of channels may be currently limited, it is constantly growing, and the feature provides a cutting-edge means for analyzing disinformation on an alternative social media platforms. Utilizing these features, the results highlight both the vast spread and sources of disinformation related to Qanon and the practicality of multidisciplinary collaborations to track and mitigate the growing threat of disinformation online.

The authors collected data through two distinct time periods: March 2020 through November 2020 and August 2021 through February 2022.

Findings

Finding 1: There was a total of 39,692 articles regarding QAnon released between March 1, 2020, to November 1, 2020, illustrating an 843.4% increase in QAnon publications compared to the eight months prior.

Articles from both trusted and disinformation sources were included in this search. Focusing on disinformation more specifically, the research analysis revealed that there was a total of 17,152 disinformation publications from the original Q drop on October 28th, 2017, to November 1, 2020 (Chart 1).² However, 46.8% of all those disinformation publications were released after March 1, 2020, suggesting an alarming rise in QAnon disinformation in the months preceding the 2020 U.S. Presidential election movements.

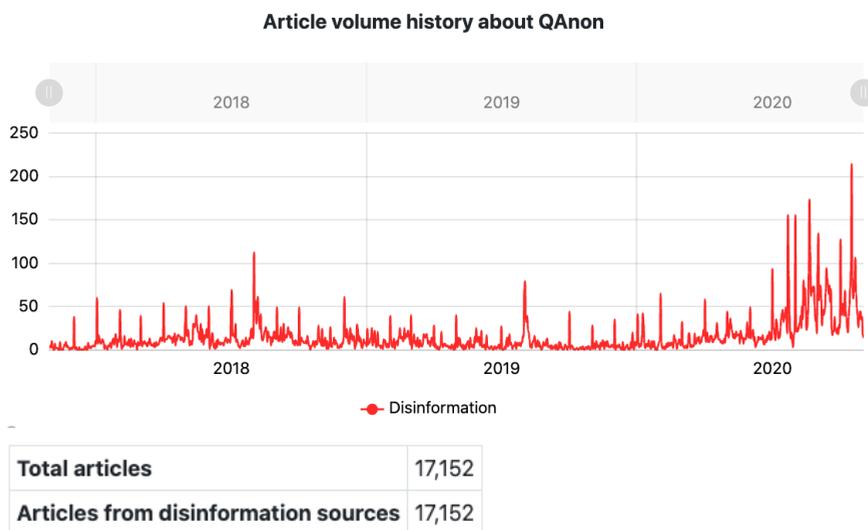


Chart 1—Volume of QAnon articles published by disinformation sources from October 28th, 2017, to November 1, 2020. Source: Storyzy

Finding 2: The top 10 most cited disinformation sources about QAnon were distributed by a range of independent fringe websites, both from the United States and abroad.

Websites such as the Westernjournal.com, the gatewaypundit.com, Breitbart.com, pjmedia.com, and the federalist.com are all classified as extreme right-wing news sources. These sources listed in Table 1 are the most quoted or cited disinformation and misinformation sources for other articles.

2. For a discussion on “Q drop,” please see Garry, A., Walther S., Rukaya R., Mohammed A. (2021, March). QAnon conspiracy theory: Examining its evolution and mechanisms of radicalization.

This analysis revealed that fringe and biased websites play a large role in the creation of disinformation that then gets re-shared to alternative spaces like Telegram.

10 most cited sources about QAnon	
 rt.com	219 links
 dailycaller.com	194 links
 rightwingwatch.org	161 links
 rawstory.com	144 links
 westernjournal.com	111 links
 thegatewaypundit.com	104 links
 breitbart.com	99 links
 zerohedge.com	92 links
 pjmedia.com	79 links
 thefederalist.com	76 links

Table 1—Top 10 Most Cited Sources about QAnon from March 1, 2020 - November 1, 2020. Source: Storyzy

Finding 3: A comparative analysis of messages across different languages revealed that certain languages and regions of the world (Germany, for example) have a higher prevalence of QAnon-related disinformation.

The authors utilized “The Storyzy Telegram” feature and analyzed 2,748 English messages featuring QAnon from March 1, 2020, to November 1, 2020 (Chart 2). Comparatively, there were nearly 8,075 QAnon messages on Telegram in German during the same timeframe, illustrating the disparity among the spread of misinformation across different cultures and geographic locations respectively. Close to 6,000 of those messages were from disinformation sources, which indicates that disinformation outlets are the primary driver in circulating this rhetoric.

There were only 265 messages found in Spanish, 1,055 in Russian, 594 in Italian, and 57 in French. Therefore, certain languages and thus the communities and countries where those languages are predominately spoken

may be more prone to spreading QAnon disinformation than others. While this number is most likely a small sample of the true amount of content on Telegram, it reveals a significant discrepancy in the percentage of messages that contain disinformation. This analysis revealed that 1,985 out of the 2,748 (72.2%) messages were from disinformation sources as opposed to trusted sources.

This is a much larger percentage of messages being from disinformation sources on Telegram compared to the overall topic analysis of QAnon content from other sites and blogs. When analyzing the topic of QAnon on other sites and blogs during the same time period, only 20.2% of all publications were from disinformation sources. This suggests that despite most recent publications regarding QAnon being from trusted sources, as conversations move to alternative spaces, the prevalence of disinformation significantly increases. Users on Telegram create their messages based heavily off disinformation sources which are then shared throughout the Telegram channels. As Table 2 suggests, not only is disinformation more prevalent than truthful information, but disinformation sources also receive significantly more views as well.

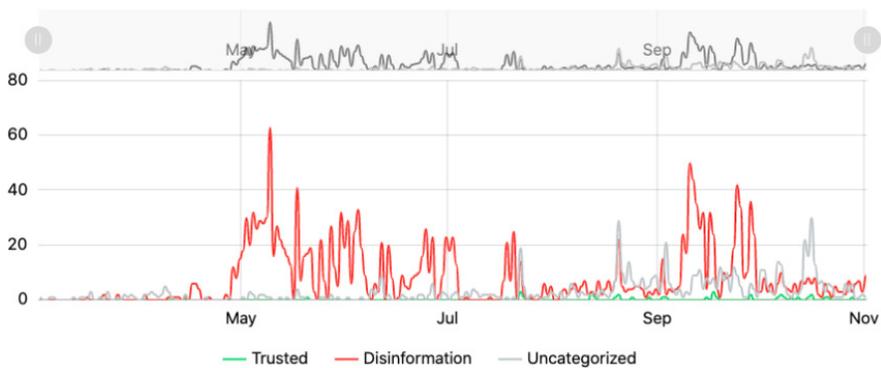


Chart 2—Telegram Message Volume from March 1, 2020 - November 1, 2020. Source: Storyzy

	Messages	Total views	Yesterday views
Total messages	2,748	10,091,948	0
Messages from trusted sources	35	130,125	0
Messages from disinformation sources	1,985	7,940,170	0
Messages from uncategorized sources	728	2,021,653	0

Table 2—Telegram Message Volume from March 1, 2020 - November 1, 2020. Source: Storyzy

Finding 4: Of the channels currently in the Storyzy database, the QAlerts App channel on Telegram is the most widely viewed QAnon related channel during the eight-month period prior to the election (Table 3).

This channel has over 55k subscribers worldwide and simply re-posts the Q drops to Telegram. This channel does not have a discussion option; however, this is one of the most widely shared channels. The content from QAlerts App channel is posted on other Telegram channels that do allow for conversation. Other channels such as QAnon International Info, Qlobal Change USA, and the Conspiracy Hole continue to spread conspiracy and extreme rhetoric surrounding several topics. Most notably, the conversations generally share conspiracy beliefs about the vaccine being a government orchestrated plan to wipe out large portions of the population, election fraud speculation, as well as racist and anti-Semitic rhetoric. Most of the most viewed channels leading up to the election were classified as disinformation and garnered over 7,800,000 views accumulatively.

10 most viewed users about QAnon from 29/02/2020 to 01/11/2020	
 qalertsapp	6,980,100 views
 bitchbotboiilive	964,268 views
 qanoninternationalinfo	348,664 views
 qnewsofficialtv	197,132 views
 bigshocktheory	179,098 views
 potusitaly	146,536 views
 bloomberg	112,700 views
 qlobalchangeusa	108,200 views
 theconspiracyhole	88,190 views
 schrangtv	84,300 views

Table 3—10 Most Viewed Telegram Channels from March 1, 2020 - November 1, 2020. Source: Storyzy

Finding 5: The 10 most shared sources about QAnon and which independent websites were most linked to on Telegram channels (Table 6).

The QAlerts.pub (same site as qalerts.app) has significantly more shares than any other websites, highlighting this site’s role in the flow of QAnon disinformation from the original Q drop, to notifying adherents, and then to being shared on alternative platforms. Speculations can be made as to why certain sources were used more than others, and what elements in each forum were more appealing among adherents. Examining the nuances between each website can reveal minute details and tactics that are leveraged to manipulate users online and further radicalize them through conspiracy-based rhetoric.

10 most shared sources about QAnon	
● qalerts.pub	1,324 share(s)
● operationq.pub	103 share(s)
● qalerts.app	46 share(s)
● nbcnews.com	21 share(s)
● bloomberg.com	15 share(s)
● spiegel.de	14 share(s)
● theepochtimes.com	13 share(s)
● businessinsider.fr	13 share(s)
● qalerts.net	12 share(s)
● theguardian.com	11 share(s)

Table 6—10 Most Shared Sources about QAnon from March 1, 2020 - November 1, 2020.
Source: Storyzy

Finding 6: There is a solid social network among misinformation sources and their appeal among users on Telegram.

Figure 3 below reveals the Telegram communities that share disinformation sources. The squares represent websites while the circles represent Telegram channels. This social network visualizes which sites and links get shared to individual Telegram channels. The social network visualization below highlights the range and complexity of disinformation-related sources online, across independent websites, news hubs, and blogs, to just name a few. For example, content is being pulled from independent QAnon-related websites, far-right websites, pseudoscience websites, and independent far-right biased news sources to the Telegram channels. This allows a range of disinformation to then be accumulated onto any singular QAnon channel, exacerbating the range of social problems that QAnon can manipulate into its ideology and rhetoric. The extent of the connectivity and disinformation data flow as it relates to Qanon serves as a barrier to objective reporting, threatening the very fabric of legitimate news and information sources in the fight against conspiracy theories, mis- and disinformation, and hate-based rhetoric.

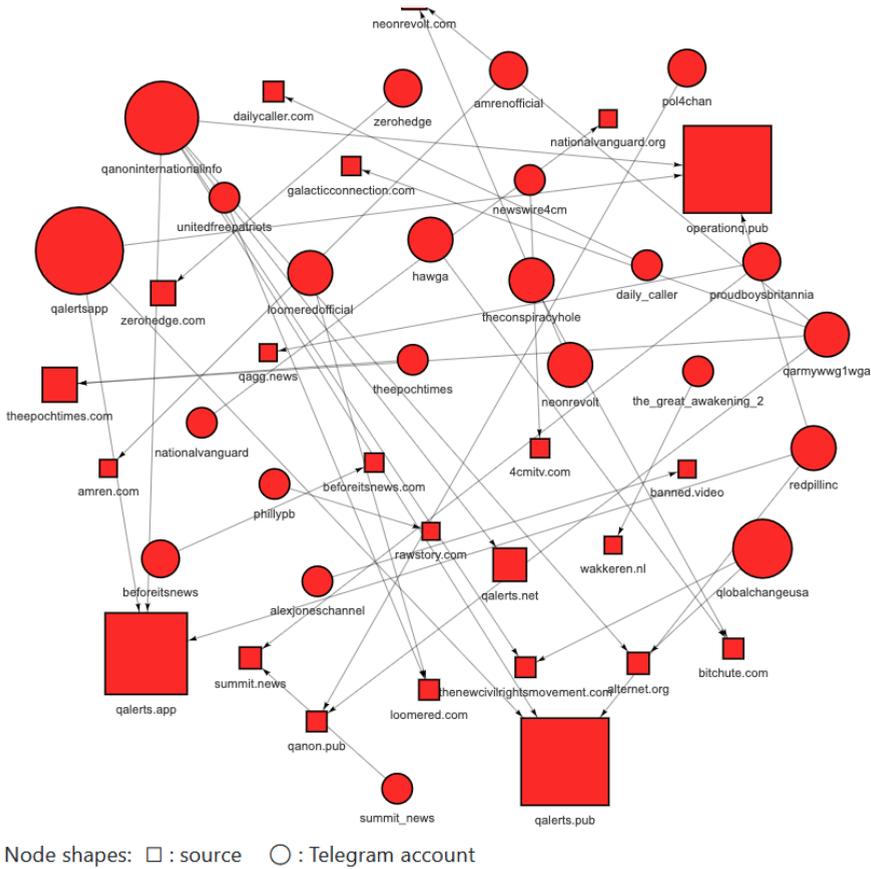


Figure 3—Social Network of Telegram Channels and Linked Sources from March 1, 2020- November 1, 2020. Source: Storyzy

Finding 7: In the eight-month period prior to the 2020 presidential election, disinformation related to QAnon was on the rise.

Disinformation was being produced at greater rates by independent websites and new sources, as well as being consumed on Telegram more rapidly than trusted sources. This finding is in line with previous research indicating that falsehoods and disinformation often spread more rapidly and more widely than verified facts (Innes, 2019). The digital environment, especially those on alternative social media platforms such as Telegram, has become

a primary means for spreading QAnon rhetoric and other lines of extremist thinking. The QAnon movement is unique in the fact that it has evolved from simple disinformation online to one of the most prominent lines of political thinking in the United States and has currently manifested into at least 23 instances of violence, including the Capitol Insurrection on January 6th, 2021.

A major component of the original QAnon study focused on tracking the flow of QAnon disinformation after the initial Q drop. The researchers found that after the original Q drop was shared to 4chan or 8kun, the update would make its way down to independent QAnon sites like Qalerts.app. These sites would serve as aggregates for all Q drops and often notified users of new Q drops. As users became aware of the new information, it was then that they could be reposted to individual blogs and Telegram channels to truly create further echo chambers and conversations. The digital disinformation platform allowed the authors of the article to retrospectively analyze the most popular sources reaching Telegram and the quantity of disinformation on these QAnon Telegram channels.

Finding 8: Continued threat of mis-and disinformation sources and their appeal among users on Telegram.

The recent data collected between August 2021 and February 2022 suggest significant user engagement with Qanon content on Telegram, considerably more significant compared to data collected between March 1, 2020-November 1, 2020 (See Table 7). For instance, “Topic Analysis” feature on Storyzy platform revealed over 11 million views of the content on Qanon classified as a disinformation source by the platform.

	Messages	Total views	Yesterday views
Total messages	11,559	29,928,063	70,845
Messages from trusted sources	25	34,926	0
Messages from disinformation sources	1,311	11,330,027	7,400
Messages from uncategorized sources	10,223	18,563,110	63,445

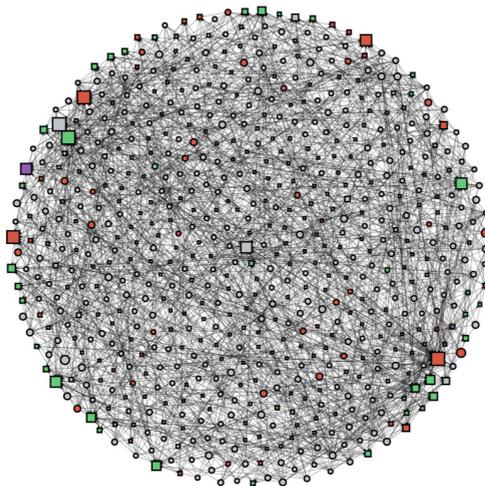
Table 7—Telegram Message Volume on QAnon from August 24, 2021 - February 25, 2022. Source Storyzy

The network link analyses below (Fig. 4, 5) demonstrate signaling behavior between users and information to and within alternative platforms. Put differently, the network-based visualization below highlights certain aspects of the information travel patterns as it relates to Qanon between Telegram channels.

Squares: Outlet, sources

Nodes: Telegram accounts

Node shapes: □ : source ○ : Telegram account
 Node types: ● Trusted ● Disinformation ● Think Tank ● NGO ● Academic ● Association ● Government ● IGO ● Political party ● Religious
 ● Uncategorized ● Neutral



Capturing and dissecting the complexity of production and dissemination of conspiracy theories, hate-based rhetoric, and mis- and disinformation online

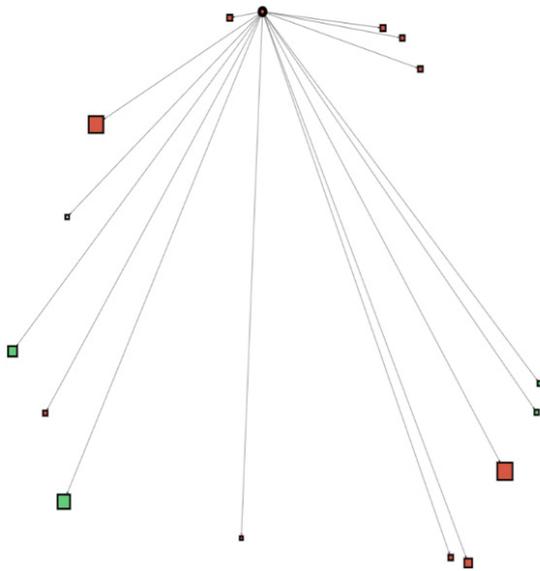


Fig. 4. The network visualization reflects Telegram accounts that share a particular post or source of information. For instance, the Telegram account @patriotarmy shared this specific link/outlet: <https://www.bitchute.com/video/UimTMFGQb0Br> on its Telegram channel Patriot Army – Telegram. This particular Telegram account also shared <https://www.bitchute.com/video/UimTMFGQb0Br> on its Telegram channel Patriot Army – Telegram. This particular Telegram account also shared <https://www.thegatewaypundit.com/2021/11/exclusive-john-pierce-william-shiplee-take-defense-qanon-shaman> . See table below for additional breakdown. Source: Storzzy

@patriotarmy		
	https://www.bitchute.com/video/UimTMFGQb0Br	2 message(s)
	https://www.dailymail.co.uk/news/article-10272869/Hunter-Biden-Tucker-Carlson-close-friends-emails-show.html	1 message(s)
	https://prayingmedic.com/2018/12/21/qanon-the-16-year-plan-to-destroy-america	1 message(s)
	https://www.redvoicemedia.com/video/2021/11/general-flynn-qanon-is-a-disinformation-campaign-by-the-cia	1 message(s)
	https://thefederalist.com/2022/01/06/reno-911-the-hunt-for-qanon-is-a-work-of-art	1 message(s)
	https://www.ntd.com/judge-sentences-qanon-shaman-jacob-chansley-to-over-3-years-in-jail_702915.html	1 message(s)
	https://www.thegatewaypundit.com/2021/11/exclusive-john-pierce-william-shiplee-take-defense-qanon-shaman	1 message(s)
	https://www.zerohedge.com/political/qanon-shaman-gets-41-months-prison-jan-6-riot	1 message(s)
	https://rumble.com/vq5rdz-the-lin-wood-general-flynn-feud-occult-prayers-moabs-and-operation-fiddler.html	1 message(s)
	https://www.theepochtimes.com/government-prosecutors-seek-51-months-in-jail-for-qanon-shaman-jacob-chansley_4096940.html	1 message(s)
	https://www.dallasobserver.com/news/leaving-las-vegas-the-next-q-con-may-take-place-on-a-plantation-in-south-carolina-owned-by-lin-wood-13216455	1 message(s)
	https://www.zerohedge.com/markets/justice-coming-qanon-shaman-appealing-result-his-guilty-plea	1 message(s)
	https://watkins.bitnitro.net	1 message(s)
	https://www.rollingstone.com/politics/politics-news/ron-watkins-qanon-digital-rosa-parks-1247187	1 message(s)
	https://www.dailymail.co.uk/news/article-10252807/Lin-Wood-Releases-Michael-Flynn-audio-calling-QAnon-CIA-disinformation-plot-total-nonsense.html	1 message(s)
	https://prayingmedic.com/2018/03/30/qanon-march-30-2018-who-is-q	1 message(s)
	https://www.salon.com/2021/11/27/qanon-congressional-candidate-sets-up-bitcoin-fundraising-operation-that-may-be-illegal_partner	1 message(s)
	https://thepostmillennial.com/breaking-qanon-shaman-three-years	1 message(s)
	https://thepostmillennial.com/biden-blames-his-problems-on-qanon-and-donald-trump	1 message(s)

Source: Storzzy

Squares: Outlet, sources

Nodes: Telegram accounts

Node shapes: □ : source ○ : Telegram account

Node types: ● Trusted ● Disinformation ● Think Tank ● NGO ● Academic ● Association ● Government ● IGO ● Political party ● Religious
● Uncategorized ● Neutral

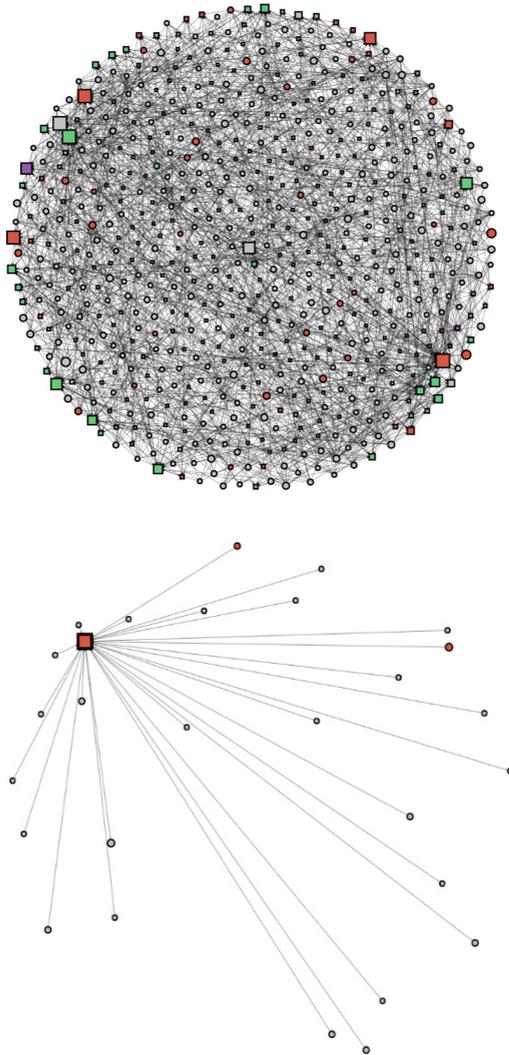


Fig. 5. The network visualization reflects Telegram accounts that share a particular post or source of information. For instance, @highdmary Telegram channel has shared 21 messages containing this particular post <https://www.bitchute.com/video/iAd0YVsSpSI1>. Other Telegram channels such as @qflys, @archiveanon, @redcapmaga, and @rightmindedpatriots similarly shared the same post.

bitchute.com	
https://www.bitchute.com/video/iAd0Yv5pS11	21 message(s)
https://www.bitchute.com/video/RJfEmK7nN5K	21 message(s)
https://www.bitchute.com/video/gZ05EilgD91L	21 message(s)
https://www.bitchute.com/video/UimTMFGQb08r	20 message(s)
https://www.bitchute.com/video/723dnNnGcqJ7	3 message(s)
https://www.bitchute.com/channel/1uDxpDogKMe9	3 message(s)
https://www.bitchute.com/channel/n78PbEkvWv2g	3 message(s)
https://www.bitchute.com/channel/9c7qjvwx7YQT	3 message(s)
https://www.bitchute.com/video/m5LnaM350Tu	2 message(s)
https://www.bitchute.com/video/RC1NFMc35I	2 message(s)
https://www.bitchute.com/video/pomPkaMtfy3	2 message(s)
https://www.bitchute.com/video/f9Y2Qjyhng1	1 message(s)
https://www.bitchute.com/video/m4KDau6VhmGm	1 message(s)
https://www.bitchute.com/video/LNS1mrfQmgB	1 message(s)
https://www.bitchute.com/video/f8Q0Xmk1qw1	1 message(s)
https://www.bitchute.com/video/msswwonlohp5g	1 message(s)

Source Storyzy

Limitations and Conclusion

This chapter engages with the important issue of the role of smaller social media platforms, as part of a wider media ecosystem engaged in constructing and communicating hate-based rhetoric, conspiracy theories, and mis- and disinformation. The authors have attempted to make the point that much of the current political, popular and media conversation about regulating social media focuses upon the role of mainstream social media platforms, namely Twitter, Facebook and to a lesser extent YouTube, which might be somewhat misplaced and misleading. Despite the renewed interest on the subject by policymakers and practitioners in the field, it remains a largely neglected topic that requires robust analysis to draw real world implications of the presented argument. The authors have attempted to present an initial discussion focused upon link analysis, signaling behavior between users and information to alternative accounts, and the like. An attempt was also made to employ network-based visualization tools to convey certain aspects of the information travel patterns, while highlighting the need for more robust research to data-driven analysis and interpretation in mapping the links between nodes.

The authors have attempted to make the case that a sole focus of mainstream social media platforms like Facebook and Twitter vis-a-vis debates

on social media regulation may not necessarily capture the complexity of the production and dissemination of hate-based rhetoric, conspiracy theories and mis-and disinformation online. This is especially important considering that Telegram has been explicitly positioning itself in the media ecosystem with a libertarian and freedom of speech protecting ethos, in the United and other western countries included, which does reduce the opportunities for regulation securing traction.

Storyzy database has over 2 million indexed sources of information, with over 42,000 classified as disinformation sources tracked in over 41 languages. Storyzy defines disinformation as “any form of intentional manipulation aiming at publishing and spreading false information, conspiracy theories or propaganda so as to deceive the audience,” (Storyzy 2022) classified under 10 categories, with “propaganda, hate, satire, false information, conspiracy, viral content, extremist, pseudoscience, and tabloid” being some of them (Storyzy 2022). Classification of disinformation sources is achieved through a mix of human-led expertise and artificial intelligence (AI). Such data classification process is deemed important to enable initial data identification and classification (human-led expertise) and then rely on AI to sift through significant data points and generate conclusions that go beyond mere analyst’s collection and interpretation of data.

Storyzy focuses on assessing reliability but not veracity, meaning that not all pieces published by a disinformation outlet are necessarily false and some pieces published by trusted sources can spread false information. For instance, RT.com may be considered a vector for disinformation but equally some of their output may be considered factually accurate. In this sense, how and why an outlet can objectively be defined as a “disinformation source” may be tricky at times. Likewise, even for Q based Telegram channels, some of what they convey may be true, and much of the traffic may be adherents engaging in ‘social chatter’ and/or commenting or debating particular claims and counterclaims. Such content does not necessarily constitute disinformation. Furthermore, data analysis platforms and tools have strengths and weaknesses that shape and mold what data is and is

not served up to the analyst. These structuring influences are important to consider when making a reasoned assessment of the validity and reliability of the findings presented. Additionally, some of the issues raised in the chapter are symptoms of the unique affordances of Telegram as a platform, and not necessarily of non-mainstream platforms more generally.

Other significant limitation in studying the generation and spreading of hate-based rhetoric and mis-and disinformation using open-source gathering tools is that individuals and groups may be relying on newly developed text-based instant messaging platforms. For example, between 2018-2019, a few months after Telegram increased the enforcement of its ToS to tackle the spread of jihadist content in the platform, many IS-affiliated online jihadists started testing marginal communication channels like BCM Messenger, Hoop Messenger, Riot.im, Rocket.Chat, TamTam, and Gab chat (the official instant messaging client for Gab.com); they became the echo chambers of far-right extremist activity absent content moderation and content removal policies (Clifford, 2020). Such platforms leverage the anonymity of their channels through end-to-end encryption to attract users, meaning no message can be accessed outside the communicating group, not even the platform itself. Most of the researchers carried out focus on using open-source data, retrievable from open media platforms like Facebook and Twitter, leading to an (inevitable) lack of information on smaller, alternative social media platforms. In this regard, the future research should remain on providing more robust accounts of the methodologies used (or that could be used) to trace how specific hate-based or mis-and disinformation content ‘travels’ and spreads across multiple platforms.

Lastly, without adequate intervention, the internet and social media will continue to serve as breeding grounds for the cultivation and spread of hate-based rhetoric and mis-and disinformation. Vehement support for mis-and disinformation, hate-based rhetoric and conspiracy-based causes can result in violence and opposition, as seen in the January 6th, 2021, riots. Smaller, alternative platforms present greater concerns as they provide limited space for counter-narratives and are likely to encounter new user

migrations from mainstream social media platforms. Failing to consistently address the issue may inadvertently contribute to the flow of hate-based rhetoric and mis- and disinformation online long term.

References

- Aliapoulos, M., Bevensee, E., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., & Zannettou, S. (2021, February). An early look at the parler online social network. *ArXiv preprint*. <https://arxiv.org/abs/2101.03820>
- Ali, S., Saeed, M. H., Aldreabi, E., Blackburn, J., De Cristofaro, E., Zannettou, S., & Stringhini, G. (2021, June). Understanding the Effect of Deplatforming on Social Networks. In *13th ACM Web Science Conference 2021*. <https://doi.org/10.1145/3447535.3462637>
- Bogle, A. (2021). Buying and selling extremism. *International Cyber Policy Center*. http://ad-aspi.s3.amazonaws.com/2021-08/Buying%20and%20selling%20extremism_v2.pdf.
- Clifford, B. (2020, November). Migration Moments: Extremist Adoption of Text-Based Instant Messaging Applications. *GNET*. https://gnet-research.org/wp-content/uploads/2020/11/GNET-Report-Migration-Moments-Extremist-Adoption-of-Text%E2%80%91Based-Instant-Messaging-Applications_V2.pdf
- Collins, B., Zadrozny, B., Jurecic, Q., & Douek, E. (2020, August 7). Podcast: Ben Collins and Brandy Zadrozny explain QAnon. *Brookings*. <https://www.brookings.edu/techstream/podcast-ben-collins-and-brandy-zadrozny-explain-qanon/>
- Christou, M. (2020). Is the radical right spreading the coronavirus? *Open Democracy*. <https://www.opendemocracy.net/en/countering-radical-right/radical-right-spreading-coronavirus/>
- de Zeeuw, D., Hagen, S., Peeters, S., & Jokubauskaite, E. (2020, November). Tracing normification: A cross-platform analysis of the QAnon conspiracy theory. *First Monday*, 25(11). <https://journals.uic.edu/ojs/index.php/fm/article/view/10643>

- Facebook. (2019, July). Our approach to misinformation. <https://transparency.fb.com/features/approach-to-misinformation/>
- Gallagher, A., Davie, J., Hart, M. (2020). Cyber swarming, memetic warfare, and viral insurgency: How domestic militants organize on memes to incite violent insurrection and terror against government and law enforcement. <https://www.isdglobal.org/wp-content/uploads/2020/07/The-Genesis-of-a-Conspiracy-Theory.pdf>
- Garry, A., Walther S., Rukaya R., Mohammed A. (2021, March). QAnon conspiracy theory: Examining its evolution and mechanisms of radicalization. *Journal for Deradicalization*, No. 26, 152-215. <https://journals.sfu.ca/jd/index.php/jd/article/view/437/265>
- Gillespie, T. (2018). Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media. *Communication Today*. <https://yalebooks.yale.edu/book/9780300173130/custodians-internet>
- House, E., Stanley, I., Wright, A. (2021, May). Rumble sends viewers tumbling toward misinformation. *WIRED*. <https://www.wired.com/story/rumble-sends-viewers-tumbling-toward-misinformation/>
- Hameleers, M., Powell, E.T., van Der Meer, G.L.A.T., & Bos, L. (2020, February 5). A picture paints a thousand lies? The effects and mechanisms of multidimensional disinformation and rebuttals disseminated via social media. *Political Communication*. <https://www.tandfonline.com/doi/full/10.1080/10584609.2019.1674979>
- Innes, M. (2019, November 29). Techniques of disinformation: Constructing and communicating “soft facts” after terrorism. *Wiley*. <https://onlinelibrary.wiley.com/doi/epdf/10.1111/1468-4446.12735>
- Lapowsky, I. (2018, March 23). Tumblr IDs 84 Accounts That Spread Propaganda. *Wired*. <https://www.wired.com/story/tumblr-russia-trolls-propaganda/>.
- Lewis, B., & Marwick, E. A. (n.d.). Media manipulation and disinformation. *Data & Society*. <https://datasociety.net/library/media-manipulation-and-disinfo-online/>

- Mena, P. (2019, July). Cleaning up social media: The effect of warning labels on likelihood of sharing false news on Facebook. *Policy & Internet*, 12(2), 165-183. <https://doi.org/10.1002/poi3.214>
- Meta. (February 2022). "Our approach to misinformation." <https://transparency.fb.com/features/approach-to-misinformation/>
- Melesehevich, K., Schafer B. (2018, January). Online information laundering: The role of social media. Alliance for Securing Democracy. <https://securingdemocracy.gmfus.org/online-information-laundering-the-role-of-social-media/>
- Munn, L. (2019). The High/Low Toggle. Alt-Right Code Switching on 8chan. *Navigationen-Zeitschrift für Medien-und Kulturwissenschaften*. <https://mediarep.org/handle/doc/14754> <https://securingdemocracy.gmfus.org/online-information-laundering-the-role-of-social-media/>
- Pasquetto, V., Olivieri A.i, Tacchetti, L., Spada, A., Riotta, G. (2022, April). Disinformation as Infrastructure: Making and maintaining the QAnon conspiracy on Italian digital media. *Proceedings of the ACM on Human-Computer Interaction*. <https://doi.org/10.1145/3512931>
- Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020, February). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science*. <https://doi.org/10.1287/mnsc.2019.3478>
- Rogers, R. (2020). Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication*. <https://journals.sagepub.com/doi/full/10.1177/0267323120922066>
- Russel, J., Collins, B. (2018, March 1). Russians Used Reddit and Tumblr to Troll the 2016 Election. The Daily Beast. <https://www.thedailybeast.com/russians-usedreddit-and-tumblr-to-troll-the-2016-election>
- Shearer, E., & Gottfried, J. (2017, September 7). New use across social media platforms 2017. *Pew Research Center*. <https://www.pewresearch.org/journalism/2017/09/07/news-use-across-social-media-platforms-2017/#:~:text=News%20Use%20Across%20Social%20Media%20Plat->

forms%202017%201,still%20get%20news%20from%20more%20tra-
ditional%20platforms.%20

Storyzy. (2022). <https://storyzy.com>

Thomsen J., (2018, April). Reddit identifies nearly 1,000 accounts linked to Russian troll farm. *The Hill*. <http://thehill.com/policy/technology/382558-redditidentifies-nearly-1000-accounts-linked-to-russian-troll-farm>

Tworek, H., Leerssen, P. (2019, April). An Analysis of Germany's NetzDG Law. *The Transatlantic Working Group Papers Series*. https://cdn.annenbergpublicpolicycenter.org/wp-content/uploads/2020/06/NetzDG_TWG_Tworek_April_2019.pdf

Trujillo, M., Gruppi, M., Buntain, C., Horne, B.D. (2020, May). What is BitChute? Characterizing the "Free Speech" Alternative to YouTube. *Proceedings of the 31st ACM Conference on Hypertext and Social Media*. <https://dl.acm.org/doi/10.1145/3372923.3404833>

Urman, A., & Katz, S. (2020, July). *What they do in the shadows: Examining the far-right networks on Telegram*. *Information, communication & society*. <https://www.tandfonline.com/doi/abs/10.1080/1369118X.2020.1803946?journalCode=rics20>

Zeng, J., Schäfer Mike S. (2021) Conceptualizing "Dark Platforms." Covid-19-Related Conspiracy Theories on 8kun and Gab. *Digital Journalism*. <https://www.tandfonline.com/doi/full/10.1080/21670811.2021.1938165>

Van Der Vegt, W. I., Gill, P., Macdonald, S., & Kleinberg, B. (2019). Shedding light on terrorist and extremist content removal. *RUSI*. https://www.researchgate.net/publication/334230631_Shedding_Light_on_Terrorist_and_Extremist_Content_Removal

Walther, S., & McCoy, A. (2021). US extremism on Telegram: Fueling disinformation, conspiracy theories, and accelerationism. *Perspectives on Terrorism*, 15(22), 100-124. <https://www.universiteitleiden.nl/binaries/content/assets/customsites/perspectives-on-terrorism/2021/issue-2/walther-and-mccoy.pdf>

THE TWILIGHT ZONE: CASE STUDIES IN MISINFORMATION & MASS MEDIA

Alyssa M. Brumis

Introduction

Misinformation and disinformation have always existed in American society. However, over the last three decades, with advancements across mass communication, like the Internet and social media, the problem has become more prevalent, with drastic and deadly consequences. Misinformation is spread across mass media platforms, reaching larger audiences than ever before (O'Connor & Weatherall, 2018). Both the ways in which we consume (mis)information and the speed at which we do so has radically changed, making it a complex and multifaceted social problem. Bennett and Livingston (2018) contend that disinformation is “a systemic problem that reverberates through the interlinked mainstream news media and alternative media ecosystem,” referred to as the DAR cycle: disinformation – amplification – reverberation (Udupa, 2020, p. 5). And this surplus of misinformation across mass media – not only accessed on our televisions, radios, and computers, but now in the palm of our hands, on our cell phones, iPads, and tablets – impacts us in ways that leave us relatively defenseless.

Due in part to motivated reasoning, or emotion-based reasoning, researchers have determined that people are more likely to believe information that reinforces our pre-existing beliefs (Cooke, 2017; Keener, 2019). Likewise, a study by Lewis & Marwick (2017) confirmed

that people are more likely to believe *misinformation* that reinforces their pre-existing beliefs. In a similar vein, processes like confirmation bias, also known as selective exposure, demonstrates that we *seek out* information to align with our pre-existing beliefs and we *avoid* information that does not (Sears & Freedman, 1967; Jonas et al, 2001; Phillips et al, 2014). Repetition also plays a role in how our brains respond and react to mis/disinformation: the more often we hear false information, the more we believe it to be true, known as the illusory truth effect (Hasher & Goldstein, 1977; Bacon, 1979; Arkes, Hackett, & Boehm, 1989; Begg, 1992; Fazio et al, 2015). And, shockingly, even when false information is debunked, Nyhan & Reifler (2010) found that the “backfire effect” may occur, which describes a “boomerang effect” that further reinforces a belief in the incorrect information, making it incredibly challenging to combat. Furthermore, the ways in which our false beliefs impact our behavior illustrates why the surplus of mis/disinformation across American mass media is particularly alarming.

As research suggests, false beliefs impact behavior (Geraerts et al, 2008; Baek, Kang, & Kim, 2019; Cheng & Luo, 2020; Roozenbeek, 2020). This could be relatively harmless depending on the topic: misinformation related to the flat-earther movement is far less threatening than misinformation related to the anti-vaccination movement, as the latter impacts public health outcomes. For example, the measles, which was previously eradicated through vaccinations, made a resurgence in America due in part to the anti-vaccination movement (Hussain et al, 2018), marking 2019 with the highest number of reported cases since 1992 (Patel et al, 2019), demonstrating the dire consequences of misinformation.

The following is a collection of three abbreviated case studies focusing on the dissemination of mis/disinformation across mass media, and the deadly outcomes as a result. The overarching theme throughout each case is the notion that the spread of mis/disinformation disrupts our connection to facts, logic, and reality, to the point that it becomes a danger. These cases explore mis/disinformation as a threat to public safety, a threat to democracy, and a threat to public health by examining the disseminators of mis/

disinformation across mass media. Exposing the effects and consequences of mis/disinformation using interdisciplinary frameworks demonstrates the dangers posed within American society by turning the country into something that more closely resembles an episode of *The Twilight Zone*, a fictitious series that aired in the 1960s depicting surreal situations in an alternate reality that led to disturbing consequences.

Misinformation & Disinformation Defined

Wardle & Derakhshan (2017) differentiated between disinformation and misinformation by defining disinformation as “false information *knowingly* shared to cause harm,” and misinformation as false information without harm necessarily intended (p. 3). The difference lies in the intent, which can make it challenging to distinguish between the two, as they are context dependent by definition. Disinformation may be spread purposefully by one source, then spread unintended by another, moving it from disinformation to misinformation, and revealing an overlap between the two.

Because my aim is not to debate intention, but instead to focus on impact and consequences of both dangerous disinformation and misinformation, I will primarily use the term misinformation to refer to false information and I will use disinformation only when there is clear intent that the false information was created and spread intentionally. I will use both mis/disinformation and misinformation when speaking about both in a more general sense.

Case I: Anti-Abortion Disinformation as a Threat to Public Safety

In the summer of 2015, the Center for Medical Progress (CMP), an anti-abortion organization, became the center of a national news firestorm after releasing what was later identified as falsified videos which purportedly depicted Planned Parenthood officials selling fetal organs for profit. Before their debunking, however, these videos were disseminated by mainstream news media, who aided in legitimizing this disinformation

campaign through their incessant coverage. Even after their debunking, conservative American media continues to promote these lies – lies that led to unspeakable violence (Harsanyi, 2019).

On Friday, November 27, 2015, workers of Colorado Springs Planned Parenthood clinic anticipated another routine day of providing medical care and treatment to the town's residents. Yet only months after these widely disseminated videos spread throughout mainstream news media, anti-abortion extremist Robert Lewis Dear turned an average day into a nightmarish hellscape that left three dead and nine wounded.

The following research question is explored:

- How does Planned Parenthood related disinformation disseminated across mainstream news media pose a threat to public safety?

Framework: Agenda-Setting Theory

McCombs & Shaw (1972) described the “agenda-setting function of the mass media” as a way to influence its audience. The foundation of this theory is the notion that the press or mass media “may not be successful much of the time in telling people *what* to think, but is stunningly successful in telling its readers what to think *about*” (McCombs & Shaw, 1972, p. 177). As Dearing & Rogers (1996) described, salience – “the degree to which an issue on the agenda is perceived as relatively important” – is key in agenda-setting (p. 8).

The agenda-setting theory is particularly useful in understanding the ways in which mass media influences public perception. For example, when mainstream news media cover particular stories more frequently and fervently, viewers perceive them as important. The more coverage of an event or issue, the more important the public believes that event or issue is. This becomes especially dangerous as it relates to the illusory truth effect, which relates to our proclivity to believe false information after repeated exposure (Hasher & Goldstein, 1977). Together, the agenda-setting theory and the illusory truth effect demonstrate why the influx of mis/disinformation across

mass media is alarming, illustrated by the rapid and incessant circulation of falsified videos targeting Planned Parenthood created by the Center for Medical Progress.

The Center for Medical Progress

In the summer of 2015, the Center for Medical Progress (CMP), an anti-abortion organization, released videos purporting to depict Planned Parenthood officials selling fetal organs, which were later identified as doctored videos. CMP's website identifies the organization as "a group of citizen journalists dedicated to monitoring and reporting on medical ethics and advances" who are "concerned about contemporary bioethical issues that impact human dignity" (CMP, 2019).

Founded by David Daleiden, previously "director of research" at Live Action, an anti-abortion organization that opposes abortion under all circumstances, Daleiden became the driving force behind CMP's summer 2015 operation against Planned Parenthood (Somashekhar, 2015). Daleiden managed to secure airtime on a number of media outlets, where he promoted his baseless claims about Planned Parenthood purportedly selling fetal organs for profit, using his falsified videos as evidence. Daleiden's inspiration for "investigating" Planned Parenthood is in itself a sham based on disinformation, or a "discredited 2000 campaign by Life Dynamics" (Kittel, 2015). As Media Matters for America described, Daleiden's strategy for CMP comes directly from the playbook of Life Dynamics – a 90's era anti-choice group that "also made wild allegations of illegal profiteering on fetal organs in order to whip up anti-choice media coverage" – which was a success, as demonstrated by the massive coverage CMP's doctored videos received (Kittel, 2015).

Mainstream News Media Coverage of CMP's Disinformation Campaign

With the help of mainstream news media, CMP's doctored videos quickly went viral during the summer of 2015. CMP's doctored videos received national attention on networks like CNN, Fox News, MSNBC, and USA Today (Planned Parenthood, 2015b). In fact, a Google search using a custom date

range (June 2015 – October 2015) with the search term “planned parenthood video” reveals the extensive reach of this coverage, which does not even account for the “on-air” programming across cable news (i.e. news sources like CNN, *The Guardian*, *The Hill*, *Vox*, *Fox News*, *The Federalist*, *Washington Post*, *New York Times*, and *ABC News*, among others, circulated these videos – see figures 1-3 below). Like the agenda-setting theory argues, this constant attention makes it challenging for people *not* to think about it. Due to the illusory truth effect, it becomes nearly impossible for people to weed out the lies from the truth. For example, American lawmakers fell victim to these falsehoods, as they began drafting bills to reduce or eliminate government funding to Planned Parenthood (Pradhan, 2015).



Figure 1. CNN disseminates CMP’s deceptive videos purporting to depict Planned Parenthood officials selling fetal organs (CNN, 2015)



Figure 2. USA Today disseminates CMP’s deceptive videos that purportedly depict Planned Parenthood officials selling fetal organs (USA Today, 2015)



Figure 3. C-SPAN hosts Daleiden to promote his doctored videos on behalf of CMP (Washington Journal, 2016)

Even *after* these videos were determined to be doctored and faked, the myth that Planned Parenthood “sells baby parts” has persisted. Months after the videos were released, Daleiden *admitted to deceptively doctoring the videos* (Planned Parenthood, 2015a). Yet more than four years later, in February 2019, *The Federalist* published an article titled “Media Are Still Trying to Pretend Planned Parenthood Didn’t Sell Baby Body Parts,” arguing that Daleiden only edited the videos as any editor would, to clean it up for public consumption (Harsanyi, 2019). Additionally, a Google search on Planned Parenthood selling baby parts would leave a person with little-to-no knowledge of the situation very confused, as CMP and other sources claiming the videos are real come up on the *first page* of the Google search. Though debunked, conservative news media continues perpetuating CMP’s lies and distortions, even after the deadly violence that resulted.

Anti-Abortion Disinformation: A Threat to Public Safety

In November of 2015, Robert Lewis Dear killed three and injured nine during his attack on the Colorado Spring Planned Parenthood clinic. Dear’s violence came just months after CMP’s doctored video release. As noted, the videos and claims were completely fabricated, but before being debunk, these lies reached mass audiences and continue to perpetuate. As we know based on our brain’s handling of disinformation, especially as it relates to the illusory truth effect, it is no surprise that even after the videos were

completely debunked, some still argue the authenticity of their claim. For example, as Media Matters for America reported, Dear referred to himself as a “warrior for the babies,” trying to ensure that there were “no more baby parts,” as shown below in figure 4 (Kittel, 2015).



Figure 4. Robert Dear’s defense describes him as wanting to stop Planned Parenthood from performing abortions (Hutchins, 2015)

Dear is not alone in his proclivity towards anti-abortion violence. Law enforcement agencies “noted an uptick in violence against reproductive health care providers” following the release of CMP’s videos (Kittel, 2015). Vox’s Emily Crockett posited: “it is clear that threats, vandalism, and violence against abortion providers and clinics have escalated since this summer, when anti-abortion activists released deceptively edited videos that accused Planned Parenthood of ‘selling baby parts” (Kittel, 2015). Similarly, according to Media Matters for America, by December of 2015, *at least* five Planned Parenthood facilities were attacked (Kittel, 2015). While Robert Lewis Dear is an extreme case, he is but one example that demonstrates how the dissemination of Planned Parenthood related disinformation across mass media has the ability to incite deadly attacks, threatening public safety in America.

Case II: Misinformation in Political Ads on Facebook as a Threat to Democracy

Beginning in mid-2020, then-President Donald Trump began spreading election-related misinformation across Facebook using political advertisements, falsely claiming the election was “rigged,” “stolen,” and “fraudulent” (Gogarty, 2020). These Facebook advertisements, paid for by the Trump administration, had the potential reach of millions of Americans on the social media platform.

Two months after the election, thousands of Americans invaded the United States Capitol in Washington, D.C., in an attempt to overthrow American democracy under the guise that the 2020 presidential election was rigged. As Joe Biden was set to be certified as the next President, rioters attacked police officers and threatened lawmakers, leading to five deaths and marking the first time in U.S. history since the Civil War that the country did not see a “peaceful transfer of power” (Polantz, 2021). In light of this event, the following research question is explored:

- How does election-related misinformation in political advertisements disseminated across Facebook pose a threat to democracy?

Framework: Media Cultivation Theory

Founded by George Gerbner, media cultivation theory (originally applied primarily to television media) is a useful theory in understanding the impact of mass media and social media in society. The term “cultivation” is used “to describe the independent contributions television viewing makes to viewer conceptions of social reality” (Gerbner, 1998, p. 180). As Gerbner explained, “mass-produced stories can reach [people] on the average of more than 7 hours a day” (Gerbner, 1998, p. 176). Fast-forward more than twenty years, we are now inundated with messages nearly every hour we are awake – mass media surrounds us.

Much like television, social media has infiltrated our lives – from our cell phones to our laptops, we have instant access to social media 24-hours a day. Because of this, communication theorists, such as Morgan, Shanahan, and Signorielli (2015) contend that cultivation theory is even more applicable in our “new media” environment (p. 674). Similarly, Nevzat (2018) examined how to “revive and adjust” cultivation theory to apply to social media, which is perhaps even more discrete in ‘cultivating our reality,’ as the content we consume is controlled by algorithms behind-the-scenes, meaning we as consumers do not have as much control over the social media content we encounter as we do other mass media platforms, like radio or television. But similar to television programming, social media sites, such as Facebook, promote and disseminate news stories, making it a new avenue for users to turn when seeking information.

Facebook as News Source

According to the Pew Research Center (2019), 7 out of 10 adults use Facebook. Of those 7 out of 10 adult users, 74% use Facebook at least once a day. Of the American adult Facebook users, at least 4 out of 10 relies on Facebook for their news diet (Gramlich, 2019). And, in 2017, Pew Research reported that as many as 45% of Americans received their news from Facebook (Gesenhues, 2017). As media cultivation theory would argue, the high frequency at which American adults use Facebook (and daily) leave us more susceptible to believing the messages it disseminates, whether true or not (Nevzat, 2018). Unfortunately, particular Facebook policies allow the company to disseminate mis/disinformation, as shown by their political advertisement exclusion.

Facebook’s Political Ad Exclusion & Targeted Advertising

The Internet and social media have allowed mis/disinformation to circulate like never before, and Facebook has become quite popular among those seeking to disseminate and spread mis/disinformation (Bradshaw &

Howard, 2019). While Facebook relies on independent fact-checkers to verify posts, the company has an exclusion for political advertisements. For non-political advertisements, Facebook policies “prohibit ads that include claims debunked by third-party fact checkers,” but politicians are exempt from this policy, due to another company policy “that prevent it from taking action on the speech of politicians.” This means that political advertisements are not included in the company’s fact checking policy, thus giving politicians a free pass to lie (Breland, 2019).

This becomes more alarming when coupled with Facebook’s targeting advertising. As Vaidhyanathan (2019) explained: “Currently, two people in the same household can receive different ads from the same candidate running for state senate. That means a candidate can lie to one or both voters and they might never know about the other’s ads,” which “limits accountability.” Furthermore, Vaidhyanathan (2016) described the harm this can do to election-integrity: “the potential for abuse is vast. An ad could falsely accuse a candidate of the worst malfeasance forty-eight hours before election day and the victim would have no way of knowing it even happened” (p. 179). Vaidhyanathan’s words serve as foreshadowing for the damage caused by the election-related lies spread through Trump’s Facebook advertisements.

Trump’s Facebook Advertisements

Then-President Donald Trump successfully used Facebook’s political ad-exclusion to disseminate his political advertisements calling the 2020 election into question. With over 500 advertisements leveraging unfounded claims of voter fraud, each set with a reach greater than one million, Trump’s message attacked the integrity of the 2020 presidential election, claiming the election was “fraudulent,” claiming the “radical left is stealing the most important election of our lives,” and asking supporters to “defend” him and the election, as shown below in figures 5-6.

Active
Started running on May 18, 2020
ID: 679348816188931

About social issues, elections or politics

Donald J. Trump
Sponsored • Paid for by DONALD J. TRUMP FOR PRESIDENT, INC.

The Radical Left is trying to STEAL THE MOST IMPORTANT ELECTION OF OUR LIVES, so President Trump is calling on you to join his White House Defense Team.

Democrats have proven they'll stop at nothing to try and take down OUR President. Just take a look at some of the ridiculous things they've already tried:

- ✓ Stuffing the ballot boxes with FAKE and FRAUDULENT votes
- ✓ Rigging the Democratic Party nomination in favor of Sleepy Joe
- ✓ Impeaching President Trump when he DID NOTHING WRONG
- ✓ Countless BASELESS Witch Hunts
- ✓ Attempting to keep President Trump OFF the ballots in California
- ✓ Weaponizing the coronavirus investigative committee against the President

Please contribute ANY AMOUNT in the NEXT HOUR to join the White House Defense Team and your gift will automatically be DOUBLE-MATCHED.



HELP PRESIDENT TRUMP FIGHT BACK

CONTRIBUTE

WWW.DONALD.J.TRUMP.COM/DONATE
HELP PROTECT OUR ELECTIONS
2X MATCH>>> [Learn More](#)

Amount spent (USD): <\$100
Potential Reach: >1M people

Active
Started running on May 18, 2020
ID: 685198065589984

About social issues, elections or politics

Donald J. Trump
Sponsored • Paid for by DONALD J. TRUMP FOR PRESIDENT, INC.

The Radical Left is trying to STEAL THE MOST IMPORTANT ELECTION OF OUR LIVES, so President Trump is calling on you to join his White House Defense Team.

DEFEND PRESIDENT TRUMP

CONTRIBUTE

WWW.DONALD.J.TRUMP.COM/DONATE
HELP PROTECT OUR ELECTIONS
2X MATCH>>> [Learn More](#)

Amount spent (USD): <\$100
Potential Reach: >1M people

Figure 5. An example of Trump's advertisements challenging the legitimacy of the election, with a potential reach of up to one million people (Gogarty, 2020)

Figure 6. An advertisement with a potential reach of one million, questioning the legitimacy of the election and asking supporters to "defend President Trump" (Gogarty, 2020)

Election-Related Misinformation in Political Ads Across Facebook: A Threat to Democracy

On January 6, 2021, the consequences of these election-related lies were fully realized, as the United States Capitol came under attack by a mob of insurrectionists who stormed the building in hopes of reversing the certification of Joe Biden as the next President. Some rioters were dressed in tactical gear, armed with assault rifles, zip ties, and handguns. Many rioters carried signs with phrases like “stop the steal” and “investigate election fraud” (as shown below in figures 7-9). Other rioters vandalized offices as they searched for congressmembers after breaching the building, and other rioters assaulted law enforcement officers, leaving hundreds injured (Fisher et al, 2021). Chants included “fight for Trump,” “stop the steal,” and “invade the Capitol” (Goodman & Hendrix, 2021). The events of the day left five dead, over 140 officers injured, and four law enforcement members (who defended the Capitol during the riots) dying by suicide in the days and months following the insurrection (Schmidt & Broadwater, 2021; Farley, 2021).



Figure 7. “Stop the steal” sign displayed at the insurrection on Jan. 6, 2021 (Bloomberg, 2021)



Figure 8. “Investigate election fraud” sign displayed at the insurrection on Jan. 6, 2021 (Siegel & Faulders, 2021)



Figure 9. “Stop the steal” and “Trump won big” signs displayed at the insurrection on Jan. 6, 2021 (Perrett, 2021)

By spreading election-related lies to convince Americans that the election was “stolen” or “fraudulent,” Facebook became a disseminator of mis/disinformation that resulted in deadly carnage. Using the lens of media cultivation theory, we know that mass media messages impact their audiences, and, with a large number of Americans *relying* on Facebook for news, it is imperative that the news in all forms – political advertisements included – are based in facts and reality, not lies and distortions. Politicians and political advertisements should not be held to a different standard than the rest of society, especially as it relates to election-related mis/disinformation, as is the case with Trump’s political advertisements. The spread

of mis/disinformation in political advertisements has the propensity for deadly consequences, as demonstrated on January 6, 2021, threatening American democracy.

Case III: COVID Misinformation Disseminated by Fox News as a Threat to Public Health

In late 2019, a new, deadly virus known as the coronavirus, or COVID-19, began to spread worldwide after originating in China. By early 2020, the entire world was impacted by the coronavirus pandemic. From the beginning, Fox News downplayed the severity of the virus, comparing it to the seasonal flu (Baragona, 2020). Fast-forward to the end of 2021, almost two years since COVID-19 struck in America, the network continued to disseminate misinformation about the virus. Through their effective use of conservative frames, Fox's pundits became a driving force in disseminating coronavirus misinformation across the mainstream American news network. This is especially dangerous as early research confirms a direct link between a belief in COVID-19 misinformation and non-compliance with public health guidelines (Roozenbeek, 2020; Lee, 2020).

Nearly two years into the pandemic, there are over 5 million deaths reported globally and over 750,000 reported deaths in America (as of November 8, 2021), with over 300,000 cases reported worldwide in the last 24-hours (New York Times, 2021; WHO, 2021). Still, Fox continues to disseminate coronavirus mis/disinformation. As such, the following research question is explored:

- How does COVID-19 misinformation disseminated by Fox News pose a risk to public health?

Framework: Framing Theory

According to George Lakoff (2014): "Frames are mental structures that shape the way we see the world. As a result, they shape the goals we seek, the plans we make, the way we act, and what counts as a good or bad

outcome of our actions” (p. 2). Kuypers (2009) described why frames are critical to our understanding of reality: “Frames are so powerful because they induce us to filter our perceptions of the world in particular ways, essentially making some aspects of our multi-dimensional reality more noticeable than other aspects” (p. 181). Lakoff (2014) highlighted the importance of language in framing: “Language activates frames” and “changes the way the public sees the world. Framing is about getting language that fits your worldview” (Lakoff, 2014, p. 2).

Fox’s Conservative Framing

Fox News’ effective use of framing allows the network to promote a conservative worldview. As noted, Lakoff (2014) explained framing is about language and ideas – “the language carries those ideas, evokes those ideas” rather than explicitly stating it outright. Fox strategically frames their stories to evoke the Fox worldview, using conservative language to appeal to a conservative mindset – in particular, anti-government, anti-science, anti-expert frames. According to Lakoff (2014), “conservative language activates circuitry for the conservative worldview,” and Fox News is notorious for using solely conservative frames (Lakoff, 2014, p. 54; Brock & Rabin-Havt, 2012).

According to cognitive science, “people think in terms of frames and metaphors—conceptual structures. The frames are in the synapses of our brains, physically present in the form of neural circuitry. When the facts don’t fit the frames, the frames are kept and the facts ignored” (Lakoff, 2014, p. 121). This helps, in part, to explain why facts are so ineffective in changing minds. As Lakoff (2014) described: “We may be presented with facts, but for us to make sense of them, they have to fit what is already in the synapses of the brain” (p. 15). As Brock & Rabin-Havt (2012) explained, Fox News understands exactly how to frame particular debates. For example, during the Obama Presidency, Bill Sammons, VP and managing editor at Fox, released a memo to the Fox staff regarding the healthcare debate. He said: “let’s not

slip back into calling it “public option.” He continued, “government option versus public options has different connotations and thus people responded to it differently,” showcasing how equipped the network is in effectively framing important issues (Brock & Rabin-Havt, 2012).

Fox pundits understand how to properly frame arguments, debates, and stories, using language to evoke conservative, right-wing ideologies. Fox also understands the power of repetition, which is important in framing (and to evoke the illusory truth effect, as repetition makes lies more believable). Along with effectively framing “news” stories to fit within a particular, conservative worldview, Fox is strategic in framing the organization as well. The name Fox “News” would likely lead people to believe the information they receive is from their platforms is, in fact, news. However, the network relies heavily on opinion disguised as fact.

Opinion As Fact

While Fox News brands itself with “news” in its name, the network relies primarily on talking heads and pundits to deliver opinion-driven, fact-less monologues. Opinion pieces framed as “news” allows Fox to operate as if it was promoting factual, reason-based information rather than misinformation, pure conjecture, and partisan opinions. Along with successfully using conservative framing to evoke the Fox News version of reality, the network also uses rhetoric to brand itself as “fair and balanced,” and as the only news source that will report on the stories that other mainstream news media fail to report on (because, according to Fox News, all other mainstream news media is biased and “liberal”).

Fox News successfully frames their coverage as “fair and balanced,” which was the network’s motto up until 2017, when it was replaced with “most watched, most trusted,” (Grynbaum, 2017). This framing has aided the network in peddling the notion that other networks are all “the liberal media,” biased, unfair, and untrustworthy. Furthermore, Brock & Rabin-Havt (2012) described the Fox News ‘playbook:’ the network frames stories as exclusive

to Fox News, because other mainstream networks are not reporting on them (because they are not true, but it feeds into this loop that Fox News is the only *real*, credible news source). It creates a frame that Fox is the authority, Fox is the most credible (or only credible) news channel. [Fox News] “shows how news and opinion is blurred – but opinion wins – and the lengths the network goes to ensure a devoted audience” (Lewis, 2019). Fox’s successful framing of the coronavirus pandemic as a threat to the economy, as a scheme created by scientists, or as an overreach of government in mandating public health guidelines aided in their success in promoting COVID-19 misinformation by further activating conservative frames.

Fox’s COVID Misinformation Problem

In 2017, during a study prior to the coronavirus pandemic, PolitiFact found that over 60% of the “news” on Fox News was either completely false or mostly false, making Fox News one of the largest disseminators of mis/disinformation across mainstream (news) media platforms – as compared to MSNBC at 45% and CNN at 22% (PolitiFact, 2017; Sharockman, 2014). What happens during a global pandemic when the majority of information on a mainstream news media network is primarily false? As figure 10, below, illustrates, the network was found spreading coronavirus-related misinformation 253 times over only a five-day timespan (Savillo, 2020). Furthermore, research conducted by Hamel et al (2021) found a direct correlation between “people’s trusted news source” and their “belief in COVID-19 misinformation,” naming Fox News among the top three networks with viewers who were more likely to believe COVID misinformation (along with One America News and Newsmax).

Instances of coronavirus misinformation on Fox News over five days

July 6 - 10, 2020; 6 a.m. to midnight

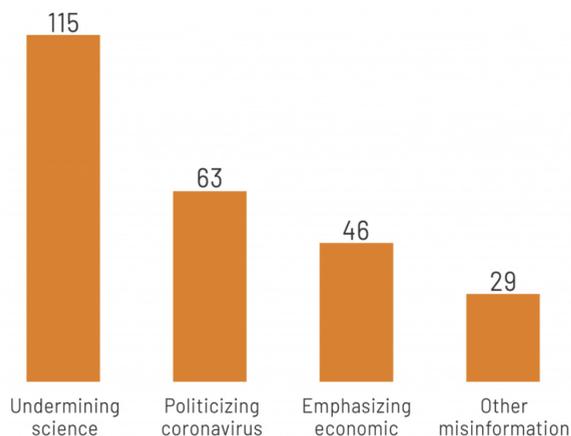


Figure 10. Media Matters for America's five-day study found 253 instances of Fox News spreading COVID misinformation (Savillo, 2020)

While Fox News disseminated science and health-related misinformation long before the coronavirus pandemic, COVID-19 brought a new wave of science and health-related misinformation (Pazzanese, 2020). And not surprisingly, Fox's decades-long anti-science and anti-government framing has assisted the network in promoting coronavirus misinformation. For instance, the following examples of coronavirus-related misinformation have been featured on Fox News, which is by no means exhaustive:

- Promoting false claims that the American government is overestimating the number of deaths related to COVID-19 (Brewster, 2020)
- Falsely claiming over 3,000 people have died from the coronavirus vaccine (McCarthy, 2021)
- Framing the COVID vaccine as "dangerous" and an example of personal freedom versus government overreach (Monroe & Savillo, 2021)
- Promoting unfounded claims that the coronavirus was intentionally leaked by scientists from a lab in China (O'Brien, 2021)

- Falsely claiming a “natural immunity” is more beneficial than the COVID vaccine (Media Matters for America, 2021)
- Framing mask mandates as an overreach of government power rather than a public safety measure (Porter, 2021)
- Framing health officials and scientists as untrustworthy as it relates to COVID-19 and the vaccine (Monroe & Savillo, 2021)
- Falsely claiming the COVID vaccine “may not work” and that the public is being lied to (by experts) (McCarthy, 2021b)
- Falsely claiming COVID is “not a major risk to children” (Media Matters for America, 2021)

To effectively spread misinformation, the network relies heavily on opinion framed as facts or news stories, distributed by their most notorious pundits: Laura Ingraham, Tucker Carlson, and Sean Hannity, the biggest purveyors of COVID-19 misinformation on the network.

Fox Pundits Disseminate COVID-19 Misinformation

The notion that COVID-19 safety measures are about control or government overreach, or the idea that scientists and experts cannot be trusted, plays into the credibility of non-experts, like Fox’s pundits, Ingraham, Carlson, and Hannity. These three pundits disseminated COVID-related misinformation using conservative frames, while peddling their opinions as facts (see figures 11-13 below).



Figure 11. Carlson framing COVID safety measures and vaccines as “about power and control not public health,” activating anti-government conservative frames (Porter, 2021)



Figure 12. Ingraham claims “experts” were wrong about COVID, activating anti-expert conservative frames (Perrett, 2020)



Figure 13. Hannity suggests Chinese scientists intentionally spread the coronavirus, activating anti-science and anti-expert conservative frames (Washington Post, 2020)

While originally identifying Sean Hannity as the biggest purveyor of coronavirus misinformation on Fox, Media Matters for America found during a weeks-long summer 2020 study that a quarter of the networks COVID-misinformation came from Laura Ingraham’s show. In addition, a 2020 study found a relationship between high COVID infection and mortality rates in areas with higher Sean Hannity viewership (Berkowitz, 2020). However, in the year since, Fox pundit Tucker Carlson has taken the lead in disseminating COVID misinformation, as he continues to claim that Dr. Anthony Fauci intentionally created COVID-19, and as he spreads a multitude of coronavirus conspiracy theories (McCarthy, 2021b). Pundits like Carlson, Hannity, and Ingraham feed their anti-science, anti-government viewers this idea that only *they* can be trusted – not experts or scientists, or other media outlets, and certainly not facts or reason.

Coronavirus Misinformation: A Threat to Public Health

As Funke & Sanders (2020) argued, “lies infected America in 2020. The very worst were not just damaging, but deadly,” naming coronavirus denial, downplay, and disinformation as PolitiFact’s “Lie of the Year,” an award dedicated to claims or statements “that prove to be of substantive consequence in undermining reality.” In addition to undermining reality, COVID-19 misinformation poses a direct threat to our collective public health, which is why the surplus of coronavirus misinformation disseminated by Fox News is so dangerous.

The spread of coronavirus misinformation becomes especially alarming as early research confirms a direct link between a belief in COVID-related misinformation and a non-compliance with public health guidelines. For example, research by Roozenbeek (2020) found a clear connection between susceptibility to misinformation and a reduced likelihood of complying with public health guidelines, like proper hand washing, social distancing, and wearing a mask (as well as vaccine hesitancy). This is echoed by Lee’s (2020) research, which concluded that belief in COVID-19 misinformation was associated with “fewer preventive behaviors.”

Fox’s strategic use of framing allows the network to be successful in disseminating mis/disinformation across its mass audience. According to National Security expert Garrett M. Graff, Fox News is especially dangerous because it leaves its audience incapable of making rational decisions: “its unique combination of lies and half-truths has built a virtual reality so complete that it leaves its viewers too misinformed to fulfill their most basic responsibilities as citizens” (Graff, 2019). With the coronavirus taking over 750,000 lives in America alone in less than two years – “a crisis exacerbated by the reckless spread of falsehoods,” (Funke & Sanders, 2020) – this pandemic has demonstrated that coronavirus misinformation disseminated through Fox News poses a direct threat to public health.

Conclusion

Whether disseminated across mainstream news media or across social media, mis/disinformation in mass media poses many dangers to American society. Effective solutions for combatting the spread of mis/disinformation must be created before this problem becomes too massive to solve. Unfortunately, as facts become “alternative,” and less likely to change our minds, we are tasked with finding new solutions for effectively communicating with those who consume misinformation across mainstream media and social media. The spread and circulation of mis/disinformation comes with both a hefty and dangerous price-tag, as it threatens our safety, our democracy, and even our health – not to mention our collective reality as a functioning society.

As research demonstrates, facts and logic are not substantial in changing people’s minds (Kolbert, 2017). Researchers suggest, “once formed, impressions are remarkably perservant” (Kolbert, 2017; Varol, 2017). Disturbingly, “even after the evidence for their beliefs has been totally refuted, people fail to make appropriate revisions in those beliefs,” (Kolbert, 2017). Because the media we consume influences our reality, when that media is filled with mis/disinformation, our reality becomes distorted as we begin to believe misinformation – due in part to the illusory truth effect and motivated reasoning. When facts and logic no longer change our minds, we must find alternative methods for solving this problem. This begs the question: how do we promote belief revision? How do we successfully reinform people who have been influenced by misinformation to the point that it impacts their beliefs and behavior? Ideally, through interdisciplinary research we can begin to uncover effective micro-level strategies for successfully communicating with people who have succumbed to mis/disinformation’s dangerous influence. However, macro-level strategies are incredibly necessary, as well.

Institutionalized, macro-level solutions must be implemented to combat the spread of dis/misinformation throughout American mass media. For example, O’Connor & Weatherall (2018) described how European countries

combat disinformation through the use of task forces. As they explained: “In 2015, the European Union created a specialized task force called East StratCom, with the express purpose of identifying strategies to combat fake news and propaganda” (O’Connor & Weatherall, 2018, p. 183). Similarly, effective regulation of mass media is extremely important in the United States, as the revocation of the Federal Communication Commission’s Fairness Doctrine – which required broadcasters to present issues in an honest, fair, balanced way – perhaps paved the way for the incessant dissemination of misinformation, along with the Telecommunications Act of 1996, which allows “anyone to enter the communications business” (FCC, 2019). More regulation in mass media is necessary to address the proliferation of misinformation – perhaps even a new doctrine through the FCC that requires *effective, non-partisan* fact-checking for both news media and social media platforms. As O’Connor & Weatherall (2018) explained, the “legislative framework that limits the ability of certain industries” has not yet carried over to the mass media, which is likely why the problem of misinformation has become so widespread, so lethal, and so in need of repair (O’Connor & Weatherall, 2018, p. 182).

Understanding how challenging it is to change our minds once we consume mis/disinformation and understanding the ways in which misinformation is disseminated across mass media, it is imperative that we find reasonable solutions for combatting its spread. In order to effectively diminish the proliferation of misinformation in mass media, *everyone*, including politicians, must be held to the same standard – that of honesty and authenticity – across all media platforms – television, radio, social media, newspapers, etc. Whether in the form of a disinformation campaign against Planned Parenthood disseminated throughout mainstream media, election-related misinformation in political advertisements disseminated on Facebook, or coronavirus misinformation disseminated by Fox New, failure to combat the spread of disinformation poses a risk to American public safety, public health, and democracy – and may result in a “post-truth society” that feels more like an alternate reality in an episode of *The Twilight Zone*.

References

- Arkes, H. R., Hackett, C., and Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, 2.
- Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, 5.
- Baek, Y.M., Kang, H. & Kim, S. (2019). Fake News should be regulated because it influences both “others” and “me”: How and why the influence of presumed influence model should be extended. *Mass Communication and Society*, 22, 301-323.
- Baragona, J. (2020, April 13). Fox News contributor compares coronavirus to the flu, claims it’s ‘not a pandemic.’ *Daily Beast*. <https://www.thedailybeast.com/fox-news-contributor-bill-bennett-compares-coronavirus-to-the-flu-claims-its-not-a-pandemic>
- Begg, I. M., Anas A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology*, 121.
- Berkowitz, J. (2020, June 25). Exposure to Sean Hannity may be lethal, studies on COVID-19 suggest. <https://www.fastcompany.com/90521195/exposure-to-sean-hannity-may-be-lethal-studies-on-covid-19-suggest>
- Bloomberg. (2021, January 11). Facebook bans ‘stop the steal’ posts ahead of inauguration. *LA Times*. <https://www.latimes.com/business/technology/story/2021-01-11/facebook-bans-stop-the-steal-posts-ahead-of-inauguration>
- Bradshaw, S. & Howard, P.N. (2019). The global disinformation order: 2019 Global inventory of organised social media manipulation. Project on Computational Propaganda. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf>
- Breland, A. (2019, October 3). Facebook just gave Trump permission to lie. *Mother Jones*. <https://www.motherjones.com/politics/2019/10/facebook-ads-biden-misinformation/>

- Brewster, J. (2020, April 9). A Fox News conspiracy—Are coronavirus death numbers inflated? *Forbes*. <https://www.forbes.com/sites/jackbrewster/2020/04/09/a-fox-news-conspiracy-are-coronavirus-death-numbers-inflated-attacked-by-fauci-birx>
- Brock, D. & Rabin-Havt, A. (2012). *The Fox Effect: How Roger Ailes turned a network into a propaganda machine*. Anchor Books.
- Center for Medical Progress (CMP). (2019). About us. <https://www.centerformedicalprogress.org/about-us/>
- Chang, Y. & Luo, Y. (2021). The presumed influence of digital misinformation: Examining US public's support for governmental restrictions versus corrective action in the COVID-19 pandemic. *Online Information Review*, 45(2), 834-852.
- CNN. (2015, July 21). Planned Parenthood Sting Video. *CNN*. <https://www.cnn.com/videos/us/2015/07/21/planned-parenthood-sting-video-foreman-live-lead.cnn>
- Dearing, J. & Rogers, E. (1996). *Agenda-setting*. Sage Publishing.
- Farley, R. (2021, November 1). How many died as a result of Capitol riot? *FactCheck.org*. <https://www.factcheck.org/2021/11/how-many-died-as-a-result-of-capitol-riot/>
- Federal Communications Commission (FCC). (2019). Telecommunications Act of 1996. *FCC*. <https://www.fcc.gov/general/telecommunications-act-1996>
- Fisher, M., Flynn, M., Contrera, J., Loennig, C.D. (2021, January 7). The four-hour insurrection: How a Trump mob halted American democracy. *Washington Post*. <https://www.washingtonpost.com/graphics/2021/politics/trump-insurrection-capitol/>
- Funke, D. & Sanders, K. (2020, December 16). Lie of the year: Coronavirus downplay and denial. *PolitiFact*. <https://www.politifact.com/article/2020/dec/16/lie-year-coronavirus-downplay-and-denial/>
- Geraerts E., Bernstein, D.M., Merckelbach, H., Linders, C., Raymaekers, L., Loftus, E.F. (2008). Lasting false beliefs and their behavioral consequences. *Psychological Science*, 19(8), 749-753.

- Gerber, G. (1998). Cultivation analysis: An overview. *Mass Communication & Society*, 3(4).
- Gesehues, A. (2017, November 8). Pew Research Center says 45% of Americans get their news from Facebook. *Marketing Land*. <https://marketingland.com/pew-research-center-says-45-americans-get-news-facebook-228001>
- Gogarty, K. (2020, May 19). Facebook is letting the Trump campaign publish at least 529 ads with false claims of voter fraud. *Media Matters for America*. <https://www.mediamatters.org/facebook/facebook-letting-trump-campaign-publish-least-529-ads-false-claims-voter-fraud>
- Goodman, R. & Hendrix, J. (2021, January 25). ‘Fight for Trump:’ Video evidence of incitement at Capitol. *Just Security*. <https://www.justsecurity.org/74335/fight-for-trump-video-evidence-of-incitement-at-the-capitol/>
- Graff, G.M. (2019, December 11). Fox News is now a threat to national security. *Wired*. <https://www.wired.com/story/fox-news-is-now-a-threat-to-national-security/>
- Gramlich, J. (2019, May 16). 10 facts about Americans and Facebook. *Pew Research Center*. <https://www.pewresearch.org/fact-tank/2019/05/16/facts-about-americans-and-facebook/>
- Grynbaum, M. (2017, June 14). Fox News drops ‘fair and balanced’ Motto. *New York Times*. <https://www.nytimes.com/2017/06/14/business/media/fox-news-fair-and-balanced.html>
- Hamel, L., Lopes, L., Kirzinger, A., Sparks, G., Stokes, M. & Brodie, M. et al. (2021, November 8). KFF COVID-19 vaccine monitor: Media and misinformation. *Kaiser Family Foundation*. <https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-media-and-misinformation>
- Harsanyi, D. (2019, February 20). Media are still trying to pretend Planned Parenthood didn’t sell baby body parts. *The Federalist*. <https://thefederalist.com/2019/02/20/media-still-trying-pretend-planned-parenthood-didnt-sell-baby-body-parts/>

- Hasher, L. and Goldstein, D. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16.
- Hussain, Azhar, Syed Ali, Madiha Ahmed, & S. Hussain. (2018). The Anti-vaccination movement: A regression in modern medicine. *Cureus*, 10(7).
- Hutchins, C. (2015, November 29). 'Baby parts' or other explanations for motive in Planned Parenthood shooting? *Colorado Independent*. <https://www.coloradoindependent.com/2015/11/29/robert-lewis-dear-motive/>
- Jonas, E., Schulz-Hardt, S., Frey, D., Thelen, N. (2001). "Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information." *Journal of Personality and Social Psychology*, 80(4).
- Kittel, O. (2015, December 17). Misinformers of the year: The Center for Medical Progress. *Media Matters for America*. <https://www.mediamatters.org/planned-parenthood/misinformers-year-center-medical-progress>
- Kolbert, E. (2017, February 19). Why facts don't change our minds. *The New Yorker*. <https://www.newyorker.com/magazine/2017/02/27/why-facts-dont-change-our-minds>
- Kuypers, J.A. (2009). *Rhetorical Criticism: Perspectives in Action*. Lexington Books.
- Lakoff, G. (2014). *The ALL New: Don't think of an elephant! Know your values and frame the debate*. Chelsea Green Publishing.
- Lee, J.J., Kang, K.A., Wang, M.P., Zhao, S.Z., Wong, J.Y.H., O'Connor, S., Yang, S.C. & Shin, S. (2020). Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: Cross-sectional online study. *Journal of Medical Internet Research*, 22(11).
- Lewis, A. & Marwick, A. (2017). Taking the red pill: Ideological motivations for spreading disinformation. Annenberg School for Communication. (December 15-16, 2017). Understanding and Addressing the Disinformation Ecosystem. [Conference].

- Lewis, B. (2019, October 25). I watched Fox News every day for 44 months – here’s what I learned. *The Guardian*. <https://www.theguardian.com/media/2019/oct/25/fox-news-watching-what-i-learned>
- Mangan, D. (2021, July 19). Trump says of Capitol rioters who sought to overturn Biden’s win: ‘What I wanted is what they wanted.’ *CNBC*. <https://www.cnn.com/2021/07/19/trump-says-he-and-capitol-rioters-wanted-same-thing-overturning-biden-win.html>
- McCarthy, B. (2021, May 7). Tucker Carlson said 3,362 people have died from the COVID vaccine. *Statesman*. <https://www.statesman.com/story/news/politics/politifact/2021/05/07/tucker-carlson-misleading-claim-deaths-after-covid-19-vaccine-covid/4989132001/>
- McCarthy, B. (2021b, May 15). Tucker Carlson falsely claims COVID-19 vaccines might not work. *PolitiFact*. <https://www.politifact.com/fact-checks/2021/apr/15/tucker-carlson/tucker-carlson-falsely-claims-covid-19-vaccines-mi/>
- McCombs, M. & Shaw, D. (1972). The agenda-setting function of mass media. *The Public Opinion Quarterly*, 36(2).
- Media Matters for America. (2021, July 13). Laura Ingraham promotes “natural immunity” over the COVID vaccine. <https://www.mediamatters.org/fox-news/laura-ingraham-promotes-natural-immunity-over-covid-vaccine>
- Morgan, M., Shanahan, J., & Signorielli, N. (2015). Yesterday’s new cultivation, tomorrow.” *Mass Communication & Society*, 18.
- Monroe, T. & Savillo, R. (2021, July 16). Fox has undermined vaccination efforts in nearly 60% of all vaccination segments in a 2-week period. *Media Matters for America*. <https://www.mediamatters.org/fox-news/fox-has-undermined-vaccination-efforts-nearly-60-all-vaccination-segments-2-week-period>
- New York Times*. (2021, November 8). Coronavirus in the US: Latest map and case count. <https://www.nytimes.com/interactive/2021/us/covid-cases.html>

- Nezvat, R. (2018). Reviving cultivation theory for social media. The Asian Conference on Media, Communication & Film 2018. International Academic Forum.
- Nyhan, B. & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–30.
- O'Brien, C. (2021, May 24). As reporting on coronavirus lab leak theory grows, critics accuse media of suffering 'amnesia' on topic. *Fox News*. <https://www.foxnews.com/media/critics-coronavirus-lab-leak-theory-critics-accuse-amnesia>
- O'Connor, C. & Weatherall, J.O. (2018). *The Misinformation Age*. Yale University Press.
- Patel, M. Lee, A.D, Clemmons, N.S., Redd, S.B., Poser, S., Blog, D., et al. (2019, October 4). National Update on Measles Cases and Outbreaks — United States, January 1–October 1, 2019. *Center for Disease Control and Prevention*. <https://www.cdc.gov/mmwr/volumes/68/wr/mm6840e2.htm>
- Pazzanese, C. (2020, October 30). Why isn't the right more afraid of COVID-19? *Harvard Gazette*. <https://news.harvard.edu/gazette/story/2020/10/what-caused-the-u-s-anti-science-trend/>
- Pennycook, G., Cannon, T., & Rand, D. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology*, 147(12).
- Perrett, C. (2020, May 12). Fox News is telling its employees to work from home for at least another month while its TV hosts push for businesses to reopen. *Business Insider*. <https://www.businessinsider.com/fox-news-closed-while-on-air-talent-push-reopen-2020-5>
- Phillips, B.K., Prybutok, V.R., and Peak, D.A. (2014). Decision confidence, information usefulness, and information seeking intention in the presence of disconfirming information. *Informing Science: The International Journal of an Emerging Transdiscipline*, 17.

- Planned Parenthood. (2015, October 21). Daleiden admits to deceptively editing videos. <https://www.plannedparenthood.org/about-us/newsroom/press-releases/daleiden-admits-to-deceptively-editing-videos>
- Planned Parenthood. (2015b, December 31). Memo: Deceptive videos about Planned Parenthood. <https://www.plannedparenthood.org/about-us/newsroom/press-releases/deceptive-videos-about-planned-parenthood>
- Polage, D.C. (2012). Making up history: False memories of fake news stories. *Europe's Journal of Psychology*, 8(2).
- Polantz, K. (2021, October 30). New January 6 court filings reveal what Trump is trying to keep secret from Congress. *CNN*. <https://www.cnn.com/2021/10/30/politics/donald-trump-house-democrats-january-6-documents/index.html>
- PolitiFact*. (2017). Fox's File. <https://www.politifact.com/punditfact/tv/fox/>
- Porter, T. (2021, February 23). How Fox News hosts started attacking COVID-19 vaccines after Biden took over the rollout from Trump. *Business Insider*. <https://www.businessinsider.com/fox-news-hosts-vaccines-biden-trump-2021-2>
- Pradhan, R. (August 8, 2015). Defund Planned Parenthood? Not so easy. *Politico*. <https://www.politico.com/story/2015/08/defund-planned-parenthood-not-so-easy-120924>
- Roozenbeek, J., Schneider, C.R., Dryhurst, S., Kerr, J., Freeman, A.L.J., Recchia, G., van der Bles, A.M., van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 7(10).
- Savillo, R. (2020, July 16). Fox News pushed coronavirus misinformation 253 times in just five days. *Media Matters for America*. <https://www.mediamatters.org/coronavirus-covid-19/fox-news-pushed-coronavirus-misinformation-253-times-just-five-days>

- Schmidt, M.S. & Broadwater, L. (2021, February 11). Officers' injuries, including concussions, show scope of violence at Capitol riot. *New York Times*. <https://www.nytimes.com/2021/02/11/us/politics/capitol-riot-police-officer-injuries.html>
- Sears, D.O. & Freedman, J.L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, 31(2).
- Siegel, B. & Faulders, K. (2021, September 10). House committee investigating Capitol riot reviewing thousands of documents. *ABC News*. <https://abcnews.go.com/Politics/house-committee-investigating-capitol-riot-reviewing-thousands-documents/story?id=79951203>
- Sharockman, A. (2014). Fact-checking Fox, MSNBC, and CNN: PunditFact's network scorecards. *PolitiFact*. <https://www.politifact.com/punditfact/article/2014/sep/16/fact-checking-fox-msnbc-and-cnn-pundit-facts-networ/>
- Somashekhar, S. (2015, October 14). Meet the millennial who infiltrated the guarded world of abortion providers. *Washington Post*. https://www.washingtonpost.com/national/meet-the-millennial-who-infiltrated-the-guarded-world-of-abortion-providers/2015/10/14/25aaf862-678b-11e5-9223-70cb36460919_story.html
- Tandoc, Jr., E.C. (2019). The facts of fake news: A research review. *Sociology Compass*, 13.
- USA Today. (2015, July 30). Planned Parenthood under fire over undercover video. <https://www.usatoday.com/videos/news/nation/2015/07/30/30199123/>
- Vaidhyanathan, S. (2019, November 2). The real reason Facebook won't fact-check political ads. *New York Times*. <https://www.nytimes.com/2019/11/02/opinion/facebook-zuckerberg-political-ads.html>
- Vaidhyanathan, S. (2016). *Anti-Social Media: How Facebook Disconnects Us and Undermines Democracy*. Oxford University Press.
- Varol, O. (2017, September 8). Facts don't change people's minds. Here's what does. <https://heleo.com/facts-dont-change-peoples-minds-heres/16242/>

- Wardle, C. & Derakhshan, H. (2017). "Information disorder: Towards an interdisciplinary framework for research and policymaking." *Council of Europe*. <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>
- Washington Journal*. (2016, January 22). Planned Parenthood Videos. <https://www.c-span.org/video/?403139-4/washington-journal-david-daleiden-planned-parenthood>
- Washington Post*. (2020, May 23). Fox News has never been so right. <https://www.washingtonpost.com/opinions/2020/05/23/fox-news-has-never-been-so-right/>
- World Health Organization. (2021, November 8). COVID dashboard. <https://covid19.who.int/>

WHEN COMMUNICATION MEETS INTERNATIONAL RELATIONS PERSPECTIVES: UNDERSTANDING DISINFORMATION IN A MULTICENTRIC POLITICAL ENVIRONMENT

Alessandra Massa & Giuseppe Anzera

Introduction. Towards a Reality-Politik?

In this theoretical chapter, we explore how some interpretive concepts tracing back to the analytical tradition of International Relations (hereafter IR) can help explain the rising phenomenon of disinformation in our current political environment. Information plays an increasingly decisive role in political processes, affecting domestic issues and rewriting the reputation of actors involved in international affairs. Electoral confrontations, internal conflicts, and international disputes are just some of the phenomena affected by the shock wave of disinformation. *Reality* – intended as a set of collectively recognizable and objectively definable phenomena – has progressively evolved into an object of discussion within the discipline. Therefore, if the analytical current of realism is motivated by describing reality, subsequent theoretical approaches have challenged the validity of such a description. Political phenomena can be understood as the result of social construction (Anzera & Massa, 2021a).

The discussion about *factuality versus the reality* of IR revolves partly around the suggestions of constructivism. In the field of IR, constructivism is a theoretical and interpretive approach that challenges many mainstream assumptions. The latter provides explanations

or venture into predictive hypotheses that assume that states' relationship and behavioural attitudes are inherent. Constructivism claims that actors and agency, interests, and structures are defined by social norms and ideas and not exclusively by objective or material conditions (Barkin, 2003). Constructivist analyses of international politics have questioned both the social construction of meaning (including knowledge) and the construction of social reality (Guzzini, 2000). This approach suggests that limits may exist concerning what we know about reality: does it exist *per se*, and therefore can it be analysed empirically, or can it not be definable outside our point of observation?

Power within the international realm is an issue of classical realism, and the empirical studies conducted within it are designed to understand its determinants (Morgenthau, 1948; Waltz, 1979). Furthermore, it is essential to examine whether this power can be perceived in its tangible forms. It must be represented, explained, and exercised. Constructivism intervenes in the mechanisms that come into play in the social construction of international reality (Wendt, 1992). How that reality is portrayed in institutional venues and popular representations has consequences for how that reality is experienced, with tangible implications for the unfolding events of IR. Guzzini (2000) suggests that the Second Gulf War discourses exemplify reality-making operations. Narratives have shaped the conflict through the use of analogies. This narrative process has involved portraying enemies, causes, and legitimization. Often reductionist, if at times evocative, narratives, rather than explaining reality, mythologize it. Political efforts to make storytelling understandable lead to events being turned into narratives that blur the lines between fact and fiction. During conflicts, media products contribute to creating new realities (Virilio, 2002), in which hybridizations among political, military, industrial, and entertainment apparatuses (Der Derian, 2009) affect a composite system of representations manipulated by (powerful) parties. Then, the constant attempt by interested actors to redefine the role of truth in international politics characterizes the contemporary era. In the last decades, the interpretive tradition of IR has focused on a variety of subjects and practices other than the traditional focus on

tangible or material aspects of power politics. The emergence of new actors entitled to intervene in IR (civil society actors, large corporations, and media companies are some of the new global players) defines brand-new forms of interaction. The perceived marginality of material capabilities and warfare implied a revised definition of power, focusing on the softer, intangible aspects. International events have gained a place in citizens' informative consumption due to technological innovations and changes in productive routines. These changes have affected professional journalism and content production. Political actors recognize the importance of the display of international politics in public opinion and institutional consultations. Political power is asserted in the link between reality and truth (Anzera & Massa, 2021b).

The World We Live In. Actors, Power, Information: A Few Definitions

In this section, we will describe the topics we will cover in the chapter to explain contemporary disinformation. The definition of what *International Relations* (the object of a specific discipline) means is a necessary premise. IR can be broadly defined as the external relationships between sovereign states. Traditional topics of IR include wars, economic agreements, arms control, environmental issues, and other pertinent matters (Reus-Smit, 2020).

While this definition seems broad, it is narrow when applied to contemporary political phenomena: IR must consider actors besides states today. However, some domestic events also impact the international order – consider the U.S. and UK electoral consultations over the past few years. Therefore, it is difficult to distinguish the domains of IR and the actors entitled to participate in them. Sovereign states, at a standard level of abstraction, are the main characters of IR. Nevertheless, the same scientific definitions of these actors, so central and complex, seem to be blurring as territoriality faces multiple challenges. The cohesion and recognition of different sociocultural units (what we commonly refer to as nations or national communities) are also changing. Who constitutes those units? As individuals, we are continuously influenced by IR – although we are not

always aware of its weight in current affairs – while we can also act and push back against IR processes. Because of this, “international relations’ have no frontiers traced out in reality, they are not and cannot be materially separable from other social phenomena” (Aron, 1966, p. 4).

Because of their complexity, debates on the core issues of IR are characterized by a multifaceted truth. IR topics are sometimes far from public attention, while political positions are not always clear and unequivocal. While IR events are often long-lasting, the media portrays them as sudden and isolated. So, politically interested actors may intervene in the public account of international affairs.

During the discussion, we will define disinformation as a general climate of manipulating information by politically motivated actors. Disinformation is the systematic collection of media texts – a more or less organized compilation of *problematic content* – intentionally produced to be deceptive (Bennett & Livingston, 2020). The goal of disinformation in international information is to make public and political actors’ narratives questionable and arouse general distrust. The term disinformation should be preferred over the term fake news: the latter describes a broad range of phenomena, from satire to commercial information (Tandoc et al., 2018). Moreover, this label risks not capturing the complexity of *information disorder*. Different types of information are intertwined, with varying adherence to reality and different intentions to cause harm (Wardle & Derakhshan, 2018).

Disinformation is a phenomenon that brings IR back into the knowledge of generalist audiences (albeit, at times, stereotypically). As evidenced by news outlets and academic research, Russia and China’s geostrategic interests emerged when their interference in online platforms and information spaces was exposed. However, the political weight of online platforms’ infrastructure (Plantin et al., 2018) and mediation (Van Dijck, 2013) is not always transparent to users. Despite their global importance (Van Dijck et al., 2018), these platforms are accused of promoting disinformation through their technical components. By reconstructing the ruthless mechanisms

of Facebook, Frenkel and Kang (2021) demonstrate how the algorithmic logic follows profit maximization and suppresses ethical business disagreements. However, the management of tensions with platforms results from the expression of new and old geopolitical fractures (Van Dijck et al., 2018).

The chapter is organized as follow. As a guide for understanding the text, the next section presents a series of definitions. A section of the paper will explore a number of classic IR themes that help place disinformation within an understanding of IR not solely based on material resources. Multicentrism and skills revolution as formulated by Rosenau (1990); Keohane and Nye's concept of complex interdependence (2012[1977]); and soft power (Nye, 1990) will be discussed. Next, we will discuss the cultural and narrative turns in IR over the past few decades. Lastly, Mearsheimer's (2013) reconstruction of lies in international politics will be used to illustrate a realist approach.

I. A Toolkit for Understanding the Role of (Dis)Information

Powerful people navigating the turbulence

Since World War II, global politics has been experiencing a tumultuous period that Rosenau (1990) describes as *turbulence*. It identifies post-international politics as a result of the inadequacy of state-centric politics to encompass and summarize the trends of the contemporary globalized world, which is characterized by profound and continuous transformations. In particular, the actors that participate in the political games have changed. The issue of international politics has become a topic of discussion among politicians and citizens, and many different subjects have created connections and relationships.

The breaking points are threefold: 1) the historical and political alterations that caused the complication of the post-war order; 2) the bifurcation of global macro structures that have led to the redesign of "two worlds of global politics", in which multicentrism challenges state-centered politics; and 3) the micro-level revolutions that have emerged, expanding individual emotional and cognitive capacities toward global politics.

These factors are combined with what Rosenau defined in the Nineties as *modern technologies*, namely microelectronics. By reducing time and shrinking space, these technologies dematerialize traditional economic processes and accelerate the transmission of information.

As the word suggests, *turbulence* is synonymous with *uncertainty* in international politics and its evolving structures. Due to this uncertainty, analytical and cognitive tools can prove inadequate for explaining a multifaceted world that often defies traditional categorization. There is a relevant clue in the increasing number of *turns* in the IR discipline's narrative, which we will discuss later.

When new problems emerge, states often cannot provide satisfactory answers. Most solutions require multilevel governance tools that are sophisticated and unprecedented, as well as confrontations with unexpected actors. Meanwhile, citizens are becoming less likely to blindly accept state-provided solutions and their traditional operative frameworks as a result of the development of individual skills. As individuals' abilities evolve, issues of competitive information have become central to managing IR's non-material dimension, escalating in some forms of contemporary disinformation. Individuals develop analytical skills and the ability to master complex cognitive maps. In evaluating the legitimacy of actors in power, performance is emphasized rather than political loyalty, which is variable and not always focused on the state.

Rosenau prefigures a scenario where active citizens refer to their skills to decode world politics. Media intermediation facilitates citizen participation in single-issue networks constituted by (sometimes) organized but temporary ties. Using technological devices, in this scenario, macro issues (those that pertain to codified formal politics) and micro issues (relating to individual expressions and autonomous positioning in the world) are connected in unprecedented ways:

[...] new technologies have had a profound, if not always desirable, impact upon how individuals perceive, comprehend, judge, enter, avoid, or

otherwise interact with the world beyond their workplace and home. For example, the new electronic technologies have so greatly collapsed the time in which organizations and movements can be mobilized that the competence of citizens feeds on itself, in the sense that they can virtually 'see' their skills and orientations being cumulated into larger aggregates that have consequence for the course of events (Rosenau, 1990, p. 15).

The insights described above suggest that mediating tools, such as online platforms, can enhance individual connections and influence politics, generating unexpected effects, as witnessed during the Arab Spring uprisings (Comunello & Anzera, 2012; Wolfsfed et al., 2013). Technology plays a dual role since it is also a transformative force for elites. Rosenau depicts citizens as *powerful people* with the right to participate in world politics. The same political elites, however, develop new skills, resulting in practices similar to contemporary astroturfing (Kovic et al., 2018; Keller et al., 2020):

Not only are 'ordinary' people often interviewed for their reactions, but in all parts of the world, they are frequently asked to turn out for rallies or protests that are scheduled for live television coverage. Or, if they have spontaneously converged upon a site and formed a leaderless public, leaders are quick to seize the chance to get them on the global stage by calling the attention of the mass media to their presence, thereby transforming them into a participatory aggregation that can serve their shared purposes. Irrespective of the ways in which they may have been stage-managed, however, the participants are likely to experience such occasions as moments when they were actors on the global stage (Rosenau, 1990, p. 344).

As global politics becomes more uncertain, authority and its decisions are questioned, and citizens' abilities to interpret contemporary phenomena grow, new processes are being cultivated. These processes affect the grip of misinformation on modern society. Due to citizens' autonomy and inclination to contest power, they tend to distrust the official political system and

its sources of information. Despite overestimating their analytical skills, citizens are always looking for alternative sources. This overestimation could undermine information management capacity in an increasingly complex political environment. Citizens' autonomous and emotional discernment is the key to transforming them into *affective publics* (Papacharissi, 2015), and this judgment is sometimes seen as an aggregate that can be modulated based on fluctuating interests. Exaggerated news often targets such audiences, playing on their sense of credulity and encouraging oppositional and populist instincts.

II. Complex interdependence and information power

Originally published in 1977, "Power and Interdependence" by Keohane and Nye has undergone several revisions to test and confirm its interpretive validity¹. Complex interdependence offers a powerful framework for explaining how interconnections among different players, fueled by economic relations, have contributed to describing flows and destabilizations in the world order.

Complex interdependence departs from the realm of realism. All economic ties that affect the relations between nations cannot be explained by the belief in the inevitable outcome of wars and the perception of perpetual competition and ruthlessness. Instead, complex interdependence involves the coexistence of three dimensions:

1. There are multiple communication channels between societies, including those developed at the informal level. Governments, political elites, and international organizations (including non-state actors and large corporations) participate in these channels.
2. State-to-state relations include multiple, non-hierarchical issues, and military matters do not necessarily dominate them.

1. In the discussion, we will refer to the fourth edition, published in 2012.

3. A large army cannot resolve all problems: for example, it cannot combat economic crises or climate chaos.

By deconstructing the phenomena associated with interdependence, some preconditions that have contributed to the rise of misinformation can be clarified. Media are allowed to engage in politics as actors other than the state intervening in international affairs. Additionally, the proliferation of relevant political issues has dramatically expanded the topics covered by legacy and digital media. This has required an increased focus on the part of audiences. Finally, this setup determines the actors' unpredictability in the political game. It is possible for anyone with organizational and economic strength, driven by interests, to actively participate in international debates, at least theoretically.

The marginalization of military forces influences conflict representation. The role of armies in international disputes appears and disappears. The rise of arrested wars results from political leadership's willingness to display conflicts (Hoskins & O'Loughlin, 2015). Mediatization (Hepp, 2019) encourages policymakers to consider the media in policy formulation and presentation. This centrality of representation is particularly problematic for events that exacerbate conflict and, at the same time, constitute a second-hand reality for most people. As well as amplifying the nonlinearity of political events, this interdependence can be viewed as a tangled network of flows. Conflicts are difficult to hide from public opinion today, but their understanding can be arrested by increasing the fragmentary nature of confrontations and creating gray areas where disinformation can thrive (Hoskins & O'Loughlin, 2015). For example, how the Syrian conflict has been portrayed to Western viewers has been marked by uncertainty, deliberate disinformation, and discrediting sources (Merrin, 2018).

Keohane and Nye suggest that the decline of traditional security management has two main consequences: first, it implies a flexible kind of power quite different from the traditional concept. Two, political interests can be domestic, transnational, or governmental, suggesting a variety of

presidiums, including the one for information. Two conditions are essential when interdependence arrives. The first is the *sensitivity* associated with public opinions' reactivity to political, social, and economic themes and issues, regardless of governmental actions. A second condition is actors' vulnerability to external interference, such as informational interference (Chong, 2007).

The increased use of information and interdependence will not erase traditional power struggles, as Keohane and Nye are aware:

As traditionalists maintain, much will be the same: states will play important roles; vulnerability will lead to bargaining weakness and lack of vulnerability to power; actors will seek to manipulate cyberspace, as they manipulate flows across borders, to enhance their power. Yet as modernists insist, the information revolution is not 'déjà vu all over again': cyberspace is truly global; it is harder to stop or even monitor the flow of information carrying electrons than to do so for raw materials or goods; and dramatic reductions in the cost of information transmission make other resources relatively scarce (Keohane & Nye, 2012, p. 212).

A state can, however, cultivate its power through the competitive role of information: "the information revolution creates a new politics of credibility in which transparency will increasingly be a power asset" (Keohane & Nye, 2012, p. 213). According to the authors, there are three types of information.

1. *Free information.* All forms of information produced and distributed without charge fall under this category. When recipients believe what is said and consider the information reliable, those with this information gain an advantage.
2. *Commercial information.* When it comes to this type of information, whoever arrives first sets standards and acquires advantages (Castells, 1996). Modern examples include the ability to interact with online platforms. In recent electoral propaganda, a symbolic gain is achieved through occupying commercial information spaces before competitors.

3. *Strategic information*. It is only when a competitor does not possess this kind of information that it generates advantages. Using listening techniques, A/B testing methods, and advanced targeting can lead to asymmetric information about competitors, as was demonstrated by Cambridge Analytica (Vaidhyanathan, 2018).

The information revolution must not be welcomed with naive enthusiasm. Emerging actors are not automatically favoured by the production of information as a competitive tool. Rogue formations can also benefit from political disorder. Arquilla and Rondfeldt (2008) claim that organized oppositional and terrorist movements can destabilize the *noopolitik* arrangements. In a digital world that favours economies and scale distribution (Van Dijck & Poell, 2013; Van Dijck et al., 2018), states do not disappear. However, more organized and well-structured actors are favoured (Castells, 2009).

Moreover, “it is usually better to be a first-mover than a fast follower” (Keohane & Nye, 2012, p. 217). Therefore, those who reach international audiences first can expect better results. However, misleading news is one of the primary strategies for eroding trust and credibility when novelty prevails over the accuracy, and those who narrate first set the tone for subsequent conversations. In fact, “credibility is the crucial resource, and asymmetrical credibility is a key source for power” (Keohane & Nye, 2012, p. 219), opening the road to *soft power* (Nye, 1990) that will be discussed in the next paragraph.

III. The softness of power. Towards non-material issues

In IR, power has long been identified with the ability to operate freely and with the capacity to push other states to do things they would not otherwise do. In the past, the ability to control other actors was often attributed to the possession of specific strategic resources by relating the strength of a state to quantifiable dimensions (population, territory, natural resources, wealth and economy, military forces) able to anticipate the performance of a country during a conflict (Nye, 1990). Due to complex interdependence,

power today lies in resources and the ability to influence state behaviour. International politics today is characterized by a magmatic context, making intangible forms of power all the more crucial. International institutions, national cohesion, and universalist culture foster the shift in power from capital to meaning (Nye, 1990; 2004). Soft power refers to the ability to make other nations want what one wants through cooptation, in addition to the use of violence and coercion (corresponding to hard power). Nye introduced the concept of soft power in 1990. In the aftermath of the Cold War, the concept effectively explained the persistence of U.S. global power. Attraction through intangible devices is the focus of this idea. Persuasion also constitutes power. The arrangements that facilitate such power are its resources.

The hard power dimension is directly related to the behavioural dimension (coercion) and the resource dimension (military arsenals, financial availability). Additionally, it is a measurable and predictable kind of power, at least theoretically (Nye, 2004). Unlike hard power, soft power does not require coercion or tangible incentives. Countries can achieve their goals in international politics because other countries admire them, emulate them, strive to achieve their level of prosperity and openness, and wish to follow them. Therefore, soft power is the ability to influence others' preferences. It is acquired by shaping the aspirations of actors in international politics in an emulative process (*agenda setting*), including by designing the priorities of discussions in institutional settings. Topics on the agenda are defined by lexical choices and salient issues (*framing*). When soft power is used effectively, it leads to the acquiescence of other actors. Soft power is political capital to be spent for future achievements. As a result, states and political actors strive to cultivate credibility. The creation of soft power involves a skilful blending of representations of foreign policy positions, public displays of values, and popular culture, according to Nye (1990, 2004).

Gallarotti (2011) argues that soft power depends on constructing a positive image through a variety of domestic and international sources. Some countries are attractive due to foreign policies they consistently implement

over time or domestic policy choices. When the states exercising soft power rely on this kind of influence and appeal, confrontation is no longer necessary. Consequently, nations tend to follow the policies and practices of the countries they choose to emulate.

The use of soft power is neutral: it can be cultivated or used for any purpose. For example, Russia has undertaken several information management initiatives recently. In accordance with the dictates of the non-linear war (Galeotti, 2016; Schnauffer, 2017), these operations focus on specific niches aimed at aggravating aspects of Russian politics that coincide directly with the ideology of the particular group, including anti-LGBT policies (for conservative parties), the legacy of communism (for socialist parties), and anti-Eurasianism (for nationalist parties) (Pomerantsev, 2015). Russian Internet Research Agency's disinformation campaign to destabilize U.S. domestic politics is a direct consequence of this familiarity with information management (Bastos & Farkas, 2019).

A second aspect relates to the credibility of information conveyed by states. Credibility is built by combining foreign policy and international standing with journalistic recognition and trust (Anzera & Massa, 2021b): for example, the BBC's international service is a canonical tool of soft power, while the Russian broadcaster RT is not.

In conclusion, soft power is a useful tool for understanding disinformation since it legitimizes the tools of communication management within the non-material sources of power available to governments. It is impossible to separate the positive image of a state from the effectiveness of its political practices and its actual performance, according to Nye (1990; 2004). Nevertheless, autonomous fact-checking is more challenging nowadays due to the complexity of verifying all news sources, especially when content is accessed online. The socio-technical affordances and user behaviours lead to the formation of echo chambers and filter bubbles within which partial information, not necessarily false, can confirm or amplify audiences' beliefs (Bruns, 2019). In addition, online platforms' lack of shared governance

makes it difficult to imagine solutions to counter informational interference that respects ethics and the common good, partly due to their constant claims of political autonomy.

And then comes culture. Interpretative turns and representations of international politics

Numerous turning points have impacted theoretical interpretations of IR in recent years. Breakpoints allow themes and issues typically marginalized within the discipline to become visible and more explicative (Baele & Bettiza, 2021).

Information competition is not necessarily related to anchoring issues in truth (as, for example, soft power suggests) but rather to influencing public perceptions of political matters. These interpretations also explain how international policy issues are brought to the attention of online users. By adopting popular geopolitics' interpretative framework, languages, themes, and practices seem to popularize issues in international politics.

In *classical geopolitics*, geographical factors are examined as determinants of a state's relations with its neighbors. *Political geography* (Agnew, 2002) analyzes how politics manages the relevance of geographical factors; conversely, *critical geopolitics* (Ó Tuathail, 1999) reveals how geographic configurations are discursive abstractions with social roots. As stated by Ó Tuathail and Agnew (1992, p. 192), "geopolitics [...] should be critically re-conceptualized as a discursive practice by which intellectual of statecraft 'spatialize' international politics in such a way as to represent it as a world characterized by particular types of places, people and dramas". In critical geopolitics, discursive elements are analyzed as individual abilities based on sociocultural resources to interpret world events and meaning (Tuathail & Agnew, 1992). The interplay of language and culture parallels economic and political structures in understanding the outcomes traditionally attributed to *territorial power* (Collins, 1978). Due to its direct connection to spatial representation, popular geopolitics can also be found in less traditional

forms of controlling territory. *Popular geopolitics* can be conveyed through the public discourse of ordinary people. Entertainment is one way in which a geopolitical orientation is spread: popular geopolitics analyzes the everyday experiences of ordinary people in relation to geopolitics (Dittmer & Bos, 2019, p. XIX). Popular culture provides a window into common perceptions of geopolitics. There are countless venues in which politics is played out, and these venues involve actors and themes which are quite different from what is traditionally studied in the analysis of phenomena that do not occur domestically (Hamilton, 2016):

Geopolitical knowledge is produced at a multiplicity of different sites throughout not only the nation-state, but the world political community. From the classroom to the living-room, the newspaper office to the film studio, the pulpit to the presidential office, geopolitical knowledge about a world is being produced, reproduced, and modified (Ó Tuathail & Agnew 1992, pp.194-195).

As the media's coverage of international affairs increases, IR is becoming more popular. This logic is being reflected in contemporary digital diplomacy. Through his Twitter account, Donald Trump has frequently spewed inaccurate information or caused information chaos by insulting, slurring, and characterizing opponents without supporting evidence (Özdan, 2019).

Moreover, the aesthetic turn emphasizes the importance of representing international phenomena (Bleiker, 2001). The postmodern turn (Hammond, 2006) posits that viewers are now accustomed to major international events being presented as media events (Dayan & Katz, 1992). As a result, the line between information and entertainment becomes thinner. As Bleiker (2001) pointed out, aesthetic representation contrasts with mimetic representation. From the perspective of aesthetics and authenticity, mimetic representation attempts to capture politics as it manifests. Aesthetic approaches acknowledge that an intentional gap separates representations and their objects. This gap is not to be minimized by aesthetic recognition, but it is to be acknowledged as the site where politics is actually made

– in the difference between a phenomenon and its representation. By reconfiguring the boundaries of visibility and publicity (Bleiker, 2017), aesthetic engagement transforms power into a negotiation between meanings and presence in areas of debate that affect decision-making (Barnett & Duvall, 2005; Lukes, 2004 [1979]). A narrative approach to the social sciences represents a new way of understanding them. Understanding political dynamics requires considering agency, storytelling (including normative storytelling), framing social movements, and creating metanarratives, which differ from only considering material phenomena (Roberts, 2006). As Roberts claims (2006, p. 710), “narratives are not just stories told but experiential events lived by the agents themselves. Narratives help to construct personal and social identity, provide sense and order to experience, and frame and structure action”.

Narratives, therefore, become a tool of political management once they are adapted to the trappings of strategy. *Strategic narratives* play a major role in shaping how political affairs are narrated and understood: by reframing past events and changing how they are told, actors can direct present action towards reaching their desired outcomes. In their study, Miskimmon et al. (2013) found that narratives can affect the following:

- *the system* defining IR’s structure;
- *identity building* by identifying and negotiating in a dialogical way the attributes that define each actor within the IR field;
- *issues*, shaping the framing around which international discussions are structured.

A narrative must adhere to the truth of the facts being described. In this process, elements of an event are selected and emphasized, or characteristics of international actors are defined. According to constructivism (Wendt, 1992), the international reality is not predetermined but results from negotiations between all parties involved. Social and discursive phenomena influence the perceptions of their counterparts and their relationship.

Incorporating information into the public disclosure of geopolitics highlights its role in security dynamics. The dissemination of information and its consequences exaggerates and accentuates some issues that develop into security concerns. In the securitization approach (Buzan et al., 1998), security issues are constructed discursively, utilizing rhetorical activations to link them together. Disinformation, on the other hand, has increasingly become a security concern because it has the potential to destabilize domestic systems or manipulate international conflict knowledge.

As the Russian-Ukrainian conflict unfolds, disinformation is strategically used to influence international public opinion. Several information and disinformation campaigns have been conducted on European territory since the disputed annexation of Crimea in 2014 (Khaldarova & Pantti, 2016). In an increasingly fragmented society, Russian narratives reflect skepticism about the international order. The dissemination of fake news has highlighted Putin's domestic politics, historical dynamics, and contestation with other political powers. As a result, media battles over frame control and legitimacy erupted. It is a situation in which globally situated perspectives become amplified through a multiplicity of situated perspectives (Szostek, 2018), reproducing geopolitical tensions and enhancing global media flows (Boyd-Barrett, 2017) that are directed toward personal news feeds (Makhortyk & Bastian, 2022).

Accordingly, the construction of information as a security issue should be investigated from the point of view of identifying disinformation by governments as a means of exerting greater control over news distribution, thus implementing censorship mechanisms (Neo, 2020; Wang & Huang, 2021) or as a geopolitical fault line. A case in point is the European Union's initiatives aimed at fighting disinformation which has primarily been used to sharpen soft power in response to interference in information spaces that threatens to destabilize Europe's soft power (Petroni et al., 2017). Several European initiatives focus on contrasting the disinformation operations coordinated by Russia to reassure the former Soviet republics. As demonstrated in the

newsletter UEvsDisinfo, this action does not take place exclusively through institutional denunciations but also through popular disseminating actions, which also serve as a repository able to objectify and archive practices of information pollution.

If it works. A realist interpretation of disinformation: Mearsheimer's typology

Many international events have illustrated how even political leaders use structured deception to pursue their foreign policy goals. Among the most relevant examples is the U.S. invasion of Iraq in 2003, whose official motivation, Saddam Hussein's possession of hazardous weapons of mass destruction, was later revealed by public opinion to be completely unfounded. According to Bennett and colleagues (2007), the Bush administration used the force of its ideas in order to define reality. The government's spin prevailed over dissonant, influential, fragmented voices. As a result, journalistic apparatuses were unable to distance themselves from government sources: seeking alternative sources would have implied distancing from official positions, thereby complicating verification routines.

In his 2013 book, Mearsheimer aims to understand political leaders' behavior when they deceive to safeguard or defend the national interest. It is crucial to keep in mind that this interpretation, which focuses on lies and wilful distortions within formal political practices, should not be seen as a celebration of leaderism. In these cases, deception is not used for personal gain but to pursue the realist instinct to advance a nation's interest. These lies have a domestic purpose, despite being directed at IR management. States' security is a concern of policymakers. Thus, international institutions are perceived to be ineffective in regulating the activities of high politics in this scenario. Thus, information curation does not appear to pose a significant ethical tension towards the international community. Deception can be classified into two basic forms, according to Mearsheimer. Usually, *concealment* involves hiding information that would otherwise be

useful for the public to judge decision-making processes. On the other hand, the *spinning* process emphasizes and amplifies items to influence domestic and international audiences.

It is common practice to use these forms of deception in international communication. Diplomatic immunity applies to high-ranking government officials, even when their information campaigns are improper or deliberately misleading. For example, Sanovich (2018) explains how Russian disinformation campaigns succeed by repurposing Soviet networking capabilities and practices for today's complex information ecosystem.

According to Mearsheimer, there are some key characteristics of information campaigns focusing on international issues. As a first step, it is important to distinguish (in terms of characteristics, scope, and effect) between lies told by states to each other and those targeted at international audiences. The reason political leaders lie is often related to high political ends, like the state's survival, rather than for their gains, like winning elections. Although lies between states are a long-standing practice in international relations, Mearsheimer argues that they are less frequent than imagined and not a standard practice. The lies states tell their citizens are worse than those directed at international actors since they undermine trust and have widespread consequences. Deliberate deception campaigns are more likely to occur against states like the U.S., engaged in global strategic issues. Media play a marginal or absent role in this elite-driven process. There is a strong influence of powerful elites over the media. In spite of the author's claim that the final recipient of some strategies of deception is public opinion, the model of information assumed is hierarchical, and information flows are determined by a bottom-up cascade (Entman, 2003). The media and its operators serve as lapdogs, passive replicators, and amplifiers of the power of the government (Louw, 2005). There is no acknowledgement of how journalism exhibiting watchdog, adversary, and investigator traits can hinder decision-makers' plans. The high stakes on which deception is often focused justify this consideration. Studies referring to the U.S. context indicate a

reluctance to criticize government during conflicts. The media emphasizes patriotism when military intervention is involved (Hallin, 1986).

Anarchy, however, forces decision-makers to use deception for the sake of strategic advantage for their country, thus preventing analysis from assessing moral or ethical values (Waltz, 1979). Furthermore, deception differs from lying. Lie means saying something deliberately false to make others believe it is true. In contrast, deception involves partial restitution of facts in an attempt to prevent others from learning the whole truth. Mearsheimer distinguishes several forms of deception:

1. Lies between states: these serve primarily to reduce tension. It may result in exaggerating armament capacity or concealing its possession. The real intentions of states and objectives of agreements are hidden during negotiations, and bluffs are used.
2. Alarmism: the perception of a threat is emphasized to gain public acceptance and support for proposed policies. Threats can be created or exaggerated from scratch, or hype can be stoked to move them up the public agenda.
3. Strategic cover-ups: politicians lie about a policy that has gone badly to protect a country's interests or to hide the terms of a controversial policy.
4. Nationalist myths: every nation cultivates myths that represent its past positively while negatively portraying rival groups.
5. Liberal lies: leaders create an arbitrary set of norms that describe acceptable and unacceptable conduct in times of peace and war in order to gain legitimacy abroad.

In these elite-driven processes, the rules and interests of politics and national interests balance concerted policy action. For the successful implementation of policies and operations, representation is essential. However, the media and widespread dissemination seem to be marginal or non-existent. Accordingly, the international reality is a set of partial truths subordinated

to the leadership's perceptions of external or existential threats: disinformation and misinformation are strategic and regulatory tools to be used sparingly and wisely.

What is next? For an integrated reading of disinformation

Three approaches have been discussed throughout the chapter. The first approach, which we call *multicentric institutionalism*, is based on information as a valuable asset whose role is likely to grow as actors and capabilities of individual actors multiply. Moreover, information policies can activate or direct public opinion mobilization, creating favourable climates. Information is part of the conflictual framework because its weaponization depends on credibility. Nye (2004) argues that credibility is expendable capital, as it determines the line between legitimate soft power operations and propaganda.

According to the second interpretation, which emphasizes the *constructivism of representations*, the intersubjective process plays a crucial role in global politics' social construction. Representations operate on security processes, and security processes are determined by the social positions of those who describe them and the extraction of those who observe them. The question might arise whether it is possible to know the world as it appears or whether it depends on who determines the lenses through which global politics are filtered.

The third interpretation is *pragmatic realism*. Lies are just another tool for governing anarchy. As selfish competitors strive to achieve their goals in a world where ethical and institutional boundaries do not always exist, it is natural to resort to such tactics. However, lying is not always necessary. In some cases, all that is needed is to spin the messages. The lies are difficult to detect because they are embedded in an elitist interpretation of high world politics, which by default does not involve public scrutiny.

These approaches return a piece of the complexity of information management in the modern age. What, then, remains to be studied? La Cour

(2020) suggests that IR could improve our understanding of 1) disinformationn facilitators, 2) how disinformation is used, and 3) how to respond to disinformation.

Furthermore, disinformation must be understood in its social context. Most often, especially when it comes to non-traditional IR issues, ordinary citizens and audiences remain outsiders or marginal figures. It is, therefore, necessary to design a diffuse discipline since disinformation acts on micro, meso, and macro levels.

References

- Agnew, J. (2002). *Making Political Geography*. London-New York: Routledge.
- Anzera, G., & Massa, A. (2021a). Using international relations theories to understand disinformation. Soft power, narrative turns, and new wars. In López-García, G., Palau-Sampio, D., Palomo, B., Campos-Domínguez, E., & Masip, P. (eds.), *Politics of Disinformation: The Influence of Fake News on the Public Sphere*. Hoboken: John Wiley and Sons, pp. 35-47.
- Anzera, G., & Massa, A. (2021b). *Media digitali e relazioni internazionali. Tecnologie, potere e conflitti nell'era delle piattaforme online [Digital Media and International Relations. Technologies, Power, and Conflicts in the Online Platforms Age]*. Milano: Guerini & Associati.
- Anzera, G., Gianturco, G., & Massa, A. (2019). Security branding and digital narrations. Security issues representation on Twitter, *International Review of Sociology/Revue Internationale de Sociologie*, 29(1), 80-101. doi: 10.1080/03906701.2018.1532656.
- Aron, R. (1966). *Peace & War. A Theory of International Relations*. London-New York: Routledge.
- Baele, S. J., & Bettiza, G. (2021). “Turning” everywhere in IR: on the sociological underpinnings of the field’s proliferating turns. *International Theory*, 13(2), 314-340. doi:10.1017/S1752971920000172.

- Barkin, J.S. (2003). Realist Constructivism. *International Studies Review*, 5, 325-342. doi:10.1046/j.1079-1760.2003.00503002.x.
- Barnett, M., & Duvall, R. (2005). Power in international politics. *International Organization*, 59(1), 39-75. doi:10.1017/S0020818305050010.
- Bastos, M., & Farkas, J. (2019). "Donald Trump is my president!": the Internet Research Agency propaganda machine. *Social Media+ Society*, 5(3). doi: 2056305119865466.
- Bennett, W. L., & Livingston, S. (Eds.). (2020). *The Disinformation Age: Politics, Technology, and Disruptive Communication in the United States*. Cambridge: Cambridge University Press.
- Bennett, W.L., Lawrence, R.G., Livingston, S. (2007). *When the Press Fails*. Chicago-London: University of Chicago.
- Bleiker, R. (2001). The aesthetic turn in international political theory. *Millennium*, 30(3), 509-533.
- Bleiker, R. (2017). In search of thinking space: Reflections on the aesthetic turn in international political theory. *Millennium*, 45(2), 258-264. doi.org/10.1177/0305829816684262.
- Boyd-Barrett, O. (2017). *Western Mainstream Media and the Ukraine Crisis*. London-New York: Routledge.
- Bruns, A. (2019). *Are Filter Bubbles Real?*. London: Polity Press.
- Buzan, B., Wæver, O., & De Wilde, J. (1998). *Security: A New Framework for Analysis*. Lynne Rienner Publishers.
- Castells, M. (1996). *The Rise of the Network Society*. Hoboken: John Wiley & Sons.
- Castells, M. (2009), *Communication Power*. Oxford: Oxford University Press.
- Chong, A. (2007). *Foreign Policy in Global Information Space. Actualizing Soft Power*. Basingstoke: Palgrave Macmillan.
- Collins, R. (1978). Long-term social change and the territorial power of states. *Research in Social Movements, Conflicts, and Change*, 1, 1-34.
- Comunello, F., & Anzera, G. (2012). Will the revolution be tweeted? A conceptual framework for understanding the social media and the Arab Spring. *Islam and Christian-Muslim Relations*, 23(4), 453-470. doi: 10.1080/09596410.2012.712435.

- Dayan, D., & Katz, E. (1992). *Media Events*. Harvard: Harvard University Press.
- Der Derian, J. (2009). *Virtuous war: Mapping the Military-Industrial-Media-Entertainment-Network*. London: Routledge.
- Dittmer, J., & Bos, D. (2019). *Popular Culture, Geopolitics, and Identity*. London: Rowman & Littlefield.
- Entman, R. M. (2003). Cascading activation: Contesting the White House's frame after 9/11. *Political Communication*, 20(4), 415-432. doi.org/10.1080/10584600390244176.
- Frenkel, S., & Kang, C. (2021). *An Ugly Truth: Inside Facebook's Battle for Domination*. Bridge Street Press.
- Galeotti, M. (2016). Hybrid, ambiguous, and non-linear? How new is Russia's "new way of war"? *Small Wars & Insurgencies*, 27(2), 282-301. doi.org/10.1080/09592318.2015.1129170.
- Gallarotti, G. M. (2011). Soft power: what it is, why it's important, and the conditions for its effective use. *Journal of Political Power*, 4(1), 25-47. doi.org/10.1080/2158379X.2011.557886.
- Guzzini, S. (2000). A reconstruction of Constructivism in International Relations. *European Journal of International Relations*, 6(2), 147-182. doi:10.1177/1354066100006002001.
- Hallin, D. (1986). *The "Uncensored War": The Media and Vietnam*. University of California Press.
- Hamilton, C. (2016). World politics 2.0: An introduction. In Shepherd, L. J., & Hamilton, C. (Eds.), *Understanding Popular Culture and World Politics in The Digital Age* (pp.3-13). London-New York: Routledge.
- Hammond, P. (2007). *Media, War, and Postmodernity*. London-New York: Routledge.
- Hepp, A. (2019). *Deep Mediatization*. London-New York: Routledge.
- Hoskins, A., & O'Loughlin, B. (2015). Arrested war: the third phase of mediatization. *Information, Communication & Society*, 18(11), 1320-1338. doi.org/10.1080/1369118X.2015.1068350.

- Karpf, D. (2018). Analytic activism and its limitations. *Social Media + Society*, 4(1). doi: 2056305117750718.
- Keller, F. B., Schoch, D., Stier, S., & Yang, J. (2020). Political astroturfing on Twitter: How to coordinate a disinformation campaign. *Political Communication*, 37(2), 256-280. doi.org/10.1080/10584609.2019.1661888.
- Keohane, R.O., Nye J. S. (2012, first edition 1977), *Power and Interdependence*. Boston: Pearson.
- Khaldarova, I., & Pantti, M. (2016). Fake news: The narrative battle over the Ukrainian conflict. *Journalism Practice*, 10(7), 891-901.
- Kovic, M., Rauchfleisch, A., Sele, M., & Caspar, C. (2018). Digital astroturfing in politics: Definition, typology, and countermeasures. *Studies in Communication Sciences*, 18(1), 69-85. doi.org/10.24434/j.scoms.2018.01.005.
- La Cour, C. (2020). Theorising digital disinformation in international relations. *International Politics*, 57(4), 704-723. doi: 10.1057/s41311-020-00215-x.
- Louw, E. (2005). *The Media and Political Process*. London: Sage.
- Lukes, S. (2004 [1979]). *Power: A Radical View*. Basingstoke: Palgrave Macmillan.
- Makhortykh, M., & Bastian, M. (2022). Personalizing the war: Perspectives for the adoption of news recommendation algorithms in the media coverage of the conflict in Eastern Ukraine. *Media, War & Conflict*, 15(1), 25-45. doi:10.1177/1750635220906254.
- Mearsheimer, J. J. (2013). *Why Leaders Lie: The Truth About Lying in International Politics*. Oxford University Press.
- Merrin, W. (2018). *Digital War: A Critical Introduction*. London-New York: Routledge.
- Miskimmon, A., O'Loughlin, B., & Roselle, L. (2013). *Strategic Narratives: Communication Power and the New World Order*. London-New York: Routledge.

- Morgenthau, H.J. (1948). *Politics Among Nation. The Struggle for Power and Peace*. New York: Knopf.
- Neo, R. (2020). A cudgel of repression: Analysing state instrumentalisation of the 'fake news' label in Southeast Asia. *Journalism*. doi:1464884920984060.
- Nye, J.S.(1990). Softpower. *Foreign policy*, (80), 153-171. doi: 10.2307/1148580.
- Nye, J. S. (2004). *Soft Power: The Means to Success in World Politics*. New York: Public Affairs.
- Ó Tuathail, G. (1999). Understanding critical geopolitics: Geopolitics and risk society. *The Journal of Strategic Studies*, 22(2-3), 107-124. doi: 10.1080/01402399908437756.
- Ó Tuathail, G., & Agnew, J. (1992). Geopolitics and discourse: Practical geopolitical reasoning in American foreign policy. *Political Geography*, 11(2), 190-204.
- Özdan, S. (2019). Subverting the art of diplomacy: Bullshit, lies and Trump. *Postdigital Science and Education*, 2, 95–112. doi: 10.1007/s42438-019-00075-6.
- Papacharissi, Z. (2015). *Affective Publics: Sentiment, Technology, and Politics*. Oxford: Oxford University Press.
- Petroni, S., Massa, A., & Anzera, G. (2017). *Lo specchio di Aletheia. Fake news e politica internazionale [Mirror of Aletheia. Fake News and International Politics]*. Roma: Nuova Cultura.
- Plantin, J. C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293-310. doi: 10.1177/14614448166661553.
- Pomerantsev, P. (2015). The Kremlin's Information War. *Journal of Democracy*, 26(4), 40-50. doi: 10.1353/jod.2015.0074
- Reus-Smit, C. (2020). *International Relations: A Very Short Introduction*. Oxford: Oxford University Press.
- Roberts, G. (2006). History, theory and the narrative turn in IR. *Review of International Studies*, 32(4), 703-714. <http://www.jstor.org/stable/40072179>.

- Ronfeldt, D., & Arquilla, J. (2008). Noopolitik: A new paradigm for public diplomacy. In Snow, N., & Taylor, P. N. (eds) *Routledge Handbook of Public Diplomacy* (pp. 352-365). London-New York: Routledge.
- Rosenau, J. N. (1990). *Turbulence in World Politics: A Theory of Change and Continuity*. Princeton: Princeton University Press.
- Sanovich S. (2018). Russia: The origin of digital disinformation. In Woolley S. C., Howard P. N. (eds.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (pp. 21-40). Oxford: Oxford University Press.
- Schnauffer, T. A. (2017). Redefining hybrid warfare: Russia's non-linear war against the West. *Journal of Strategic Security*, 10(1), 17-31. doi: 10.5038/1944-0472.10.1.1538.
- Szostek, J. (2018). Nothing is true? The credibility of news and conflicting narratives during "information war" in Ukraine. *The International Journal of Press/Politics*, 23(1), 116-135. doi:10.1177/1940161217743258.
- Tandoc Jr, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news" A typology of scholarly definitions. *Digital Journalism*, 6(2), 137-153. doi: 10.1080/21670811.2017.1360143.
- Vaidhyanathan, S. (2018). *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy*. Oxford: Oxford University Press.
- Van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media*. Oxford: Oxford University Press.
- Van Dijck, J., Poell, T., & De Waal, M. (2018). *The Platform Society: Public Values in a Connective World*. Oxford: Oxford University Press.
- Virilio, P. (2002). *Desert Screen: War at the Speed of Light*. London: Athlone Press.
- Waltz, K.N. (1979). *Theory of International Politics*. Reading, Mass: Addison-Wesley.
- Wang, C., & Huang, H. (2021). When "fake news" becomes real: The consequences of false government denials in an authoritarian country. *Comparative Political Studies*, 54(5), 753-778. doi: 10.1177/0010414020957672.

- Wardle, C., & Derakhshan, H. (2018). Thinking about “information disorder”: Formats of misinformation, disinformation, and mal-information. In Ireton, C., Posetti, J. (Eds.) *Journalism, “fake news” & disinformation* (pp. 43-54). Unesco.
- Wendt, A. (1992). Anarchy is what states make of it: The social construction of power politics. *International Organization*, 46(2), 391-425. <http://www.jstor.org/stable/2706858>.
- Wolfsfeld, G., Segev, E., & Sheaffer, T. (2013). Social media and the Arab Spring: Politics comes first. *The International Journal of Press/Politics*, 18(2), 115-137. doi:10.1177/1940161212471716.

FAKE NEWS, POST-TRUTH, AND JOURNALISM: WEAKNESSES AND STRATEGIES IN 2018 BRAZILIAN ELECTIONS

Luísa Guimarães Torre

Introduction

The announcement of Brexit and the election of Donald Trump, two events that took place between June and November 2016, showed the world that there was something going on in communication mediated by digital networks. The first episode, the UK's decision to leave the European Union through a referendum, and the second, an election marked by the rise of a billionaire media tycoon to power in the United States, a victory that none polling survey could predict, brought to the media two terms so far little known by the general public: fake news and post-truth. As the press and the experts tried to find answers to these developments, it became clear that there was a new factor at this table: the circulation of deliberately false or severely distorted information through social media.

In a post-industrial society in which information plays a fundamental role in social structures (Castells, 2012), where what is perceived as real becomes more and more what is disseminated in digital life, rumors have come to occupy an important space in interpersonal communication mediated by information technologies. Rumors are a form of communication that has existed in human societies for centuries. Kapferer (1993) says that rumor is the oldest media in the world, the “emergence and circulation in the social body of information

that has not yet been publicly confirmed by official sources, or that has not been contradicted by them” (p. 16). However, in the face of a new communication paradigm, with the introduction of the internet and social networks as important channels of communication between individuals, news circulation finds unprecedented scale and amplification. Those rumors that once spread by word of mouth now take on a new dimension. Rumors, now in the media, came to be called fake news and are a key part of the broader phenomenon of information disorder, which comprises misinformation, disinformation and mal information that started to circulate in the public sphere (Wardle & Derakhshan, 2018).

To this phenomenon of disinformation, another is added: for some individuals, the truth has become something closer to what they believe than to the facts, what Keyes (2004) called post-truth – elected, in 2016, the word of the year by Oxford Dictionaries. According to the dictionary, the entry would mean “relative to what denotes circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion or personal belief”. This is a fertile ground for deceiving discourses to be disseminated in social media networks, which constitute the essence of fake news. The concept of fake news has been extensively studied in recent years, but it is still a concept about which there is great debate in social sciences and in communication, with its inadequacy related to the word news (“news”) in its name (Ribeiro and Ortellado, 2018). However, in this study, we will use the fake news nomenclature, since the phenomenon we will study is treated in this way by the press, the focus of analysis of this work.

In a world of abundance of information, journalism is also changing. It becomes yet another voice in the midst of different enunciation poles and loses centrality in the production, selection and distribution of information (Bentes, 2015). If the logic of mass communication is hierarchical, limited to the transmission of content by a sender with a very limited possibility of debate on the part of the receiver, this has changed. Today everyone is on an equal footing, responding and negotiating (Wolton, 2010). The receiver is the new protagonist.

On social media networks, the way to consume news and information starts to follow new criteria. In place of newspaper publishers' assessment, recommendation algorithms are used to create personalization and define what is most relevant to your users. These mechanisms, not always transparent, will determine what appears first in the users' feed based on criteria such as popularity, interaction and similarity to already consumed content, among others – the composition of these formulas, however, is still, to a large extent, a business secret (Pasquale, 2015). While allowing broader access to information, social networks are presented in a format that makes it difficult to verify and prove the authenticity of content (Allcott and Gentzkow, 2017). This will also impact the dissemination of deceiving articles, as opposed to the more traditional model of disseminating journalistic information, which has been the dominating model for sharing news for decades.

In this way, traditional or professional journalism then faces transformations and changes that are also felt in the sphere of power. In this new form of society, organized in a network, power is also exercised in a network (Parente, 2004). It's a different kind of power, exercised in environments of high freedom, within an invisible structure, where capitalism is more alive than ever (Musso, 2004). In this context, several news organizations launched sections dedicated to verifying information that circulates in social media and on the internet, a methodology named *fact-checking*, to refute or confirm the content of the checked pieces.

Power relations are discursive, and discourse constitutes society and culture (van Dijk, 2017), a speech is a form of social action and is always relational. It can also be ideological, serving to build and legitimize relations of power and dominance (Wodak, 2001). As part of the social fabric, ideologies are also reproduced in news published by the press and – why not – in fake news that circulate on social networks. That is why, when analyzing the discourse, we will also notice the marks and ideologies expressed in these messages.

In this study, we want to understand: how are discourses built by fake news emitters, in articles classified by fact-checkers as so, and by journalists, in *fact-checking* articles, in terms of structure? Our objective is to unveil ideological polarization and power relations hidden in those structures. For that, we used critical discourse analysis as our methodology and investigated the discourse of Fato ou Fake, the fact-checking section of a traditional media outlet from Brazil, the Globo group, as well as the discourse of fake news emitters captured and debunked by the media group and published in this section.

The fake news and *debunkings* published in Grupo Globo's website G1, in the "Fato ou Fake" section, between August 31 and October 6, 2018 – during the legal electoral campaign of Brazilian elections – are the object of analysis of this study, that aims to identify discursive strategies used by both parties to test the hypothesis that they were engaged, in the context of the 2018 Brazilian general elections, in a discursive dispute, presenting polarized discourse structures – which was also a power dispute.

For a concept of fake news

The concept of fake news is being widely debated and studied in recent years, especially since 2016, when the term became popular after Donald Trump's election to the presidency of the United States. However, it is still a terminology about which there is great debate in the social sciences and in communication (Ribeiro; Ortellado, 2018). According to Ribeiro and Ortellado (2018), in literature, the debate is divided between those who defend the use of terminology, shaped by the political debate and in journalistic coverage, and those who believe that it is an inadequate term because it carries the word "news" in its name.

In fact, if a piece of information is not accurate or false, it is contradictory to call it news, since making journalism is the act of attributing faith that something really happened, based on reports from reliable and verified sources (Chaparro, 2007). However, the term ended up becoming the

most common to refer to a phenomenon that involves a series of implications for the informative scene and the public sphere in several countries around the world.

Allcott and Gentzkow (2017) conceptualized one of the most cited definitions of fake news, defining it as “news articles that are intentionally and verified false and that can deceive readers” (p. 213). It is a definition, therefore, restricted and focused on the intentionality of the content. Zuckerman (2017) argues that fake news can be understood in three ways: real facts that did not deserve to receive the amount of attention they received; political propaganda, which mixes true, false and misleading speeches with the aim of strengthening one side and weakening the other; or as *disinformatzya*, information that aims to pollute the news ecosystem and make it harder for the public to believe what media companies publish. It’s a broader concept based on the scope, intentionality of the content and its impact, but it doesn’t seem to cover all developments we find in the current disinformation phenomenon, especially when analyzing fact-checking sections, including the sample selected for this study.

The debate over the definition is therefore in questioning whether the term fake news refers only to demonstrably false news content or other types of disinformation and whether the concept should include only intentionally false content or also those factual misunderstandings, as calculation errors. Ribeiro and Ortellado (2018) note that what is conventionally called fake news sites produce not only false news, but mainly hyper partisan “combat information” in the form of news stories, that is – “it can only be a convenient clipping from the day’s news, a story with a sensational headline, a fact taken out of context, an exaggeration or speculation presented as fact – occasionally, it might even be a lie.” (p. 73)

However, it should be noted that, if at the beginning the term was adopted to designate only news sites that used the credibility of the journalistic format to spread false information, today on *fact-checking* pages in the press, content labeled as fake news are as diverse as possible: photos, videos, graphic

montages, prints of *Whatsapp* messages, prints of posts on social networks, in addition to montages on the layout of journalistic sites, among others. The term, therefore, cannot be seen in a limited way since it ends up covering a huge diversity of formats, contents, intentions and impacts.

It seems to be more appropriate to understand this complexified scenario as a conjuncture, or what Wardle and Derakhshan (2018) call “information disorder”, a systematic and comprehensive problem, which includes satire and parody, clickbait headlines, the misleading use of subtitles, visuals or statistics, as well as real and verified content shared out of context, the use of a journalist’s name or media vehicle logo to credible information without connection to them, digitally manipulated content and also deliberately fabricated content.

“From all of this, it emerges that this crisis is much more complex than the term fake news suggests” (Wardle and Derakhshan, 2018, p. 47). This broader concept, which combines the ideas of disinformation, misinformation and malinformation, classifies false content based on the nature of its falsehood, with “fake news” being part of this concept as an idea of information that emulates the journalistic discourse

From the observation of this phenomenon, we realize that fake news also share several characteristics of the rumor, the oldest media in the world, according to Kapferer (1993), especially due to its collective nature.

Kapferer (1993) explains that the rumor is essentially unverified information. The definition of a rumor as unverified information that circulates in this way results in an impasse in distinguishing a rumor from other information transmitted orally or through the media. The dividing line between what is true and what is rumor is subjective and results from persuasion itself, says Kapferer (1993): the rumor is only shared if people believe in it; if the information is classified as false, it dies; this is the paradox. Therefore, the parameter of true or false is useless, in the view of Kapferer (1993): “The rumor mechanism exists because people think that information is true

and they think it's important to pass it on. [...] The dynamic of the rumor, therefore, is independent of the problem of its authenticity". (p. 15)

Speed is one of its hallmarks, as rumors have a "virotic" capacity (Sodré; Paiva, 2011, p. 31): "...the rumor predominates by the seductive force of the imagination, it is more "viral" than the communication of the fact. On the electronic network, the legs of the lie may be short, but they are certainly very fast".

But the most important feature of the rumor, says Kapferer (1993), is its collective character:

There was a rumor – this collective act of speaking – because the group seized the information. [...] What is necessary to explain in the genesis of a rumor process is the adhesion, the mobilization of the group. Even if there is an initial speaker, the one that supports the rumor is the other people, the ones who heard it and will talk about it. (p. 24)

That is, despite the fact that there is an initial source, what creates the rumor is the behavior of the group that, upon learning about the fact, starts to talk about it. The rumor materializes not in its genesis, but in being passed from one person to another. It is also a group opinion, says Kapferer (1993): "it is not a matter of repeating, but of interpreting, of evaluating the implications of the initial fact, of defining public opinion, what the group thinks subjectively." (p. 46)

Rumor is, therefore, a vehicle of social cohesion, as it is through it that the group communicates among itself what should be thought to continue adhering to the ideas circulating there. The collective character of the rumor also gives clues as to why few individuals verify the information before passing it on. According to Kapferer (1993), the more the rumor is spread, the more easily it convinces. This can be seen in the phenomenon of the disinformation, as the deceiving information sees its power of convincing people increase as it circulates throughout social media groups and reach larger audiences. In the case of fake news, this group behavior

becomes even more fundamental for the presumption of veracity of these rumors that are now mediated because of the working dynamics of social networks, in which repetition, attention and emotion are the engines of the spread of information.

The so-called fake news, therefore, share the essence and characteristics of the rumor, with some differences. First, transmission is no longer oral and becomes mediated by technology, and information can be disseminated in different formats, such as text, image, video, audio and others. Second, information is now recorded in a format and stored on communication devices used by individuals, such as smartphones and computers. Therefore, we defend that fake news are a kind of mediatized rumor.

With that in mind, we can say that fake news, or the mediatized rumor, are deliberately false, manipulated content or real content shared out of context, unverified, shared with or without the intention to deceive, including political propaganda or hyper partisan information, featuring its strength in its nature of a group behavior. It is shared not based on its authenticity but based on the urgency of its utility to the group. This kind of information ends up polluting the news ecosystem, making it harder for the public to believe what media companies publish, being an important part of the disinformation disorder phenomenon.

Journalism in the age of “algorithmic gatekeeping”

In a scenario of networked communication and in a world of abundance of information, there is a change in the traditional role of journalists as gatekeepers of information. Journalism becomes one voice among many poles of enunciation (Pinto, 2000) in the media ecosystem and loses both control over the news (Chaparro, 2007) and centrality in production, selection and news distribution (Bentes, 2015). This does not mean that media organizations have become irrelevant – what happens is a complexification of the system (Primo, 2011).

An individual who accesses social media networks will no longer get information only through traditional journalistic vehicles, but through a range of options. “An interactor in cyberculture consumes any and all information they have contact with, according to their particular interaction strategies on the network”, explains Primo (2011, p. 141). Your worldview emerges from the intersection of all this information. Journalists have become just one of the groups and social actors involved in the news, as audiences are able “to even ignore them entirely to access first-hand information from a variety of other organizations and sources.” (Bruns, 2011, p. 137)

Many people now use social media as their primary news source. The processes that determine the way information and discourses circulate in social media and how it reaches citizens are fundamental to understanding the impact that the fake news phenomenon has, in particular, when it comes to social media networks. In them, the content undergoes a kind of curatorship that will determine the relevance of the information for what is published through the participation of the network itself, “where the network itself gives visibility to what it considers important, while despising what is not” (Recuero; Zago; Soares, 2017, p. 6). This process, Bruns (2011) called gatewatching, a practice that understands the action of actors in social media, where each user of the network observes the flow of information and defines which ones deserve more visibility, interacting, commenting or sharing that publication with their friends and followers, in a manner analogous to the gatekeeping process.

Journalists leave the scene and algorithms come into play: structures based on data, which generally work based on probabilistic techniques, which serve to classify information. Based on criteria that are not clear, heavily settled on digital surveillance, algorithms play a crucial role in deciding “how information circulates, how people meet and interact, and how behaviors are effectively carried out.” (Rieder, 2018, p. 136) Algorithmic classification is the foundation of the dynamics of personalization of information and recommendation systems, which, in turn, are the basis for the functioning of

social media networks. These recommendation mechanisms will also contribute to amplify or to silence discourses in the social media networks, in a process of personalization taken to the extreme that is not carried out by human beings, but by machines and mathematical processes.

Not only amateur content is classified by these criteria, but also news produced by media organizations. As a result, Heinderyckx and Vos (2016) say, various forms of “algorithmic gatekeeping” have become part of the media industry. Gatekeeping in the digital age shifts from a logic of relevance to the logic of popularity.

In Brazil, it is important to highlight that most people that consume news on the internet do it through social media. According to the Reuters Institute Digital News Report 2021¹, the consumption of information via social networks has been gaining strength year after year. In Brazil, 63% of respondents said they use social media as a news source, surpassing the 61% who said they use the TV as a news source. In social media, there is a predominance of Facebook (47% use it as a news source) and WhatsApp (43% use it the same). On the other hand, trust in news overall stood at 54%, up from 2020’s results of 51% (it was 59% in 2018, which dropped to 48% in 2019), well above news found on social media (34%).

This growth comes in an age of rising right-wing populism which uses fake news as a political communication strategy. Populist politics heavily disseminate the idea that the internet is a free space, where it is possible to exercise direct democracy and have direct access to information that large corporations, organizations and governments want to hide. It is this ideology that supports the classic populist maxim of direct and genuine representation of the will of the people without the traditional intermediation, which, by the same rhetoric, are characterized as agents that distort and corrupt any possibility of establishing an honest, transparent relationship between politicians and the people (Aggio, 2020).

1. Available on <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021>

As consequences of the widespread dissemination of fake news, there are particularly two important impacts to journalism: the discrediting of news and the artificiality of the public sphere. First, in social media networks, which limit quality control standards to determine what constitutes news, the spoofing and imitation of news sites blurs boundaries between what is true and what is false. Furthermore, they are not always there to persuade journalists or broad audiences, but to challenge the status of verifiable information produced by professional news makers, explains Ireton and Posetti (2019): “This confusion means that many news consumers are increasingly empowered to choose or create their own ‘facts’, sometimes aided by politicians seeking to protect themselves from legitimate criticism.” (p. 18). In this way, misleading content is “affecting citizens’ understanding of reality and undermining trust, informed dialogue, common sense of reality, mutual consent and participation”, says Ireton (2019, p. 39).

Second, it is possible to create “an artificial public sphere, which does not match the real expression of the representation of the media or even society” (Campos, 2018, p. 31). Making the public sphere artificial is a threat to democracy as individuals become guided by discourses that do not find an echo in reality and begin to attach great importance to matters that should not receive it. Since the discourses constitute the society and power relations (van Dijk, 2015), it is essential to understand what kind of discourses are being disseminated by disinformation creators, to perceive which strategies are used to deceive citizens with inaccurate or fabricated information.

The discourse of truth and its power

Journalistic discourse is not an autonomous discourse, according to Chaparro (1996): “Journalism, in its nature, has the ability to capture, understand, reorganize and disseminate the discourses that society produces, adding to them the credibility of critical mediation. By this understanding, journalism is an environment of macro-interlocutions” (p.1). In either way, conferring veracity to an information is part of the essence of the journalistic discourse. It is an assumption “that organizes social expectations in

relation to journalism – that journalistic discourse contains the essential predicate of veracity”, states Chaparro (2007, p. 11). Furthermore, the quality of attributing faith is within the nature of journalism.

The discourse of truth in the press is built by using quotation marks, presenting conflicting and contradictory possibilities, presenting auxiliary evidence, structuring information in an appropriate sequence, and exposing issues from different angles (Tuchman, 1993), in addition to the use of photography or video, the error correction policy, codes of ethics, among others – all to become credible to the public and provide proof of the veracity of their reports (Träsel; Lisboa; Reis, 2018).

From a philosophical point of view, according to Träsel, Lisboa and Reis (2018), “journalism only becomes knowledge if it manages to sustain its veracity through justification – truth and justification are, therefore, the pillars of the epistemological statute of practice” (p. 7). Thus, the discourse of truth is a discourse that heavily permeates the journalistic discourse.

Truth is central to journalism also because it is an instrument of power. Both Foucault (2005) and Castells (2015) talk about the power of establishing what is truth and what is a lie, and about how communication can infuse ideas and discourses in the mind of the individuals.

The “political economy” of truth has five historically important characteristics: “truth” is centered on the form of scientific discourse and the institutions that produce it; it is subject to constant economic and political incitement (a need for truth both for economic production and for political power); it is the object, in various ways, of immense diffusion and immense consumption (...); it is produced and transmitted under the control, not exclusive, but dominant, of some major political or economic apparatus (university, army, writing, media); finally, it is the object of political debate and social confrontation (the “ideological” struggles). (Foucault, 2005, p. 11)

As noted by Moraes (2010), the media discourse interferes in the cartography of the world, and its goal is to induce public opinion to think that what the social communication vehicles disclose, and their voices have more relevance and are truthful. Therefore, the establishment of what is the truth (and of what is lie) is so disputed by various actors in the public sphere, including politicians, and that's why the phenomenon of fake news is so strongly inserted in the political information scenario.

However, we entered a post-truth era as said by Keyes (2004), where the truth is closer to an individual's personal beliefs than to the facts. What is new is not the rumors, the lies, the manipulation and political falsehoods, "but the public's response to it. Indignation gives way to indifference and finally to collusion." (D'Ancona, 2018, p. 34)

This is the essence of post-truth culture, says D'Ancona (2018): its popular strength depends not on evidence but on the feeling it evokes. In this new world, there is a priority given to emotion over evidence. A relevant aspect is that disinformation campaigns serve to sow doubt and discredit traditional sources of information such as the press and political and scientific authorities. According to D'Ancona (2018), pressure groups that spread disinformation stimulate the public to question the existence of a reliable truth, turning everything into relativism. While the journalistic discourse is essentially connected to the truth, disinformation campaigns constantly aim to attack this fundamental.

It is in this scenario that algorithms connect individuals to the things they like or might come to like; something responsive to personal taste, but blind to veracity, according to D'Ancona (2018): "The web is the definitive vector of post-truth, precisely because it is indifferent to lies, honesty and the difference between the two." (p. 55)

As for the post-truth, Constante (2019) says it is "a political device, a power structure and an influencing force" (p. 2). The author assesses that the social networks "create, instruct, influence, determine, change, disturb, build

and raise new values and perceptions in this reality that is built day by day; in fact, their basis is the great architecture of persuasion” (Constante, 2019, p. 2)

This would be the fundamental impact of fake news. In the post-truth era, they are promoted and fueled by the use of the internet and social media, engineered by advertising and communication experts, who aim to manipulate public opinion (Constante, 2019). In this way, the individuals’ perception would move away from reality, which could cause a radical transformation in society.

Power relations are discursive

Since power relations are discursive and discourses constitute society and culture, who have control over the discourse and more characteristics of it are, by definition, also more powerful (van Dijk, 2017).

The information discourse, the one conveyed by the media, is a “fundamental” discourse, through which a social bond is established in societies without which there would be no recognition of identity. The information propagated by the media, and its transmission, is also a transmission of knowledge, knowledge about the world through language. And, in this way, whoever possesses this knowledge also holds a certain type of social power. (van Dijk, 2017; Charaudeau, 2013)

Communicating, informing, everything is a choice (Charaudeau, 2013). Each discursive choice in the construction of the event report will give the information discourse a different meaning: what is retained or neglected, what is emphasized or what is hidden, the game of said and unsaid, what is explicit and what is implied; the meaning is put into discourse through these choices that are not always perceptible by everyone.

In short, choosing meaning effects to influence the other, that is, in the end, choosing discursive strategies (Charaudeau, 2013).

Thus, this study will pursue an answer to the following question: which are the discursive strategies that permeate the construction and meanings of the information media discourse and the discourse of fake news? From a critical point of view, as we propose in this study, we will resort to Critical Discourse Analysis (CDA) to unveil these strategies. According to Meyer (2001), there is a consensus that ACD is not understood as a single method, but an approach that consists of different levels, and at each level a number of selections are made. However, whatever the case, addressing research problems will always be critical, as the CDA will always make explicit power relations that are often hidden. This is also the reason why we choose this method: clarifying the power relations hidden in both discourses.

In this work, we are interested in discovering how this fake news phenomenon affects the journalistic discourse to assess if there is a dispute between the two enunciation poles, journalists and fake news emitters. We searched for ideological polarization not only in the discourse content but also in the structures and strategies of the discourses emitted by both poles of enunciation.

Van Dijk (2017) says that a discourse can be ideological as it is an essentially evaluative condition and provides guides for the social perception and interaction of groups. Ideology is found at the level of meaning (sense) and reference (as actors and events are named). Ideologies are not innate to individuals, “but learned, shared and used to identify, form and maintain social groups and their power” (van Dijk, 2017, p. 15). If the media discourse is loaded with ideology, why would fake news, which emulate the information discourse, not also carry ideologies and power relations? And how is this articulated within these discourses? This is what we will seek to understand from the analysis below.

The approach suggested by van Dijk (2017) occurs through the triangle discourse-cognition-society, with discourse being perceived as a ‘communicative event’, ‘cognition’ being understood as both social, beliefs

and purposes, as well as evaluations and emotions and any other 'mental' or 'memory' structure, and 'society' is used to include both the local micro-structures of face-to-face interactions, as well as more societal and political structures diversely defined in terms of groups. In a similar approach, Fairclough (1995) proposes a critical analysis in three dimensions that involve: discourse, discursive practice and social practice. According to the author, each discursive event is a text in spoken or written language; it is an instance of discursive practice that involves the production and interpretation of text; and it is a piece of social practice.

In this work, we will analyze the following indicators: 1) macro-topics (what the article mainly talks about); 2) their contextual models (the relationship of discourse structures with the structures of local and global contexts); 3) interdiscursivity (the genres to which the discourse belongs and references to other discourses); 4) lexical units (the meanings of words, the structures of propositions and coherence and other relations between propositions); 5) generic structures (ideas that repeat several times and words that have a stronger ideological charge, and the observation if lexical repertoire is typical of some type of discourse); 6) local forms (sentence syntax and formal relations between sentences or sentences in ordering sequences, primacy, pronominal relations, active-passive voice, nominalizations, sequences, transitivity, impersonality and passivity); 7) if there is a displacement of the opinion to another enunciator, (and if this is being used in the sense of demonstrating the non-agreement with that speech or of stamping your own speech).

To understand discursive strategies, Wodak (2001) proposes to analyze: the naming of actors; the predication; the argumentation used to justify political inclusion or exclusion; the perspective, framing or representation of the discourse; intensification and mitigation, to control the illocutionary force of the statements. These strategies are also investigated in this analysis.

The global ideological strategies will also be observed in this study: the volume of propositions (detailing) that characterize the event; the importance given to a proposition (its organization at the level of macrostructures or microstructures); relevance to language users or participants, controlled by contextual models; the presence or absence of information from the models; the attribution, that is, the way in which the acts are attributed to the actors; and the representations of us (ingroups) and of them (outgroups), as proposed by van Dijk (2017).

These categories will serve as a basis for analysis so that we can answer the investigation question: what are the discursive strategies fake news broadcasters and professional journalists use to influence public opinion? Our hypotheses are that there is a discursive dispute between fake news issuers and traditional media, which is also a dispute of power; and that the discourse of traditional journalism in the midst of this dispute will be aimed at reaffirming credibility.

In this study, using critical discourse analysis categories pointed out, we will search for discursive strategies used both by the fake news issuers and by professional journalists in eight fact-checking articles (debunkings) published by Grupo O Globo on the G1 website, at the section Fato or Fake (<https://g1.globo.com/fato-ou-fake/>). The articles were published between August 31, 2018, and October 6, 2018, when the Brazilian presidential electoral campaign ran. Eight articles that mentioned news outlets, simulated news from traditional journalistic vehicles, or because it is a publication by a vehicle that defines itself as journalistic, were selected from a total of 40 debunkings that talked about politics in that period. This cut was made because of the objectives of this research piece. The articles that will be analyzed will be:

Table 1*Debunkings selected for the analysis*

Date	Id.	Article title	Link
04/09/2018	A1	É #FAKE que livro citado por Bolsonaro no JN é o que aparece com carimbo de escola de Maceió	https://g1.globo.com/fato-ou-fake/noticia/2018/09/04/e-fake-que-livro-citado-por-bolsonaro-no-jn-e-o-que-aparece-com-carimbo-de-escola-de-maceio.ghtml
21/09/2018	A2	É #FAKE print de texto que diz que Jean Wyllys foi convidado por Haddad para ser ministro da Educação em eventual governo	https://g1.globo.com/fato-ou-fake/noticia/2018/09/21/e-fake-print-de-texto-que-diz-que-jean-wyllys-foi-convidado-por-haddad-para-ser-ministro-da-educacao-em-eventual-governo.ghtml
21/09/2018	A3	É #FAKE capa da revista Veja sobre escândalo das pesquisas eleitorais compradas	https://g1.globo.com/fato-ou-fake/noticia/2018/09/21/e-fake-capa-da-revista-veja-sobre-escandalo-das-pesquisas-eleitorais-compradas.ghtml
27/09/2018	A4	É #FAKE que revistas publicaram capas com declarações de diretor da OEA sobre fraudes nas urnas	https://g1.globo.com/fato-ou-fake/noticia/2018/09/27/e-fake-que-revistas-publicaram-capas-com-declaracoes-de-diretor-da-oea-sobre-fraudes-nas-urnas.ghtml
29/09/2018	A5	É #FAKE documento atribuído ao Exército que exige perícia nas urnas eletrônicas antes e depois das eleições	https://g1.globo.com/fato-ou-fake/noticia/2018/09/29/e-fake-documento-atribuido-ao-exercito-que-exige-pericia-nas-urnas-eletronicas-antes-e-depois-das-eleicoes.ghtml
01/10/2018	A6	É #FAKE mensagem que cita a Globo e diz que manifestação pró-Bolsonaro é a maior da história	https://g1.globo.com/fato-ou-fake/noticia/2018/10/01/e-fake-mensagem-que-diz-que-manifestacao-pro-bolsonaro-e-a-maior-da-historia.ghtml
04/10/2018	A7	É #FAKE print de post que diz que Procuradoria Geral da República indeferiu candidatura do deputado Iran Lima	https://g1.globo.com/fato-ou-fake/noticia/2018/10/04/e-fake-print-de-post-que-diz-que-procuradoria-geral-da-republica-indeferiu-candidatura-do-deputado-iran-lima.ghtml
06/10/2018	A8	É #FAKE que candidato a vice de Bolsonaro propôs confisco da poupança	https://g1.globo.com/fato-ou-fake/noticia/2018/10/06/e-fake-que-candidato-a-vice-de-bolsonaro-propos-confisco-da-poupanca.ghtml

Before starting the analysis of the discourse of the selected articles, it is important to note that Grupo Globo's own journalistic vehicle classifies various types of information as fake news, including content that is unrelated to the journalistic text, such as posts on social networks such as Facebook and Twitter, Youtube videos, Whatsapp messages and images assembled with sentences, among others. This type of classification contributes to blurring the boundaries between what is credible information and what is a personal opinion or comment, without a journalistic nature, attributing a news value to, for example, a message sent by an individual to a group of friends containing a commentary on political facts.

Results discussion

After these initial observations, in our analysis, we saw that, in general, we can observe some recurrences in the analysis of both fake news and debunkings made by Grupo Globo vehicles.

The A1 article is a tweet posted by the son of the candidate Jair Bolsonaro, Eduardo Bolsonaro, which is also a politician, in which he comments the participation of Bolsonaro in the TV news broadcast *Jornal Nacional*, where the candidate showed a book which he said would teach kids "gender ideology". Bolsonaro said that this book was being adopted by public schools, which was a lie. While a part of society, which has a liberal view on moral issues sees no problem in these kind of debates, another conservative part believes that what they call "gender ideology" is something to be eradicated, which naturally generates much controversy. This clash and more specifically the conservative discourse is present as an interdiscourse within what Bolsonaro and Eduardo say. In this way, by building the image of Bolsonaro as someone credible and true and by polarizing his opponents, Eduardo implicitly proposes that the liberals are the ones who lie. At the level of local meanings, we can see that when Eduardo says "A book that Bolsonaro tried to show in *Jornal Nacional*", he uses the expression "tried to show", implicitly suggesting that the press did not give him space to express himself, as well as putting his father in the role of victim of a press

that wants to bring him down. In addition, in the second sentence, irony is used as an argumentation strategy. When Eduardo says “Ué, but wasn’t it Bolsonaro’s lie?”, he means the opposite, that Bolsonaro doesn’t lie on the subject. When he says “but it wasn’t Bolsonaro’s lie”, the subject is implicit and diffuse, in reference to all those who have already said he is a liar. It is also possible to observe that the construction of the sentence puts Bolsonaro passively in relation to the lie.

The second article, A2, is the debunking of an image that looks like a photo of a print that contains an alleged news from the G1 portal, which reads the title “Jean Wyllys confirms Haddad’s invitation to be Minister of Education in an eventual PT government” and the subtitle “The PSOL Deputy confirmed the invitation, and should be part of the portfolio. The names of Gleisi Hoffmann, Dilma Rousseff, Ciro Gomes and Lula to compose ministries are also speculated.” It is important to contextualize that the news, which reports a fact that did not happen, if it were true, would be quite relevant to the electoral scenario, giving the context in which the country found itself in 2018, as the Deputy Jean Wyllys was a of the politicians most attacked for representing the LGBTQIA+ community. Thanks to this militancy, he is also a politician quite rejected by more conservative sectors of the electorate. When analyzing the discourse generic structures, it is possible to perceive that the language used is journalistic. It is a text that fits the genre of journalistic reporting that even bears the signature of a reporter (“Bernardo Coram, from G1 Brasília”). All these elements are used to corroborate the veracity of the image as reliable journalistic information. We note that the title of the alleged news says that “Jean Wyllys confirms invitation”, which presupposes that the deputy not only was invited into the ministry, but he also confirmed the invitation. The following sentence brings the rhetorical figure of the repetition, in a clear strategy of intensification. The last sentence reveals that, while Wyllys would already be right in the ministry, other prominent names are being speculated, including two former presidents of the country, Dilma Rousseff and Lula, as well as one of the candidates in the electoral race, Ciro Gomes. Certainly, the indirect

meaning of strong names like these (and rejected by part of the electorate) is that an eventual government of Haddad would be radicalized, full of leftists and without space for dialogue with those who oppose him. However, by the positioning of the topics, we see that the most important name is also the most controversial, that of Jean Wyllys. We see that the acts are attributed directly to Congressman Wyllys, who confirmed the invitation and must accept it. The other names mentioned are being “speculated”, that is, they are referred to indirectly and passively. It is not possible to perceive a clear polarization in the text, because, when analyzing the text, it is not possible to determine who is the outgroup and who is the ingroup.

In the case of the A3 article, we see the cover of the *Veja* magazine that bears the title “The scandal that everyone suspected was disclosed. Shortly before the election, the biggest political scandal is disclosed: the polls were bought. The question is: who ordered it? It is still possible to read other messages in the cover headlines: “data show that Bolsonaro has always been ahead”, “PT is the biggest beneficiary of the electoral farce” and “Lula, even in prison, knew about the whole scheme”. In this supposed cover of the magazine, the main topic is the denunciation of a great farce in the elections with the objective of benefiting the Workers’ Party (PT) to the detriment of the candidacy of Jair Bolsonaro. We observed that the speech is typical of the dynamics of the political campaign, since it begins with a type of construction that is not common in journalistic vehicles, “Disclosure of the scandal that everyone suspected”, denoting a very clear opinion and not typical of the domain of news discourse of journalism, since the enunciator places himself within the discourse when it is said that “everyone suspected”. When analyzing the lexical units, we see that the use of the words “everybody” and “biggest scandal” clearly reveals a strategy of intensification to make the matter even more impactful. Then the speech says “The question is: who ordered it?”, which is a false question, since in the cover headlines below we already know who is responsible for the electoral farce: the PT, which is the biggest beneficiary, and whose leader, Lula, knew everything. We see, in these calls, a clear separation between Bolsonaro, who is a victim of the corrupt

system and who has always been preferred by voters, and Lula, who is in prison, but continues to command a criminal scheme from within the jail. We see from these naming and preaching strategies that there is a global ideological strategy of polarization between the in-group, of Bolsonaro, which is good, and the out-group, which is the PT and Lula, the bad ones. The bad guys are directly blamed in the phrases. In this discourse, we can see that it is implicit that the PT has always deceived the electorate, using shady methods to stay in power, and that Bolsonaro is a victim of this scheme. We see that the speech is all constructed in the sense of persuading the receiver to think that those who think that there is no electoral farce are alienated. So the implicit meaning of this magazine cover is that if the party is so corrupt, it doesn't deserve any voter's trust.

In the A4 article, we see again covers of three national magazines, as in the previous article. The theme is also recurrent in the sample: an alleged scheme to rig the 2018 elections. This is the main topic of this fake news, which contains three covers of the magazines "Época", "Veja" and "Exame", with statements attributed to the director of the Organization of American States (OAS) Department for Electoral Cooperation and Observation, Gerardo de Icaza. The image of this fake news is a montage with the three covers, in which Icaza assumes fraud in the electronic voting machines in the elections. The covers read: "What now PT? Gerardo de Icaza opens his mouth and assumes fraud at the polls in favor of PT", at Época; "Bomb! Gerardo de Icaza, director of the OAS, admitted negotiation to defraud electronic voting machines and collaborate with PT", in the headline of Veja, with a headline with the text "Fernando Haddad says he was deceived and calls urgent meeting with party leaders"; ""The PT is capable of anything." Gerardo de Icaza causes controversy after assuming fraud in the polls in the 2018 Elections. PT despairs", in Exame. As in the previous case, there is a simulation of the format and journalistic discourse on the covers. At the same time, we observed in some phrases on the covers a language closer to the political discourse typical of electoral campaigns, when a candidate uses strategies to disqualify the candidate or the opposing party. In the case

of *Época*'s cover, the positioning of the first sentence, highlighted graphically on the cover, reveals the importance given in this words to how much the party would be cornered. The question sets the stage for the next sentence, "Gerardo de Icaza opens his mouth and assumes fraud at the polls in favor of the PT." Two expression units draw attention: "opens his mouth" and "assumes fraud", a gradation strategy aimed at generating expectation in the reader: first he decided to speak, then he assumed fraud. The expression "opens his mouth" is already a more emphatic way of saying "tells" or "speaks". The word "assume" is also used with the aim of emphasizing the revelation. Here we observe the use of the intensification strategy. On the cover of *Veja*, the same strategies are used: the opening sentence is "Bomb!". The positioning of the word indicates that the most important thing about this headline is its destructive power, like a bomb. The next phrase shows again a verb that emphasizes the responsibility of the subject in his revelation: he was not told, he "admitted" that that fact was true, an admission of his guilt for collaborating with one of the electoral parties. As in the previous cover, the first sentence ("What now?" and "Bomb!") and the verb used ("assume" and "admit") carry the greatest ideological weight. On the cover of *Exame*, the structure is a little different. However, the highlighted phrase, enclosed in quotation marks, has a similar intensifying effect. "The PT is capable of anything," says the cover, suggesting that there are no limits, in this case, to the party's dishonesty. The use of the word "everything", a hyperbole, reveals a strategy of intensification, as in the other covers. In the following sentences, the verb "to assume" is also used to impose greater accountability on the subject of the speech and to increase the strength of the revelation. All three discourses are similarly constructed with the aim of convincing the reader through the use of gradation, hyperbole and intensification. We see in the three covers the implication of blame to Icaza and the PT, who were together in the fraud, and the representation of the exogroup as that are responsible for striking down Brazilian democracy and deceiving voters.

The A5 article brings the issue of electoral fraud and concerns a report made by a supposedly journalistic website called “A Bronca Popular”, about an Army document that would require expertise in electronic voting machines to guarantee its reliability. In the title, it is possible to read: “High command of the Brazilian Army and requires expertise in the polls before and after the elections!”. At first, we observed that the site reproduces the journalistic format: it contains the title, name of the reporter who signed the article, publication date, illustrative photo and the text. It is signed by Edesio Adorno. Once again, the journalistic format is used to give credibility to the information disclosed. As we continue reading the text, we see an attempt to simulate journalistic language, however some elements show that it is a piece of political propaganda in defense of the candidacy of Jair Bolsonaro. It is possible to say, therefore, that the genre of the text is that of political discourse. The text starts talking about the attack suffered by the politician, in the midst of the political campaign, when he was stabbed in the belly, characterizing the attack as “political”, which defines the course that the speech follows, and says that it “intrigued the high command of the Brazilian Army, which follows with redoubled attention the investigations carried out by the Federal Police”. Some specification elements used in the first paragraph (“political attack”, “Army high command”, “follows with redoubled attention”) will detail the relevance of the news and leave an implicit message that there is something special about this investigation into a political act that is attracting the full attention of the most important figures in the Army. The analysis of these lexical structures leads us to believe, in these sentences, that there is an important conspiracy against Bolsonaro that makes the colonels uneasy. The macro-topic of this speech, however, is only presented in the second paragraph, as an additional fact that also makes the Army uneasy: “Another fact that worries the barracks concerns the supposed vulnerability of electronic voting machines”. Note that the text itself raises doubts about the vulnerability of the ballot boxes, since it is “supposed”. The disqualification strategy used here, however, will be deconstructed throughout the text, in an attempt to prove that the vulnerability of the polls is not supposed, but real. The Army’s concern is

written in an official document signed by “last-rank graduates of the three weapons”, saying they demand that electronic voting machines should be “subjected to expertise by specialists from the Armed Forces, before and after the election”, says the text, without naming who these people are. It is not known what leadership positions they are in or how many. We also don’t know how many are from the Army, Navy or Air Force. While the recipient of the document is referred to explicitly, the STF (the Superior Court), the authors are cited in a totally imprecise way. The article then explains that the intelligence of the Armed Forces “suspects that the sudden and inexplicable growth of Fernando Haddad (PT), in the electoral polls, is the preparation of the ground for a mega electoral fraud”. In this sentence, we see the polarization explicitly. Haddad’s growth in the polls, not only “sudden”, would have occurred in an “inexplicable” way, which indicates an electoral “mega-fraud”. The indirect meaning that we infer from these phrases is that there is no explanation for Haddad’s success other than widespread and never-before-seen electoral fraud. The “mega-fraud”, a hyperbole, turns the suspicion about the possible vulnerability of the polls into a warning for a fraud never seen before. Just as the military is not named, those responsible for the coup are also referred to in an imprecise and diffuse way. In the last two paragraphs, we realize that the text is a manifest in favor of Bolsonaro, and clearly see the marks of a speech by the radical right, when there is a quote that talks of a homeland “free from the threats of communism” and of “Venezuelanization”, a neologism referring to the country led by leftist Nicolás Maduro.

In A6, the article analyzed is a print of a Facebook post, with deleted authorship, which contains the following comment: “Globo admitted live that the demonstration for bolsonaro is the biggest in history”. Along with this sentence, a video was shared, where you can see, in the upper left, the Globo News logo and the words “Live” and “São Paulo”. At the bottom of the screen, there is the following caption: “1 million protesters on Avenida Paulista and surroundings, according to PM”. In this case, the video made by Globo News becomes fake news because of the comment on the Facebook

post that takes it out of context. The video is real and shows a real scene, which took place in 2015, during the protests against then-President Dilma Rousseff. Focusing on the analysis of the text, we observe that the comment in the post has the typical language of everyday conversation, while the second sentence, as expected, fits the journalistic discourse. Again in this article, we see the use of the verb “admitted”, which carries an ideological charge since it assumes that the subject of the discourse recognized a fact that he had not previously accepted as a hypothesis. If Globo did not first want to recognize Bolsonaro’s strength, we can assume that Globo is against the politician, which we can characterize as a polarization between the in-group (Bolsonaro and his supporters) and the out-group (Globo). In addition to admitting, Globo would have admitted “live” the fact that there is a huge demonstration in favor of Bolsonaro, which increases the power of admission. We also see in the phrase that the demonstration was “the biggest in history”, once again, the use of the intensification strategy is present. These three are the expressions, in this text, with the strongest ideological charge.

The A7 article analyzes a print that imitates a page on the G1 portal, containing information about the rejection of the candidacy of a deputy who is trying to be reelected. We see the header of the G1 portal, with the website logo on the left side, and the indication “Elections 2018” on the right, above the phrase “Moju – Pa”, as if it were a regionalized website of the city of Pará. Below, it is possible to read the title “Procuradoria Geral da República Issues opinion rejecting the candidacy registration of Deputy Iran Lima (MDB), based on the LAW OF THE CLEAN FILE (in Portuguese, “Lei da Ficha Limpa”, a law approved to fight corruption in Brazil), contrary to the decision of the TRE/PA.” and several document images with some parts highlighted graphically. TRE is the regional electoral authority of Brazil and PA is the acronym for the state of Pará. Once again, we see that the image simulates a news item from the G1 site in its format and also in the lexicon used, typical of political journalism. In this text, we see three pieces of information: the first part of the sentence (“Attorney General’s Office Issues an

opinion rejecting the candidacy of Deputy Iran Lima (MDB)) and the second part (“based on LEI DA FICHA LIMPA”) present an explanation list, while the third part of the sentence carries a relationship of adversity in relation to the first and second stretches (“contrary to the decision of the TRE/PA.”), expressed in the word “contradicting”, which allowed the deputy’s candidacy to be registered. “Based on the Clean Record Law” also implies that the deputy has a dirty record, that is, he has a great chance of not being honest. Here the argumentative strategy of cause and consequence is introduced, with the rejection being the consequence of the application of the law. It is possible to observe that the repetition of verbs that carry a negative connotation for the subject of the speech, the deputy Iran Lima (“dismissing” and “contradicting”). The information is positioned as a title, which gives it importance. In this case, we observe that the deputy is referred passively, that is, he only suffers the effects of the reported decision. The sentence, however, is constructed directly, with the TRE being directly responsible for the decision taken that both condemns the deputy and contradicts a previous decision.

The last article analyzed, A8, is a debunking whose main topic is an alleged statement by the candidate for vice president on the ticket of Jair Bolsonaro, Hamilton Mourão, about his proposition around the confiscation of people’s bank savings. The fake news in question is a print of an article from “Blog da Cidadania”, a journalistic blog edited by Eduardo Guimarães, which has a bias in favor of leftist parties and ideas. We observe that the article uses the lexicon of journalism. In this case, we also see that there is use of the lexicon of economic discourse. In the article, the main topic is the false declaration of the vice candidate Hamilton Mourão defending the confiscation of bank savings, a relevant topic, as Brazilians experienced a similar episode in 1990, when then President Fernando Collor proposed a plan to stabilize the economy and inflation, and ended up blocking and confiscating the equivalent of about US\$ 100 billion from people’s bank savings. When we read the title of the fake news, we observe that the sentence is constructed directly and the voice is active, as the utterance shows that the

subject performs the action. In this way, the subject of the discourse is directly responsible for the proposal to confiscate the savings. The use of the word “also” is curious in this context, as it brings ambiguity to the sentence. The “also” can either be a presupposition that, like Bolsonaro, Mourão also proposes the confiscation of savings, or it can mean that, among other economic measures, the candidate for vice also defends the confiscation. This imprecision, which may have been intentional, can lead to an ambiguous interpretation, which will be encoded from the reader’s mental and contextual models. When analyzing the generic structures, the word “confiscation” brings an idea of punishment, increasing the ideological charge of the phrase. The subject of the discourse is called by his surname, and therefore it is assumed that he is an important person and that the reader knows him, given the public exposure he had been suffering in those months of the electoral campaign. If he defends a measure that was harmful to Brazilians, this leads us to analyze that the subject of the discourse is characterized as someone bad, who makes a mistake that has already been committed in the past. In this way, we can see a polarization between Mourão and the victims of the Collor Plan, who had their savings confiscated – a large part of the Brazilian electorate.

These observations lead to the conclusion that, among fake news, we frequently observe forms of implicit meanings (as in A1 and A3) or inaccuracies (as in A5), partly because they are messages that are part of the dynamics of political discourses and partly because the information is not objective and leaves room for the reader to interpret more freely. Another point worth noting is the use of figures of rhetorical style found in fake news, such as irony, in A1, and hyperbole, in A3 and A5. These are figures that are not used frequently within the journalistic genre, but they fit well into the political discourse, mainly because they evoke emotion in the receiver, which is part of the electoral dynamics, for example.

In the debunkings, we observe more the use of the rhetorical style figure of repetition, used in general with a strategy of intensification from the point

of view of the article, which is to prove that the message being analyzed is a lie, being observed in all articles analyzed. In fake news, we also see the strategy of intensification (in A2, A3, A4 and A6) being used for proofing the central ideas defended in these discourses.

In addition, we realize that debunkings in general present the themes directly, in an active voice, and explicitly, which is coherent with the intended objective, which is to explain and correct wrong information. We also see, in debunkings, the direct accountability of the subjects of the discourse. In some fake news, such as A1, we see the use of passive voice to soften accusations in relation to the subject of the speech. In fake news, sometimes the press is characterized as unfair, other times it is not mentioned directly; however, in the debunkings, we can see that fake news is always named as false or misleading, explicitly and repeatedly.

While in all the debunkings, the lexical repertoire is typical of journalistic discourse, in some fake news we see a lexicon more used in political discourse (in A1, A3, A5 and A7). Although the lexicon belongs to the domain of politics, we can still observe among fake news a simulation, in some cases (in A2, A3, A4, A5, A7 and A8), of the journalistic format or even the layout of the G1 site. In our opinion, this type of reproduction is made with the objective of borrowing the credibility that is inherent to journalistic vehicles.

The characterization of the subject of the discourse, in fake news, is different depending on the context of each piece of information. In A1 and A5, for example, the subject is placed as the victim, while in A8 it is qualified as the villain. In A1, the press, for example, is preached as censorship and unfair. However, in the debunkings, we can see that the information analyzed is always named as false (“fake”), explicitly and repeatedly, in several sentences. It is also important to note that this qualification is always present in the first paragraph, in addition to the title, which gives importance to this classification. We also see in the debunkings the repetition of negative expressions to reinforce the idea of the falsity of the information in question, in a clear strategy of intensification.

In debunkings, we also see the use, in all articles, of the argumentative strategy of proof and, in fake news such as A8, the strategy of cause and consequence. The intensification strategy was observed both in fake news and in debunkings, in order to reinforce the defended point of view. In the debunkings, in a clear, explicit and objective way to defend the falsehood of the analyzed piece. In fake news, to make the impact of information greater or to highlight some more central topics. The argumentative strategy of proof was used in all debunkings, but not in fake news, which, in our analysis, indicates that they had no goal to provide proofs of truth in their discourse.

Another important observation is that debunkings always bring, in the first paragraph, the indication that the information being analyzed is false, or “#FAKE”. In other words, right at the beginning of the articles, Grupo Globo already characterizes that information that will be reported as a lie. The same occurs in the titles of the debunkings: the expression “It’s #FAKE” starts the titles of all analyzed articles, which aims to bring more attention and deposit more strength to this information. Regarding framing, we observe in all debunkings a framing of information in a negative way, as a lie or falsehood. The defense of falsehood in the articles is made in an explicit and detailed way, where the use of a large volume of propositions that point to the same thesis of false information is observed, which makes the message even more clear and direct.

We noticed that, among the global ideological strategies, polarization was found in the majority (in 7 out of 8) of fake news and also in the debunkings analyzed in this sample. In the sample, ingroups are generally characterized as true, good, and defenders of good values, while outgroups are represented as liars.

In these cases, in fake news, the press was characterized as the outgroup, while in the articles published by G1, we see that fake news emitters were always characterized as the outgroup. The irony here is that although the press is often qualified as an outgroup, the use of the journalistic format

is still a way to lend credibility to an information. On the other hand, fact-checking is used, in this sample, within the context of Grupo Globo's journalistic vehicles as a way to reaffirm the credibility of journalism and its power.

We see, clearly in the debunkings, a large volume of propositions that detail the point of view defended in the articles, in relation to the falsity of the information analyzed. Another curious point is the use of the expression “actually” (“na verdade”), as in articles A4 and A6, which reserves for the newspaper the reputation of having the truth, as opposed to the information analyzed.

Final considerations

Social media networks have become an open space for the circulation of various discourses, including those that go against what is said in the press. Some, by ideology; others, out of economic interest, thanks to the business model of sites like Google; others, by political strategy. Communication is power, and discourses are both an instrument of power and control and part of social construction of reality. The way they confirm, legitimize, reproduce or challenge power relations, ideologies and dominance are expressed in their structures, and the in-depth analysis of discourses, both true and false, will help to clarify the dynamics of new power disputes in which communication processes are inserted.

These reflections led us to assess, from our starting question, whether the hypotheses raised in this work were confirmed. Regarding the discursive strategies used, we see that while in the fake news analysis we observe forms of implicit meanings, inaccuracies and active voice in relation to some discourse subjects, we see in debunkings the antithesis of this: the direct construction, in active voice, the presentation the arguments explicitly, and the direct accountability of the subjects of the discourse. We also see the use of figures of speech such as irony and hyperbole in disinformation articles, and repetition in debunkings.

In fake news, the press is characterized as unfair and deceitful; however, in the debunkings, we can see that fake news are named as false or misleading, explicitly and repeatedly. This qualification is always present in the first paragraph, in addition to the title, which gives importance to this classification. The intensification strategy was observed both in fake news and in the debunkings, in order to reinforce the defended point of view, while the argumentative strategy of proof was only used in debunkings. We see, clearly in the debunkings, a large volume of propositions that detail the defended point of view, that of the falsity of the analyzed information.

Among the global ideological strategies, polarization was found in fake news and also in the debunking analyzed, where in-groups are generally characterized as true, good, and defenders of good values, while out-groups are represented as liars. In the case of fake news, the press was characterized as the outgroup; while in the articles produced by G1, we see that the issuers of false information were characterized as the outgroup.

By analyzing both discourses in this sample, this made us confirm our hypothesis that there is a discursive dispute between fake news emitters and professional journalism, since there is a polarization between ingroups and outgroups in the discourses of both parties, both in terms of discourse structure and content.

On the other hand, we see that Grupo Globo's effort to disqualify false information involves reaffirming its power as holder of the truth. We see that there is a relationship of dispute of forces between the press and false news emitters: one tries to take advantage of the other's power, while the other tries to disqualify it. From this observation, we realize that fact-checking is used, in this sample, within the context of Grupo Globo's journalistic vehicles as a way to reaffirm the credibility of journalism and also as a way to defend the vehicles of the business group, placing them as holders of the truth, who fight an enemy who is all those who lie.

The consequences of the fake news phenomenon for journalism are important, but not yet fully known, since the phenomenon, in its current format, is

relatively new. However, in our view, this phenomenon of fake news reveals a change in power relations in the public sphere, and disturbs the information ecosystem. It has two consequences: making the ecosystem cloudy, given the confusion between true and false information, and more complex, since power relations in the network society (Castells, 2015), where power is also exercised in networks (Parente, 2004), will make the information more or less relevant to the user based on their own beliefs and opinions and based on the group's beliefs and opinions.

References

- Aggio, C. D. O. (2020). Comunicação eleitoral “desintermediada”, mas o quão realmente interativa? *E-Compós*, 23, 1–18.
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236.
- Bentes, I. (2015). *Mídia-Multidão: estéticas da comunicação e biopolíticas*. Mauad X.
- Bruns, A. (2011). Gatekeeping, gatewatching, realimentação em tempo real: novos desafios para o Jornalismo. *Brazilian Journalism Research*, 7(2), 119–140.
- Campos, R. (2018) O sonho libertário virou o pesadelo da liberdade. (Interview with Ricardo Machado). *IHU On-Line - Revista Do Instituto Humanitas Unisinos*, 18(520), 29-32.
- Castells, M. (2012). *A sociedade em rede - a era da informação: economia, sociedade e cultura*. (Vol.1). Gulbenkian.
- Castells, M. (2015) *O poder da comunicação*. Paz e Terra.
- Chaparro, C. (1996). Jornalismo na Fonte. In: Dines, A. & Malin, M. (Eds.). *Jornalismo Brasileiro: no caminho das transformações* (pp. 132–154). Banco do Brasil.
- Chaparro, M. C. (2007) *Pragmática do jornalismo*. (3rd ed). Summus.
- Charaudeau, P. (2013). *Discurso das Mídias*. (2nd ed.). Contexto.
- Constante, A. (2019). Fake News and Post-Truth. In: Peters, M. A. (Ed.). *Encyclopedia of Teacher Education* (pp. 1–4). Springer Singapore.

- D’Ancona, M. (2018) *Pós-verdade: a nova guerra contra os fatos em tempos de fake news*. Trad. Carlos Szlak. Faro Editorial.
- Fairclough, N. (1995). *Critical Discourse Analysis: the critical study of language* (1st ed.). Longman Inc.
- Foucault, M. (2005) *A Verdade e as Formas Jurídicas*. Editora Nau.
- Heinderyckx, F.; Vos, T. P. (2016) Reformed gatekeeping. *Communication and Media*. XI(36) 29–46, 2016.
- Ireton, C. (2019). Verdade, confiança e jornalismo: por que é importante. In: Ireton, C.; Posetti, J. (Eds.). *Jornalismo, “Fake News” e Desinformação: Manual para Educação e Treinamento em Jornalismo*. (pp. 34-45). Unesco.
- Ireton, C.; Posetti, J. (2019) Introdução. In: Ireton, C.; Posetti, J. (Orgs.). *Jornalismo, “Fake News” e Desinformação: Manual para Educação e Treinamento em Jornalismo*. (pp. 15-26) Unesco.
- Kapferer, J. N. (1993) *Boatos: o mais antigo mídia do mundo*. Forense Universitária.
- Keyes, R. (2004) *The post-truth era: Dishonesty and Deception in Contemporary Life*. St. Martin’s Press.
- Moraes, D. (2010) Comunicação, hegemonia e contra-hegemonia: A contribuição teórica de Gramsci. *Revista Debates*, Porto Alegre, 4(1), 54-77.
- Musso, P. (2004). A filosofia da rede. In: Parente, A. (Ed.). *Tramas da Rede*. (pp. 17-38) Sulinas.
- Parente, A. (2004). Enredando o pensamento: redes de transformação e subjetividade. In: Parente, A. (Ed.). *Tramas da Rede*. (pp. 91-110) Sulinas.
- Pinto, M. (2000). Fontes Jornalísticas: Contributos para o Mapeamento do Campo. *Comunicação e Sociedade* 2, Cadernos do Noroeste, Série Comunicação, 14 (1-2), 277-294. Ed. Uminho.
- Pasquale, F. (2015) *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

- Primo, A. F. T. (2011) Transformações no jornalismo em rede: sobre pessoas comuns, jornalistas e organizações; blogs, Twitter, Facebook e Flipboard. *Intexto*, Porto Alegre, UFRGS, 2 (25), 130-146.
- Recuero, R., Zago, G., Soares, F. B. (2017). Mídia social e filtro-bolha nas conversações políticas no twitter. *XXVI Encontro Anual da COMPÓS*, 1–27.
- Ribeiro, M. M., Ortellado, P. (2018). O que São e Como Lidar com Notícias Falsas: Dos sites de notícias falsas às mídias hiper-partidárias. *SUR* 27, 15(27), 71–83.
- Rieder, B. (2018). Examinando uma técnica algorítmica: o classificador de bayes como uma leitura interessada da realidade. *Revista Parágrafo*, 6(1), 123-142.
- Sodré, M.; Paiva, R. (2011) Informação e boato na rede. In: Silva, G., Künsch, D. A., Berger, C., & Albuquerque, A. (Eds.). *Jornalismo Contemporâneo: figurações, impasses e perspectivas*. (pp. 21-32) EDUFBA.
- Träsel, M., Lisboa, S., Reis, G. (2018). Indicadores de Credibilidade no Jornalismo: uma análise dos produtores de conteúdo político brasileiros. *Compós*, 1–19.
- Tuchman, G. (1993) A Objetividade como Ritual Estratégico: Uma Análise das Noções de Objetividade dos Jornalistas. In: Traquina, N. (org.) *Jornalismo: Questões, Teorias e Estórias*. (pp. 74-90). Vega.
- van Dijk, T. (2017). *Discurso, notícia e ideologia: estudos na análise crítica do discurso* (2nd ed.). Edições Húmus.
- Wardle, C., Derakhshan, H. (2018). Thinking about ‘Information Disorder’: formats of misinformation, disinformation and mal-information. In: Ireton, C.; Posetti, J. (Eds.). *Journalism, fake news & disinformation: handbook for journalism education and training*. (pp. 43-54) Unesco.
- Wodak, R., Meyer, M. (2001). *Methods of Critical Discourse Analysis*. Sage Publications.
- Wolton, D. (2010). *Informar não é comunicar*. Sulina.
- Zuckerman, E. (2017). *Fake news is a red herring*. Deutsche Welle.

OPINION-ORIENTED NEWS AS A SOURCE OF POLARIZED DISINFORMATION ON THE EU: A CASE STUDY ANALYSIS DURING THE 2019 EP ELECTIONS

Rubén Rivas-de-Roca & Mar García Gordillo

Introduction

The role of facts for public opinion is being put into question in the era of fake news. This is specially observed in election campaigns, which have become an objective of disinformation (Blassnig et al., 2019). Journalists show a preference for opinionated stories in a growing worldwide trend. The emergence of online news consumption meant a change in the relationship between politics and journalism, moving away the latter from its traditional function of providing keys for the understanding of the public sphere (Bennett & Livingston, 2018).

One of the consequences of the digital era is a news coverage focused on personalities (Van Aelst et al., 2012). In this context, the journalistic production has become cannibalistic as the media outlets compete more than ever with each other (Carlson, 2018). The coverage of the European Union (EU) is strongly affected by those problems. The European project is usually linked to matters of national policy, finding a utilitarian view of its policies. Besides that, when the EU is addressed as a singular issue, it is associated with diplomacy and corporatism, reinforcing the feeling of remoteness (Baisnée, 2014).

Another additional problem reporting EU is the difficulty in generating interest on this matter. Citizen disaffection has increased at the same time that some

institutional actions to promote Europeanism were introduced, showing the inability of public communication policy to achieve a Europeanization of the journalistic treatment (Walter, 2017). Elections to the European Parliament are even considered third-order elections by political actors, which is reflected in low turnout (Haßler et al., 2021).

The huge pro-European mobilization after EU enlargement in countries like Spain did not bring a better news coverage (Papaioannou & Gupta, 2018). In the framework of an increasingly interconnected political journalism faced with post-truth phenomena, this EU communication deficit has been widely studied in the literature in recent years (Goldberg et al., 2021). The use of national approaches is pointed out as the immediate cause of this problem of journalistic treatment.

Bearing those trends in mind, this research aims to conceptualize the role of opinionated news in disinformation, as this practice takes advantage of a polarized public opinion (Wagner & Boczkowski, 2019). Beyond a theoretical approach, we use the multiple-case study as research strategy to assess the degree of opinion-oriented stories about the EU. The report of this issue suffers from cultural clashes that threaten its journalistic quality (Norris & Inglehart, 2019).

Disinformation on EU issues

Covering the EU: the challenge of remoteness

Legacy media have formally attached some importance to the EU in their teams, as can be seen in the presence of correspondents before the European institutions in Brussels (Lloyd & Marconi, 2014). However, the coverage has continued to be carried out from a national perspective and without continuity, mostly depending on events (van Spanje & de Vreese, 2014). Initiatives to develop European journalistic narratives are scarce and usually financed by the EU institutions, but they overlap with a potential European Public Sphere (Rivas-de-Roca & García-Gordillo, 2022).

The number of EU news does not seem to influence citizen evaluations of its activity, but the use of crisis frames has contributed to heightening mistrust in the European institutions (Brosius et al., 2019). An informed citizenry is a necessary step for the well-functioning of a democracy regarding the European project. Transparency about elections improves the possible effects of information (Grill & Boomgaarden, 2017), although the EU has been traditionally accused of lack of clarity.

An example of the poor journalistic treatment of EU issues is that the press looks unable to build European identity in historical relevant moments, which has been accelerated by the rise of disinformation (Kermer & Nijmeijer, 2020; Otto et al., 2021). The failed Constitution for Europe of 2005, the refugee crisis in 2015 or Brexit were not accompanied by a proper EU news coverage. Following a chronological order of these events, the first of them was the failure to approve a Constitution for Europe in 2005. After public consultations, France and the Netherlands rejected that initiative. In the campaigns of these referendums, the media focused on national debates on Europe rather than European ones; thus, the text to be ratified was almost sidelined (Papaioannou & Gupta, 2018).

Strong national public spheres coexist with a weak and nascent European Public Sphere (EPS) on common interests, which have a recent example in Commission's proposal to put an end to seasonal clock changes. The prevalence of domestic affairs is not unexpectedly considering the robustness of state political systems and the richness of national identities. What is most relevant is that each of these two types of spheres imply different levels of expectations, impacting on participation models (Herkman & Harjuniemi, 2015). In this sense, the space for dialogue at the EU level is still limited.

On this backdrop, the EU institutions also play a role in fighting disinformation. Some initiatives adopted during the latest EP elections such as “Act, React, Impact” (2014) or “This time I’m voting” (2019) were useful to promote a well-informed society (Rivas-de-Roca & García-Gordillo, 2022). In addition to that, the EU officials employ day-to-day strategies to counteract fake

news on social networks, illustrating the institutional character of the EPS and the absence of fact-checking platforms that tackle with Europe.

The rise of disinformation in the digital age

The current news coverage of the EU occurs in a post-truth era, in which political journalists are forced to justify each of their statements to legitimize their work because of the proliferation of fake news (Carlson, 2018). 2016 is considered a turning point in leaving behind facts and the emergence of criticism of traditional media, both in the United States with the victory of Donald Trump (Pérez-Curiel et al., 2021) and in the EU with Brexit referendum (Lilleker et al., 2021). The withdrawal of the United Kingdom (UK) from the EU shows the challenge that disinformation involves for the European project, having different impact by country (Hameleers et al., 2021).

Southern European audiences have shown little resilience to disinformation given the polarized tradition of these countries (Roses & Humanes, 2019). While in the 19th century a mass commercial press emerged in the UK, countries such as Spain were still immersed in the opinion press model, focused on disseminating political ideas. Despite the active role of journalism in the period of political transition to democracy, the historical backwardness of the Spanish press is evidenced in lower reading rates compared to most European countries (Hallin & Mancini, 2004).

However, digital convergence is also a concern for journalistic quality. Critics of convergence see platforms as fuel for misinformation (Innes & Innes, 2021). There is a lack of specialization due to less knowledge of the sources, to which is added the limited time frames. Journalists have to work on stories disseminated through multiple digital channels (Humprecht et al., 2020). All these phenomena influence quality of the news items (Pavlik, 2013).

The way the news is made determines the social vision of politicians. Focusing on the statements of the leaders, now shared via social networks, supposes to delve into the personalization of politics (McAllister, 2007).

Likewise, it means the distortion of the value of facts for democracy, losing importance in the face of selected quotes without public relevance (Vázquez Bermúdez, 2006). These practices are a breeding ground for disinformation, identified as a frequent practice of current political communication that reduces the trust in democratic institutions. In fact, politicians are more important than the media in the dissemination of disinformation (Heiberger et al., 2021).

Previous studies on disinformation across Europe suggest that EU institutions prioritize reporting false content on social networks (Tuñón Navarro et al., 2019). Fact checkers or think tanks are defining structures to fight disinformation at the European level. Regarding the European Commission, disinformation represents a key challenge in its current action, but it has not been possible to stop this problem of spreading false information for negative purposes as seen in the Covid-19 pandemic (Salaverría et al., 2020).

Method

This research seeks to conceptualize the role of opinionated news in disinformation regarding the EU. We defined three research questions (RQs) on this matter:

RQ1. Which are the bias and the protagonists of the headlines?

RQ2. How is the negative bias towards EU institutions built?

RQ3. What is the reaction of readers to these news items on the EU?

According to our research design, we compared the coverage of EU affairs in local media from Germany, the UK and Spain during the framework of the 2019 EP elections. The analysis was performed on a sample of news items on European issues, collected over a six-month period (from January 1 to June 30, 2019). The elections took place on 23-26 May, meanwhile the EP published its first pre-electoral survey in February. This chapter only considers the journalistic pieces on topics related the EU, that is, their institutions and processes. In total, 612 units of analysis were captured.

Our study includes six local media outlets, two for each country, in some of the most populated cities of Germany, the UK and Spain. All those countries have a different historical relationship with the European project. We selected local newspapers because of its social function in nearby communities (Hess & Waller, 2017). Since the 1980s there are more personalization and negativity in the political coverage (Kuhn & Nielsen, 2014), but the journalism cultures by countries also play a role on this matter (Obijiofor & Hanusch, 2011). For this reason, our research carries out a comparative approach across Europe.

This study focused on two main items: personalization and negativity. These variables were analyzed through several categories: headlines, topics and reader's comments. We developed an analysis sheet with exclusive categories to gather the data, following the content analysis parameters of Krippendorff (2012). The collection of the sample was carried out manually and then analyzed through the software IBM SPSS Statistics, Version 28.

The research design was applied by a single researcher, but we also carried out two previous rounds of coding training to refine the categories and achieve scientific validity. New rules were added to the coding manual after these rounds. The use of one encoder tries to provide homogeneity, as this chapter is part of an extended research project.

The categories created aim to deepen in the fact that the EU does not have its own media system, which is assessed by the literature as basis for proper news coverage. In addition to that, citizens vote according to identity and values, so one might wonder about whether these values exist at the European level. The lack of common identity leads to voting in a national perspective, explaining the low participation rates in the elections to the European Parliament. In the 2014 elections, the turnout in the EU as a whole stood at 43%, although there were several Eastern countries in which it did not reach 30%. In 2019, the first general increase since 1994 was observed, reaching half of voters (50%). Therefore, the period of the EP elections is chosen as a timeframe of our study.

It should be noted that the EU model has some democratic legitimacy through holding elections to the European Parliament every five years. A democratic system requires the existence of public opinion as a way for its citizens to make informed decisions, based on an idea of deliberative democracy. Nevertheless, the different views in the Member States and the little concern for EU issues call into question the journalistic coverage that we explore here.

Results

Headline bias

The headline is the classic entrance to the information. There are different elements to assess, but we focus on knowing the protagonist of the headline, insofar as this data reveals to whom the media's attention is oriented. Brief messages stand here as a source of political journalism, since political parties share messages in a massive way and easily convertible into headlines.

The topics or the people addressed by headlines are one of the factors that determines the first impression of EU news coverage. For this reason, we seek to know the degree of relationship that each type of headline has with the bias delivered to the European project, that is, if the personalities or thematic issues are more likely to some approaches. This makes it possible to understand which actors the Eurosceptic news is targeting.

	Positive	Neutral	Negative	Positive / negative
EU personality	15.0	65.0	3.7	16.3
Non-EU personality	10.8	45.5	33.7	10.0
European affairs	30.6	44.9	10.1	14.4
Non-European affairs	22.2	60.5	9.0	8.3
Total (average rate)	19.6	54.1	14.1	12.2

Table 1 - Distribution of the bias of the message according to the orientation of the headline (%). *In bold outstanding findings.

The frequency of headline bias remains mainly neutral, although some differences emerge. The highest positive frame is present in European affairs (30.6%), while non-EU personalities, mostly national, receive a negative bias above the average (33.7%). This data is relevant, illustrating how negativity points to these profiles of personalities outside the European bubble, which lead us to reflect upon whether the polarization of the EU comes from national spheres.

In the mix of positive and negative biases, no large differences are detected, but the first practice is slightly more common in the items that refer to the EU (EU personality and European affairs). The relationship of this finding with the customary nature of the EU and the deficit of politicization with European politics could be studied in future research. Moreover, the positive bias attributed to purely European issues may be linked to the presence in the sample of two historically pro-European countries such as Germany and Spain.

It should be taken into account that the EU news items in these media are not located in local or regional sections, but rather in political descriptors. In this regard, the approach for these pieces almost never goes through a nearby perspective, but through national or European approaches. This may be connected with the progressive politicization of the European project (de Wilde et al., 2016), which rises its influence in many areas and, hence, compels to consider political aspects in its coverage. However, findings such as the national orientation of the headlines mean that these interpretations are not clear. The media prefer a national orientation in the headlines, understood as gateway to the information, perhaps because they believe that it may foster more clicks on their web pages.

Topic bias

The majority bias on the EU news is usually neutral along the time (Lloyd & Marconi, 2014). This is not an obstacle to the existence of topics that generate polarizations of interest (table 2), with greater differences than those seen in the headlines. The classification of topics provided refers to the main issues of debate in the 2019 EP elections, triggered from an observation of the Twitter profiles of the candidates to preside over the European Commission. They were named as ‘Spitzenkandidaten’ (leading candidates).

	Positive	Neutral	Negative	Positive / negative
Institutional issues	24.2	35.5	12.2	28.1
Electoral contest	19.2	69.3	6.4	5.1
Brexit	12.0	62.4	17.9	7.7
External relations	43.5	35.4	10.3	10.8
Economy	30.4	42.8	10.9	15.9
Migration	15.0	37.6	30.7	16.7
Environment	41.5	34.6	8.5	15.4
EU-funded projects	83.1	13.6	0.0	3.3
Digital market	23.5	28.9	19.3	28.3
Social policy	48.4	36.7	6.7	8.2
Other issues	6.8	72.8	20.4	0.0
Total (overage rate)	31.6	42.7	13.0	12.7

Table 2 - Distribution of message bias according to topic (%). *In bold outstanding findings.

One of the most noteworthy data is the huge positive bias of EU-funded projects (83.1%), which also happens on a smaller scale in social policy (48.4%), external relations (43.5%), and the environment (41.5%). By contrast, the topic of migrations (30.7%) doubles the average of information with a negative tone. The news item of figure 1 shows a positive tone about European investments. This overlaps with a utilitarian perspective of what the EU does for me at the local level.

Europa subvenciona a 21 municipios sevillanos para implantar WiFi en espacios públicos



Fig. 1. Journalistic piece on EU-funded projects with a positive bias that tells the benefits of EU funding (Sevilla Actualidad (Spanish digital media outlet), June 30, 2019). Source: <https://bit.ly/32aVjKn>

Besides that, typical EU issues such as the EP electoral contest (26.9%) and the digital market (28.6%) bring together the combination of positive and negative approaches, following the trend towards moderation proper of European politics. In any case, the figures reveal that there are some issues close to the EU that tend to concentrate positive or negative approaches. Thus, depending on the prevailing agenda, the European project will be subject to different kinds of biases. We also observed that the topic influences on the type of authorship. Most of the pieces are signed by the agencies, but journalists stand out as authors for the EP electoral contest, Brexit, and social policy.

Positivity towards the EU institutions is higher in pro-European countries and is determined by the issues tackled in the journalistic field. For instance, there is greater negativity towards migration. As a consequence, the media agenda on the EU can shape the tone of journalistic messages. This means a learning for the European institutions, which must set the topics of interest if they want the media conversation to take place from a Europeanist logic.

Reader comments

News relevance is key in the gatekeeping process. A proof of the importance in the current digital pieces is the number of comments, which is related to discursive participation. The reception of a greater number of comments has the potential to cause the news production to be more audience-oriented,

so the number of these interactions is measured for the cases of each country of the sample.

	0-5 comments	6-15 comments	Over 15 comments
German local media	100.0	0.0	0.0
British local media	53.8	29.5	16.7
Spanish local media	94.5	4.5	0.9
Total (average rate)	82.8	11.3	5.9

Table 3 - Frequency of the figure of comments by country (%). *In bold outstanding findings.

As can be seen from table 3, the two German media do not receive comments, which is also found in one of the Spanish media outlets. The British media are more likely to generate interactions (29.5% of the pieces between 5 and 15 comments, and 16.7% with more than 15). These frequencies seem significant compared to the other two countries. Therefore, in the framework of this chapter the EU news suffer from a lack of interest in relation to comments, being the UK the only exception.

The reader comments on the webs are sometimes plagued by incivility messages, which reverts to the interest in participating and the feeling of quality of the news item (Prochazka et al., 2018). Additionally, the reasons to intervene in these spaces change among countries depending on their journalistic cultures. This happens at a historical moment in which the use of digital tools undermines trust in institutions, instead of spurring a true public discussion.

Specifically, the tone of the messages on the Internet is a factor in the process of crisis of journalism and democracy (Costera Meijer & Groot Kormelink, 2021). In this sense, Table 4 shows how the bias of the readers' comments is exposed. The tone of these comments is analyzed in relation to the European project, as the Eurosceptic discourses may shape their contents.

	Favorable	Neutral	Unfavorable	No comments
German local media	0.0	0.0	0.0	100.0
British local media	3.3	33.2	34.3	29.2
Spanish local media	0.8	15.4	6.6	77.2
Total (average rate)	1.4	16.2	13.6	68.8

Table 4 - Bias of comments by country (%). *In bold outstanding findings.

The results reveal that positive comments about the EU are scarce (1.4% of the total), with neutral and negative messages prevailing. It may be noted that the media that receive the most unfavorable comments belong to the country that generated the highest volume of interactions: the United Kingdom (34.3% of their comments are negative). This significant data is illustrative of how the interaction occurs under negative parameters.

While in the British media the attention attached by the EU is linked to negativity and neutral approaches (33.2%), these neutral ones are preferred in the Spanish media (15.4% versus 6.6% of unfavorable messages). According to the data, the EU is hardly associated with positive issues on most of the audience, which shows the presence of an idea of crisis attributed to the European project.

In our study, negativity is much stronger in a country with a long Eurosceptic tradition like the United Kingdom. The frequencies of news items without comments are significant, which leads to the conclusion that the corpus is not extensive enough to assess this point in depth. However, this finding also evidences that the EU is not able to mobilize public reactions as it is usual of opinion-oriented news.

Conclusions and Discussion

This chapter aimed to examine the role of opinionated journalism in the breakthrough of disinformation about the EU. The timeframe of the 2019 EP elections is used as a relevant moment for the future of the European

institutions. Prior research on EU communication has primarily focused on the political dimension of legacy media, but it is also necessary to shed light on the role of local journalism as having a closer relationship with citizens. Because of the distant character of the European project, innovation in communication should be key for both institutions and media. Nevertheless, the primary contribution of this study is to evidence that local media lack from originality in their texts about the EU, not fostering a true public discussion.

The reasons for the poor quality could be found in the fact that the topic is not really an important issue for them, together with the structural weaknesses of a local press traditionally blamed for bad practices (López García & Maciá Mercadé, 2007). Our observance found little features of an in-depth coverage to the extent that most of the pieces are brief and have few sources.

Drawing upon a corpus of 612 news items, we provide three interrelated theoretical contributions that also answer to research questions. First, the headlines show a greater presence of thematic elements than of personalities, illustrating a preferential attention for fragmented news rather than for individuals. Non-EU personalities are those with the highest percentage of negative approaches. Hence, national politics work as a polarized cleavage in the EU field, meanwhile the European affairs trigger more positive views. Beyond that, it is noteworthy that the news items collected are mainly placed in political sections, leaving behind the local or regional dimension.

Our second contribution offers insightful findings on the different biases by topic. EU-funded projects, social policy, external relations and the environment are reported from a positive view. However, the negative tone is remarkable regarding migrations. This means that the agenda-setting is essential for the European debates, since the most mentioned issues triggered different biases among the population.

Regarding the third, we further our understanding on the working of readers' responses. Our study reveals that the EU is not relevant for the audience in terms of comments, although the EP elections took place during the

research. The selected British media are an exception, having some pieces with more than 5 comments. Moreover, we analyzed the tone of these messages to explore possible Eurosceptic narratives. On this matter, positive comments are very rare. People who decide to interact through the formula of comments do so to express neutral or negative views. The unfavorable outlook is especially frequent in the UK, whose local media were those that evidenced a more negative approach towards the EU.

Based on the literature review and the three contributions provided, we argue that the prominence of opinion-oriented news about the EU could boost polarized disinformation. Our theoretical approach through the multiple-case study underlines how the current state of play of journalism influences the success of polarization in the digital sphere. In a context of fragmentation of audiences, opinion-oriented news is key for disinformation phenomena addressing liberal institutions. This disruptive communication around individualization could mean a decline of democracies such as the EU (Bennett & Livingston, 2018), since facts are no longer relevant for the audience.

The findings overlap with theoretical concepts such as the comparing media systems (Hallin & Mancini, 2004) and the journalistic cultures (Obijiofor & Hanusch, 2011). Polarization is more frequent in the UK (polarized liberal model), while the German press shows approaches that seek a balance from different sources. The low negativity in Spain could be explained through a Europeanist political culture. It would be interesting to find out if a higher polarization affects EU issues in Germany and Spain in the coming years. Eurosceptic parties are already present in their national parliaments, changing the traditional political discussion.

As outlined before, our results suggest that the European project is not evaluated as a useful topic by the local media, notwithstanding the EU rules have a strong impact on the nearby communities. The local or regional approaches are substantially ignored when covering the EU on these media. Specifically, this is concerning because the phenomenon of disinformation

is growing in contexts of proximity, supported by the advantages of geographical close groups to know better the effectiveness of mechanisms of propaganda (Correia et al., 2019).

The reduced appearance of EU personalities in the headlines may be explained through the low personalization of the European project. Despite a global tendency to personalize politics (McAllister, 2007), the EU has not succeeded in creating identification with recognizable leaders. The biggest recent example is the procedure of ‘Spitzenkandidaten’. This attempt to develop leading candidates has not implied that the campaign moved around these Europeanized actors.

We should acknowledge some limitations of the study, as the sample is small and focuses on cases of interest by city and country. The trends of the European project (elite-driven integration or corporatism) could be reasons for the particular style of disinformation detected. EU news are mostly ruled by national parameters and one of the most striking findings is the negativity in both headlines and reader comments in the UK. This has to do with a great use of political statements as a source in that country.

Taking these results together, we argue that academia will have to empirically strengthen our understanding about the local spheres as spaces of transnational discussions. The EPS is a widely studied concept. Many authors talk about the building of a single EPS or the Europeanization of national spheres. Our proposal is to revisit from a proximity approach this concept, primarily defined as a common space of deliberation for citizens that may serve against disinformation.

Accordingly, future research should expand the scope to examine the impact of journalism about the EU in local spheres. Disinformation is also related to low-quality journalism and the rise of opinion-oriented news. The disconnection between the EU and its citizens is likely rooted in structural reasons that have to be with the unpopular bias of the European integration. Nevertheless, this deficit does not prevent that the EU policies have

a great impact on the life of people. Hence, we need cross-cutting models to attain a better understanding of the network agenda and the shaping of disinformation in contexts of proximity.

References

- Baisnée, O. (2014). Reporting the European Union: A Study in Journalistic Boredom. In R. Kuhn & R. K. Nielsen (Eds.), *Political journalism in transition: Western Europe in a comparative perspective* (pp. 131-149). I.B. Tauris.
- 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>
- Blassnig, S., Engesser, S., Ernst., N., & Esser, F. (2019). Populism in Online Election Coverage: Analyzing populist statements by politicians, journalists and readers in three countries. *Journalism Studies*, 20(8), 1110-1129. <https://doi.org/10.1080/1461670X.2018.1487802>
- Brosius, A., van Elsas, E. J., & de Vreese, C. H. (2019). Bad news, declining trust? effects of exposure to economic news on trust in the European Union. *International Journal of Public Opinion Research*, edz025. <https://doi.org/10.1093/ijpor/edz025>
- Carlson, M. (2018). The Information Politics of Journalism in a Post-Truth Age. *Journalism Studies*, 19(13), 1879-1888. <https://doi.org/10.1080/1461670X.2018.1494513>
- Correia, J. C., Jerónimo, P., & Gradim, A. (2019). Fake news: emotion, belief and reason in selective sharing in contexts of proximity. *Brazilian Journalism Research*, 15(3), 590-613. <https://doi.org/10.25200/BJR.v15n3.2019.1219>
- Costera Meijer, I., & Groot Kormelink, T. (2021). *Changing News Use. Unchanged News Experiences?* Routledge.

- de Wilde, P., Laupold, A., & Schmidtke, H. (2016). Introduction: the differentiated politicization of European governance. *West European Politics*, 39(1), 3-22. <https://doi.org/10.1080/01402382.2015.1081505>
- Goldberg, A. C., van Elsas E. J., & de Vreese, C. H. (2021). One Union, different futures? Public preferences for the EU's future and their explanations in ten EU countries. *European Union Politics*, 22(4), 721-740.
- Grill, C., & Boomgaarden, H. (2017). A network perspective on mediated Europeanized public spheres: Assessing the degree of Europeanized media coverage in light of the 2014 European Parliament election. *European Journal of Communication*, 32(6), 568-582. <https://doi.org/10.1177/0267323117725971>
- Haßler, J., Magin, M., Russmann, U., & Fenoll, V. (2021). *Campaigning on Facebook in the 2019 European Parliament Election*. Springer.
- Hallin, D. C., & Mancini, P. (2004). *Comparing Media Systems: Three Models of Media and Politics*. Cambridge University Press.
- Hameleers, M., Brosius, A., Marquart, F., Goldberg, A. C., van Elsas, E., & de Vreese, C. H. (2021). Mistake or Manipulation? Conceptualizing Perceived Mis- and Disinformation among News Consumers in 10 European Countries. *Communication Research*, 1-23. <https://doi.org/10.1177/0093650221997719>
- Heiberger, R., Majó-Vázquez, S., Castro, L., Nielsen, R. K., & Esser, F. (2021). Don't Blame the Media! The Role of Politicians and Parties in Fragmenting the Online Public Space. *International Journal of Press/Politics*, 1-32. <https://journals.sagepub.com/doi/full/10.1177/19401612211015122>
- Herkman, J., & Harjuniemi, T. (2015). Unity or Heterogeneity: The Promise of a European Public Sphere? In R. G. Picard (Ed.), *The Euro Crisis in the Media: Journalistic Coverage of Economic Crisis and European Institutions* (pp. 221-241). I. B. Tauris.
- Hess, K., & Waller, L. (2017). *Local journalism in a digital world*. Palgrave Macmillan.

- Humphrecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to Online Disinformation: A Framework for Cross-National Comparative Research. *The International Journal of Press/Politics*, 25(3), 493-516. <https://doi.org/10.1177/1940161219900126>
- Innes, H., & Innes, M. (2021). De-platforming disinformation: conspiracy theories and their control. *Information, Communication & Society*, 1-19. <https://doi.org/10.1080/1369118X.2021.1994631>
- Kermer, J. E., & Nijmeijer, R. A. (2020). Identity and European Public Spheres in the Context of Social Media and Information Disorder. *Media and Communication*, 8(4), 28-39. <http://dx.doi.org/10.17645/mac.v8i4.3167>
- Krippendorff, K. (2012). *Content Analysis, An Introduction to Its Methodology*. Sage Publications.
- Kuhn, R., & Nielsen, R. K. (Eds.) (2014). *Political journalism in transition: Western Europe in a comparative perspective*. I.B. Tauris.
- Lilleker, D., Bene, M., Balaban, D., Fenoll, V., & Kruschinski, S. (2021). Populism, Its Prevalence, and Its Negative Influence on Democratic Institutions. In D. Palau-Sampio, G. López García & L. Iannelli (Eds.), *Contemporary Politics, Communication, and the Impact on Democracy*. IGI Global.
- Lloyd, J., & Marconi, C. (2014). *Reporting the EU: news, media and the European institutions*. I.B. Tauris.
- López García, X., & Maciá Mercadé, J. (2007). *Periodismo de proximidad*. Editorial Síntesis.
- McAllister, I. (2007). The personalization of politics. In R. Dalton & H. D. Klingemann (Eds.), *The Oxford Handbook of Political Behaviour* (pp. 571-588). Oxford University Press.
- Norris, P., & Inglehart, R. (2019). *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. Cambridge University Press.
- Obijiofor, L., & Hanusch, F. (2011). *Journalism across cultures: An introduction*. Palgrave Macmillan.

- Otto, K., Degen, M., Olgemöller, M., & Köhler, A. (2021). News Coverage of the Greek Sovereign Debt Crisis in German Public Service Broadcasting: A Case Study Analysis in Tagesschau, Heute, Brennpunkt and ZDF Spezial. *Journalism Practice*, 15(7), 1012-1032. <https://doi.org/10.1080/17512786.2020.1760729>
- Papaioannou, T., & Gupta, G. (2018). *Media Representations of Anti-Austerity Protests in the EU*. Routledge.
- Pavlik, J. V. (2013). *Innovation and the future of journalism*. *Digital Journalism*, 1(2), 181-193. <https://doi.org/10.1080/21670811.2012.756666>
- Pérez-Curiel, C., Rivas-de-Roca, R., & García-Gordillo, M. (2021). Impact of Trump's Digital Rhetoric on the US Elections: A View from Worldwide Far-Right Populism. *Social Sciences*, 10(5), 152. <http://dx.doi.org/10.3390/socsci10050152>
- Prochazka, F., Weber, P., & Schweiger, W. (2018). Effects of Civility and Reasoning in User Comments on Perceived Journalistic Quality. *Journalism Studies*, 19(1), 62-78. <https://doi.org/10.1080/1461670X.2016.1161497>
- Rivas-de-Roca, R., & García-Gordillo, M. (2022). Understanding the European Public Sphere: a review of pending challenges in research. *European Politics and Society*, 23(3), 380-394. <https://doi.org/10.1080/23745118.2021.1884965>
- Roses, S., & Humanes, M.-L. (2019). Conflictos en los roles profesionales de los periodistas en España: Ideales y práctica. *Comunicar*, 58(27), 65-74. <https://doi.org/10.3916/C58-2019-06>
- Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I., & Erviti, M.-C. (2020). Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19. *Profesional de la Información*, 29(3), e290315. <https://doi.org/10.3145/epi.2020.may.15>
- Tuñón Navarro, J., Oleart, A., & Bouza García, L. (2019). Actores Europeos y Desinformación: la disputa entre el factchecking, las agendas alternativas y la geopolítica. *Revista de Comunicación*, 18(2), 245-260. <https://doi.org/10.26441/RC18.2-2019-A12>

- Van Aelst, P., Sheafer T., & Stanyer J. (2012). The personalization of mediated political communication: A review of concepts, operationalizations and key findings. *Journalism*, 13(2), 203-220. <https://doi.org/10.1177/1464884911427802>
- van Spanje, J., & de Vreese, C. (2014). Europhile Media and Eurosceptic Voting: Effects of News Media Coverage on Eurosceptic Voting in the 2009 European Parliament. *Political Communication*, 31(2), 325-354. <https://doi.org/10.1080/10584609.2013.828137>
- Vázquez Bermúdez, M. A. (2006). *Noticias a la carta. Periodismo de declaraciones o la imposición de la agenda*. Comunicación Social Ediciones y Publicaciones.
- Wagner, M. C., & Boczkowski, P. J. (2019). The Reception of Fake News: The Interpretations and Practices That Shape the Consumption of Perceived Misinformation. *Digital Journalism*, 7(7), 870-885. <https://doi.org/10.1080/21670811.2019.1653208>
- Walter, S. (2017). *EU citizens in the European Public Sphere: An analysis of EU news in 27 EU member states*. Springer VS.

**LIES ARE ALL AROUND BUT WHO ARE
THE LIARS? ADDRESSING ONLINE
DISINFORMATION PLATFORMS IN
THE CZECH REPUBLIC AND SLOVAKIA**

Ján Hacek & Lucia Virostková

In January 2022, Prague City Gallery installed a curious piece of modern art close to the city library and magistrate: a fake news-sensitive street lamp titled “Visibility.” The video object is connected to a media monitoring server; the more traffic flows to Czech websites with untruthful content, the more blurred is the street lighting. “Fake news can have the same effect on our mind and judgment,” commented the curator of the exhibition (Monitora.cz, 2022). The art display is a powerful illustration of how the phenomenon of disinformation penetrates all areas of our public sphere. It is also an apt symbol of the dynamic machinery behind the distribution and amplification of fabricated reports which must be approached with a certain amount of creativity in order to be exposed and outperformed.

In the aftermath of the global pandemic of Covid-19 and throughout the major military conflict in Europe, fighting disinformation appears as a matter of urgency on both state and personal levels. As acknowledged by the European Commission (2018), public harm triggered by deception involves “*threats to democratic processes as well as to public goods such as Union citizens’ health, environment or security.*” After the coronavirus outbreak started, the peril of mass spreading of deceptive content became apparent (Radu, 2020). The phenomenon

famously referred to as “infodemic” showed its potential to “*hamper an effective public health response and create confusion and distrust among people*” (UN.ORG, 2021).

It is in the interest of state authorities to prevent the circulation of inaccurate information among citizens without hindering freedom of expression. The main role in the dissemination of fake news is evidently played by social networking services and public expectations for action to this effect are communicated at the international level. But the production of controversial content also happens “*within a variegated “alternative” news ecosystem increasingly dominated by hyperpartisan, anti-system and conspiracy news websites*” (Štetka et al., 2021). In order to recognize the right strategy to prevent the negative effects of this information environment, it is important to recognize the owners, contributors, and business models of the online media outlets.

This chapter describes the structural features of the disinformation scene in the Czech Republic and Slovakia and its key narratives that have influenced the political debate. Both member states of the EU and NATO, the two countries share a significant part of history, as well as strong parallels in the development of their economic and media systems. Due to language similarities and a high share of the bilingual population, disinformation can spread freely across borders and online media outlets. Hence by studying the background of controversial websites operating in both languages we effectively cover the existing market and online space where the same actors can influence audiences from both states.

Just before the pandemic, personal experience with disinformation-related issues in the region proved relatively weak. In the 2019 Eurobarometer, Czechs and Slovaks were surveyed as having less personal experience with fake news than the EU average. But a comparison with previous periods showed a clear trend upwards. In fact, Slovakia registered the second-highest rise in the share of citizens who often come across fabricated reports in media (Eurobarometer, 2019). Subsequent public-health concerns following the

pandemic outbreak changed the perspective on the role of the state in preventing fake news distribution. Police and government officials sought to contest online content which encouraged virus denials and mass breaches of social distancing or other hygiene regulations. Intelligence services continued to monitor and report on major disinformation outlets which were presented externally as an alternative eco-system (BIS.cz, 2021). However, some national parliamentarians personally engaged with websites known for publishing manipulative articles (Špalková Krátka, 2021).

Our objective in this chapter is threefold: apart from an empirical analysis of the structure and development of the Czech and Slovak “alternative” news online eco-system, we compare it with global disinformation tendencies after the break-out of Covid-19 pandemic. Secondly, we present the results of detailed research on transparency issues of websites with controversial content. Finally, we sum up some of the most recent strategies to counter disinformation.

Challenge of detecting and categorizing disinformation

The studies of disinformation and its threat to democracy grow in number and scope. *Disinformation* is defined as a sub-category of *misinformation*, i.e., inaccurate information which is produced intentionally; its function is to mislead people (Fallis, 2015). While the basic categories of “true” and “false” for evaluating any written statement appear as clear and straightforward, a growing body of empirical data demonstrates the use of several forms and formats of disinformation (Staender, Humprecht, 2021).

Different categories of fake news have been identified in content analyses of online reporting. Brennen et al. (2020) demonstrate that reconfiguration of accurate and false claims and their context prevailed over completely fabricated statements on Covid-19 pandemic. Gregor and Mlejnková (2021) recognise similar trends in reports on polarizing events such as the EU migration crisis in 2016 and they list other frequent manipulation techniques apart from fabrication, such as blaming, labelling, and appealing to fear in their study of Czech disinformation websites.

The alternative versus mainstream media landscape is also the subject of intensive academic discourse. Data from 10 countries examined by Hameleers et al. (2022) describe the dynamics in the relation between mis-/disinformation and distrust in news media. Several authors investigate how alternative media position themselves against mainstream media and highlight criticism of content and objectivity in both directions (Mayerhöffer, 2021; Wu, 2021). Horne D. et al (2019) refer to a strong homogeneity of tightly formed communities across social media and to content sharing only with similar news sources while alternative and mainstream news media often report on the same events but with competing and counter-narratives.

Regarding propaganda and disinformation campaigns, Starbird et al. (2018) point to empirical evidence of the Russian-government media apparatus with its political and military goals as being integrated into the alternative media ecosystem. Academic debate is also dedicated to questions of ideological polarization and its relation to specific characteristics of online communication, such as the effect of filter bubbles present on social media as the key instruments of amplification of fabricated news (Spohr, 2017).

Online disinformation eco-system in the Czech Republic and Slovakia

Vulnerabilities of media systems in the region

The global financial crisis of 2008-2010 presented a crucial moment for media systems in Central Europe following their post-communist transformation. As a result of the major economic downturn affecting the circulation and advertising expenditures of most commercial media outlets, a majority of foreign investors sold their assets in key publishing houses to local businessmen. Štetka (2015) refers to the period as the *third media ownership transformation* marked by de-globalization and oligarchization of the media sector. In the Czech Republic, the series of shifts started with the purchase of Mafra media group in 2013 by Andrej Babiš and his Agrofert, the third-largest business in the country (Vojtechovská, 2017). By late 2020, Czech businessmen held control over all relevant local news media

(Štetka and Hajek, 2020). Similar patterns of “business parallelism” (Zielonka, 2015) characterised by local media moguls active in various businesses can be identified in Slovakia where the presence of foreign investors has been practically reduced to Ringier Axel Springer, a Swiss-German publisher (Dragomir 2020). Media concentration and a long-term lack of transparency in media ownership present a high risk for market plurality in the region (Sampor, 2021; Štetka, Hajek, 2021).

The strong involvement of local oligarchs has been one of the factors behind an upsurge in public distrust in legacy media brands in both states. The 2021 Digital News Report confirmed the fall in trust score for most of the surveyed brands in Slovakia – with overall trust in news by the domestic audience at 30 percent as compared to 36 percent in the Czech Republic where the trend mildly improved after several years of decline (Chlebcová Hečková and Smith, 2021; Štetka, 2021). The negative emotion associated with media ownership has been repeatedly nurtured by politicians. Up to 2020 in Slovakia, mainstream media had a particularly problematic relationship with the then Prime Minister Robert Fico who titled journalists “Soros’ slaves” among other names (Kernová, 2020). His successor Igor Matovič followed suit shortly after taking office, implying that reports reflecting criticism against his government had been instructed by oligarchs owning the media companies (Mikušovič, 2020). In the Czech Republic, Babiš moved from business to politics in 2017 but maintained control over newspapers he owned which had prompted numerous journalists to leave their jobs out of concerns for their professional integrity and encouraged a public discourse on the ethical and economic crisis in Czech journalism (Hájek et al., 2015). At the same time, several cases of direct attempts at influencing journalists and their reporting by political and business actors had been exposed in the region (Petková, 2017).

The concentration issues and oligarch-dominated ownership of established media brands have been also fully exploited by “alternative media” in the region. The network of print outlets and fringe websites had emerged primarily in reaction to the pro-European and pro-NATO narratives embraced

by mainstream media following the enlargement of both international organisations to Central Europe. The key part of alternative media provides a platform for extremist groups and opinions, promoters of conspiracy theories, as well as advocates of anti-system measures and anti-democratic regimes, mainly in Russia and China. Combined with outlets related to public health issues, such as anti-vaccination and the Covid-19 pandemic, these sites constituted the core of the disinformation eco-system in the region (Klingová, 2021). However, apart from merely ideological and propagandist objectives, economic interests also played an important role: according to a conservative estimate, 76 million US dollars in ad revenues flow each year to disinformation sites in Europe, with Google supporting the largest number of disinformation domains, mostly via Google ads (GDI, 2020).

Disinformation campaigns and their prominent subjects

A growing body of research is dedicated to the phenomenon of “hybrid war” as represented by communication strategies between Russia and pro-Kremlin actors (Galeotti 2019; Charap 2015). Central Europe is considered a “laboratory” for testing and modifying Moscow’s disinformation campaigns (Sultanescu, 2019, p. 77). Clear evidence of “master narratives” concerning political and economic weaknesses of the European Union and Western liberal ideology can be identified in various forms and adapted to concrete events and situations in individual countries (Levinger, 2018). In Slovakia, several alternative websites which are currently considered as part of disinformation mediascape originally enjoyed broader support by conservative groups and pro-life organisations; they were esteemed for advocating traditional cultural values. These online platforms emphasised the same supportive attitude towards Russian President Vladimir Putin and his criticism of modern liberal ideology. The narrative resonated mainly during the political debates before and immediately after the referendum related to the rights of the LGBT community in 2015, and in a series of unsuccessful attempts at the ratification of the Istanbul Convention on preventing and combating violence against women in Slovakia (Gabrižová, 2019).

The same “alternative news” eco-system in both Slovakia and the Czech Republic produced misleading online content on the migration crisis and influx of Muslims in 2016 which got amplified via polarizing discussions on social networking sites (Šuplata and Nič, 2016). Finally, during the corona crisis of 2021 new websites focusing solely on public health issues emerged and contributed to the disinformation network with the same manipulation techniques: reconfigurations of texts produced by established media, foreign video translations, references to professionals from unrelated areas of expertise in order to increase the trust of readers, and interlinking of the content addressed at pro-Russian and anti-vaccination/Covid-19 doubters (European values, 2021). State security agencies confirmed the existing links between actors behind platforms with manipulative content and protest movements against social distancing and hygiene regulations. “*The Covid-19 pandemic can simply be described as the perfect black swan of intelligence operations,*” the Czech Security Information Service summed up in its 2020 report with a detailed description of threats to the democratic basis of the Czech state (BIS.cz, 2021, p. 9).

Given that mainstream media debunk fake news and expose the players behind alternative platforms, legacy brands have become targeted and verbally attacked by disinformation outlets and their discreditation campaigns in both countries. Cross-border media reports such as the Pandora Papers exposed by the International Consortium of Investigative Journalists are also presented as a directed operation aiming to damage the Czech sovereignty by the EU, Washington or George Soros, and the Open Society Foundation (Šefčíková, 2022). On the other hand, politicians engage with websites from alternative eco-system in different ways. While the major part of the political spectrum in both countries warned citizens against propagandist and untrue reports, mainly on issues related to Covid-19 and Russia or China, the political parties and their members facing criticism by mainstream media would resort to alternative outlets to address their audiences. In several cases, politicians or officials in government positions refused to communicate with legacy newspaper brands or public media and replied to alternative media outlets instead (Šníd, 2018).

Databases of disinformation websites

The rise in the volume of fabricated and manipulated content has attracted interest in fact-checking and investigating sources of fake news. One of the first coordinated initiatives across the region originated in the area of advertising, with a clear motivation to disrupt the business model of untrustworthy online outlets. Even more importantly, the project aimed to protect companies from potentially harmful linking of their brands with controversial content on the internet. In 2016, a group of scholars, journalists, and marketing specialists set up a database titled “conspirators” (Konspiratori.sk). The scheme originally listed 38 websites operating in Slovakia and in the Czech Republic.

The public database is presented by its authors as summarising platforms of “unserious, deceptive, fraudulent, conspiracy and propagandistic content” and it is the most numerous database in the region, currently listing 246 online media outlets (Konspiratori.sk, 2022). Other similar projects founded by journalists and think tanks in the Czech Republic provide less numerous databases (with up to 60 online sites) but they overlap with Konspiratori.sk list and focus on the most influential actors in the Czech market. The terminology referring to the outlets differs from “disinformation websites” to “conspirator and anti-system websites” with minor differences in categorization applied by different projects (Štetka et al., 2020).

In order to achieve higher impact and engagement of articles on social media, producers of content on the listed websites use some of the time-proven tactics, such as copying features of established online media and news websites; mixing commentary with news from subscribed news agency services; providing web design similar to established online news outlets; presenting their reports as investigative journalism without identifying real authors; providing content produced by translating articles from non-journalistic sources; sharing the same texts across the web (Hacek, 2020).

Apart from promoting a political agenda, some of the listed websites focus on health and healthy lifestyle, arms and defence, or religious issues.

In a separate analysis of controversial websites with the highest impact, Syrovatka et al. (2020) proposed a four-fold classification based on the prevalent topics of content and motivation of providers: 1. Esoteric and 2. Preachers (ideology-oriented providers, the most convinced of their truth); 3. Healers (websites focusing on lifestyle and health issues); and 4. Businessmen (current-affairs fake news produced for profits).

Transparency issues of disinformation websites

Various databases of controversial websites are based on varied criteria: while some are exclusively dedicated to pro-Russian propaganda, the Konspiratori.sk works with a set of five criteria: 1. Scientifically unproven, charlatan health – and lifestyle-related content; 2. Hoax, fake news, manipulated content; 3. Conspiracy theories and delusions with negative political, economic, or health-related consequences; 4. Vulgarisms, promoting violence and extremism, aggressive language, and defamation towards religious and ethnic groups; 5. Breach of basic journalistic values and standards – no corrections or information on authors of articles, and organizational and financial structures of online media outlets are published (Konspiratori.sk, 2022).

Websites that can qualify for at least one of the above-listed criteria are included in the database. The online providers of controversial content are reported to the non-profit organisation behind the Konspiratori.sk project by members of the wide public, with a monthly average of 30-50 websites notified (Konspiratori.sk, 2020). The board of independent experts proceeds to scrutinise websites based on the selected criteria. Interestingly, actors behind this initiative have been forced to take several items off the list due to preliminary court orders, as their online providers opposed vigorously their inclusion in the database (Kernová, 2020).

For our study on providers of controversial online content, we focused on the issue of transparency in ownership and authorship as one of the attributes Konspiratori.sk uses when evaluating suspicious websites. We collected data

based on the terms valid in the period between March 2020 and February 2021. The database contained 204 websites (N = 204) in Czech (49.6%) and Slovak language (44.9%). Some webpages have meanwhile ceased to exist (N = 12) which may signal that they had been created for a particular purpose, e.g., before the parliamentary elections. The same applies to websites (N = 17) that were sold to a different provider after they had completed their mission. The rest of them consisted of non-updated websites (N = 38).

In order to research ownership of untrustworthy sites, we analysed contact information to identify their individual representatives (e.g., contact person, editorial board) and publishing house or website operator. The website operator can be represented by a) a limited liability company, b) a civil-law association, c) a non-governmental and non-profit organization, d) an unknown person, e) a private person f) a self-employed person. On the issue of financial income – apart from calls for support published on the websites, we studied their web advertising and e-shop facilities.

The data presented here underline our previous observations about the anonymity of researched websites (Hacek, 2020). Most of the outlets (56.2%) have only impersonal contact – an email address without additional personal information or information about the website provider. Furthermore, 20.4% of websites do not provide any contact. Only 23.4% of the portals have standard contact information. As we expected, most of the websites with controversial content are almost anonymous. Only journalists working with open sources can discover the history of their ownership through public registers, e.g., when a website provider applies for financial support (remittance tax).

Very similar data arise in connection with the variable of authorship of content. Almost half of the published texts come from unknown authors (49%) and 16% of the researched websites published their content partially anonymous. Regarding anonymous content, some of the examined texts are only signed by nicknames.

Figure 1: Contact information of the researched websites

contact information	N	%
contact	32	23,4
only unpersonal contact	77	56,2
no contact	28	20,4

Figure 2: Authorship of published content on the researched websites

authorship of content	N	%
author	48	35
unknown author	67	49
irregularly	22	16

Open data analysis shows that the largest group (52%) of selected websites are provided by completely unknown providers. Civil-law associations provide 19% of the researched websites, followed by limited liability companies (14.6%), private persons (11.5%), non-profit organizations (2.2%), and self-employed persons (0.7%).

Figure 3: Type of the website provider of the researched websites

Type of website provider	N	%
civil-law association	26	19
limited liability company	20	14,6
non-profit organization	3	2,2
private person	16	11,5
self-employed person	1	0,7
unknown person	71	52

Website advertising was present in 54 percent of researched online platforms but a set of other types of financial support requested by readers could be identified, mainly calls for support to the web provider account (37.2%), followed by a combination of account and remittance tax support

possibilities (5.1%). Some websites offered readers a kind of paid membership (4.4 %) and 1.5% of providers asked for a remittance tax. On the other hand, more than half of the websites have no financial support (51.8%) and we can observe also a minority of websites (3.8%) with transparent accounts.

Figure 4: Type of the request for financial support

Request for financial support	N	%
remittance tax	2	1,5
support to account	51	37,2
remittance tax + account	7	5,1
no support	71	51,8
membership fees	6	4,4

As there are a variety of financial profits for the researched websites, in this study, we also focused on the conditions for placing banner ads. As has been mentioned, the project Konspiratori.sk was created with the objective to present a reference for online marketing. Since its introduction, only 54% of websites with controversial content have advertising while 16.8% provided e-shops. The declaration of independence was provided for 34.3 % of websites, as part of contact information.

Figure 5: The website advertising on the researched websites

web advertising	N	%
yes	74	54
no	63	46

Figure 6: The website provides an e-shop

e-shop	N	%
Yes	23	16,8
No	114	83,2

Figure 7: The website declares its independence

Declaration of independence	N	%
yes	47	34,3
no	90	65,7

As can be seen from our analysis, the network of online media outlets producing controversial and untrustworthy content in the Czech Republic and Slovakia is characterised by insufficient transparency regarding the real authors of published articles, as well as the financial background and business activities of their owners. Given that one of the frequent arguments against mainstream media by their staunch critics among politicians or from the alternative media environment is the presumed influence by local oligarchs and foreign investors over news reporting and the content of legacy brands, it is equally significant to stress this evidence of financial grey zone behind alternative and disinformation platforms.

Conclusion: fake news detection as a policy instrument

Various methods of detecting false content are examined by a growing body of research (Porter and Wood, 2021; Margolin, Hannak and Weber 2018; Amazeen, 2017). Overall, fact-checking is regarded by policymakers as a key element of a successful multidimensional approach to limiting the spread and influence of online disinformation (HLEP, 2018). In the Czech Republic and Slovakia, legacy media do not necessarily have full sections dedicated to debunking fake news but they monitor and report on the most blatant and harmful cases massively distributed on social networks. The collected information is presented as databases of untrue statements by local politicians or in form of numerous news articles in print and online media (Echo24.cz, 2022).

Transparency issues of media outlets spreading propaganda and disinformation are a frequent subject of reports by investigative journalists. In the Czech Republic, a team of newspaper and public TV reporters revealed the

owner and only contributor to the third most popular website spreading pro-Russian propaganda (ČT24.ceskatelevize.cz, 2020). Investigative reporters also exposed financial links and cooperation between ultra-conservative religious groups in central Europe and their online outlets, including several based in Slovakia (Dauksza et al., 2020).

Interestingly, citizens in both countries are more likely to say that the media itself should be responsible for combating disinformation than is the average sentiment in the EU; Czechs alone were surveyed as the most reluctant to have public authorities in charge of counter-disinformation action (Eurobarometer, 2020). However, formal measures had been taken on many levels of state and government administration in the region: the Czech police set up a special unit called “Centre against Terrorism and Hybrid Threats” in 2017; the Slovak police followed with a special section to monitor disinformation and hate speech in 2018 and later added a special social media site to debunking hoaxes (Pavleska et al., 2018). During the corona epidemic, police experts on propaganda and mis-/disinformation increasingly turned directly to social media and posted corrections in order to prevent the amplification of false stories (Minv.sk, 2018). Finally, the Slovak police listed a number of Facebook profiles and sites frequently used for spreading hoaxes, in a similar manner as the databases of disinformation websites drawn up and operated by non-state actors.

The most serious crackdown on actors from the alternative and disinformation eco-system up to date has been connected with hate speech and propaganda. Following the criminal investigation into Patriot paper (*Vlastenecké noviny*), the Czech online site was ordered a closure and its owner Radek Velička was charged with four offences including hate incitement (BIS.cz, 2021). The most severe move by the European Union as a whole followed after the outbreak of the military invasion of Russia to Ukraine in February 2022 as several broadcasting and online media generally acknowledged as platforms of Russian propaganda had been blocked. Even prior to the EU action, the Czech provider of the national domain in

coordination with the Czech government acted to block eight disinformation websites on the list of pro-Kremlin outlets for state security reasons (Echo24.cz, 2022). Similarly in Slovakia, the National Security Authority blocked providers of harmful online content following extraordinary legislation adopted by the Slovak parliament (Mikušovič, 2022). These measures – both on national and pan-European level – had been adopted in fast-track procedures and prompted a continuing public debate on various legal, transparency, and freedom of speech-related aspects. Therefore, the consequences of the legislative action remain a question for future research. At the same time, it will be useful to study whether those decisions had really achieved a reduction of the fake news circulating online and subsequently also a diminished polarization and hate speech in each country. Or coming back to the initial metaphor of the fake-news-street lamp in Prague, we shall see if our minds and judgment are clearer without the blurring effect of untruthful content spreading from those blocked media outlets.

Bibliography

- Amazeen, M.A. (2020). Journalistic Interventions: The structural factors affecting the global emergence of fact-checking. *Journalism*, 21(1), 95-111. <https://doi.org/10.1177/1464884917730217>.
- BIS.CZ (2021, November 23). Bezpečnostní informační služba Výroční zpráva 2020, Retrieved from <https://www.bis.cz/public/site/bis.cz/content/vyrocní-zpravy/2020-vz-cz-2.pdf>.
- Brenne, S.J, Felix, S.M., Howard, P.N. and Nielsen, R.K. (2020, April). *Types, Sources and claims of Covid-19 Misinformation*. Retrieved from <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-04/Brennen%20-%20COVID%2019%20Misinformation%20FINAL%20%283%29.pdf>.
- Česká televize (2020, March 3). *Odhalení dezinformátora z Aeronetu: Pátrání ještě není konec, Říka novinář Kundra*. Retrieved from <https://ct24.ceskatelevize.cz/specialy/newsroom-ct24/3057374-odhaleni-dezinformatora-z-aeronetu-patrani-jeste-neni-konec-rika>.

- Dragomir, M. (2020). *Media Influence Matrix. Slovakia*. CEU Center for Media, Data and Society. Retrieved from <https://cmds.ceu.edu/sites/cmcs.ceu.hu/files/attachment/basicpage/1324/mimslovakiafunding2020.pdf>.
- Echo24.cz (2022, February 25). *Česká vláda nechala zablokovat některé weby, které označila jako dezinformační*. Retrieved from <https://echo24.cz/a/ScjgR/ceska-vlada-nechala-zablokovat-nektere-weby-ktere-oznacila-jako-dezinformacni>.
- Eurobarometer (2018, March). *Fake news and disinformation online*. Retrieved from <https://europa.eu/eurobarometer/surveys/detail/2183>.
- Eurobarometer (2020, March). *Attitudes towards the Impact of Digitalisation on Daily Lives*. Retrieved from <https://europa.eu/eurobarometer/surveys/detail/2228>.
- European Commission (2018, April). *A multi-dimensional approach to disinformation*. Retrieved from <https://www.ecsite.eu/sites/default/files/amulti-dimensionalapproachtodisinformation-reportoftheindependenthighlevelgrouponfakenewsandonlinedisinformation.pdf>.
- European Parliament (2019, March). *Automated tackling of disinformation*. Retrieved from [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624278/EPRS_STU\(2019\)624278_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624278/EPRS_STU(2019)624278_EN.pdf).
- Fallis, D. (2015). What is disinformation? *Library Trends*, 63(3), 401-426. <https://doi.org/10.1353/lib.2015.0014>.
- Gabrižová, Z. (2019, November 29). *Slovakia still opposes EU accession to Istanbul Convention preventing violence against women*. Retrieved from <https://www.euractiv.com/section/freedom-of-thought/news/slovakia-still-opposes-eu-accession-to-istanbul-convention-preventing-violence-against-women/>.
- Galeotti, M. (2019). *Russian Political War: Moving Beyond the Hybrid*. London: Routledge. <https://doi.org/10.4324/9780429443442>.
- GDI (2020, March 24). *Why is ad tech funding these ads on coronavirus conspiracy sites*. Retrieved from <https://disinformationindex.org/2020/03/why-is-ad-tech-funding-these-ads-on-coronavirus-conspiracy-sites/>.

- Gregor, M. & Mlejnková, P. (2021). Facing Disinformation: Narratives and Manipulative Techniques Deployed in the Czech Republic. *Politics in Central Europe*, 17(3) 541-564. <https://doi.org/10.2478/pce-2021-0023>.
- Hacek, J. (2020). Antisystémoví voliči na Slovensku a ich dôvera v médiá. In Sámelová, A., Stanková, M., Hacek, J. *Fenomén 2020: Komunita v mediálnom priestore*. p. 63 – 79. ISBN 978-80-223-4974-1
- Hameleers, M., Brosius, A., & de Vreese, C. H. (2022). Whom to trust? Media exposure patterns of citizens with perceptions of misinformation and disinformation related to the news media. *European Journal of Communication*. <https://doi.org/10.1177/02673231211072667>.
- Horne, B. D., Nørregaard, J., & Adalı, S. (2019). Different Spirals of Sameness: A Study of Content Sharing in Mainstream and Alternative Media. *Proceedings of the International AAAI Conference on Web and Social Media*, 13(01), 257-266. Retrieved from <https://ojs.aaai.org/index.php/ICWSM/article/view/3227>.
- Humprecht, E., Esser, F., Aelst, P.V., Staender, A. and Morosoli, S. (2021). The sharing of disinformation in cross-national comparison: analyzing patterns of resilience. *Information, Communication & Society*. <https://doi.org/10.1080/1369118X.2021.2006744>.
- Charap, S. (2015). The Ghost of Hybrid War. *Survival*, 57(6), 51-58, <https://doi.org/10.1080/00396338.2015.1116147>.
- ICJK.SK (2020, December 30). *Zlatý lev reve z Krakova. Za presadzovaním ultrakonzervatívnych hodnôt sa skrýva dobre fungujúci stroj na peniaze*. Retrieved from <https://www.icjk.sk/93/Zlaty-lev-reve-z-Krakova-Za-presadzovanim-ultrakonzervativnych-hodnot-sa-skrýva-dobre-fungujuci-stroj-na-peniaze>.
- Kernová, M. (2020, March 3). *Súd prikázal dočasne stiahnuť web Hlavné správy zo zoznamu konšpirátori*. Retrieved from <https://www.omeiach.com/hoaxy/17314-sud-prikazal-docasne-stiahnut-web-hlavne-spravy-zo-zoznamu-konspiratori-sk-video>.
- Klingová, K., Hajdu, D., and Sawiris M. (2021, November). *Globsec Vulnerability Index 2021 – Slovakia*. Retrieved from https://www.globsec.org/wp-content/uploads/2021/11/Vulnerability-Index_Slovakia.pdf.

- Konšpiratori.sk (2020, May 11). *Prečo niektoré weby boli a už nie sú na zozname konšpirátorov?*. Retrieved from <https://dennikn.sk/blog/1887829/preco-niektore-weby-boli-a-uz-nie-su-na-zozname-konspiratorov/?ref=inc>.
- Kospiratori.sk (2022). *Protect your brand from being associated with controversial content*. Retrieved from <https://kospiratori.sk/en>.
- Krátka Špalková, V. and Činčerová, Z. (2021). „Rok 2020: Jak česká vláda prohrála s dvěma vlnami dezinformací o koronaviru.“ Bezpečnostní centrum Evropské hodnoty. Retrieved from <https://europeanvalues.cz/wp-content/uploads/2021/03/covid-report50.pdf>.
- Krátka Špalková, V., Přívrátský, J., Pastír, M., Činčerová, Z. (2021). *Výroční správa o stavu české dezinformačné scény za rok 2020*. Retrieved from https://europeanvalues.cz/wp-content/uploads/2021/11/Vyrocní_zprava_o_stavu_ceske_dezinformacni_sceny_za_rok_2020.pdf.
- Levinger, M. (2018, September 24). *Master narratives of disinformation campaigns*. Retrieved from <https://jia.sipa.columbia.edu/master-narratives-disinformation-campaigns>.
- Mancini, P., and Zielonka, J. (2012). Introduction. *The International Journal of Press/Politics*, 17(4), 379–387. <https://doi.org/10.1177/1940161212452959>.
- Margolin, D., Hannak, A. & Weber, I. (2018). Political Fact-Checking on Twitter: When Do Corrections Have an Effect?. *Political Communication*, 35(2), 196-219, DOI: 10.1080/10584609.2017.1334018.
- Mayerhöffer, E. (2021). How do Danish Right-wing Alternative Media Position Themselves Against the Mainstream? Advancing the Study of Alternative Media Structure and Content. *Journalism Studies*, 22(2), 119-136. <https://doi.org/10.1080/1461670X.2020.1814846>.
- Mikušovič, D. (2022, February 22). *Vláda chce dat dočasne NBÚ právo blokovat aj weby s dezinformáciami. Česi vypli Aeronet*. Retrieved from <https://dennikn.sk/2738940/vlada-chce-dat-docasne-nbu-pravo-blokovat-aj-weby-s-dezinformaciami-cesi-vypli-aeronet/>.

- MINV.SK (2018, October 25). *Proti podvodom a hoaxom aj na Facebooku PZ*. Retrieved from <https://www.minv.sk/?tlacove-spravy&sprava=proti-podvodom-a-hoaxom-aj-na-facebooku-pz>.
- Monitora.cz (2022). *Umelecká instalace upozorňuje na fake news díky datům z Monitory*. Retrieved from <https://monitora.cz/znalosti/umelecka-instalace-upozorňuje-na-fake-news-diky-datum-z-monitory/>.
- Naeem, B. et al. (2020). A deep learning framework for clickbait detection on social area network using natural language cues. *Journal of Computational Social Science*, 3(231 – 243), <https://doi.org/10.1007/s42001-020-00063-y>.
- OSF (2017, August). *Reklama na konspiračních a dezinformačních webech*. Retrieved from <https://osf.cz/wp-content/uploads/2017/08/reklama-na-konspiracnich-dezinf-webech-web.pdf>.
- Porter, E., Wood, T.J. (2021, September 14). The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. *PNAS*, 118 (37) e2104235118. <https://doi.org/10.1073/pnas.2104235118>.
- Radu, R. (2020). Fighting the ‘Infodemic’: Legal Responses to COVID-19 Disinformation. *Social Media + Society*, 6(3), 1, <https://doi.org/10.1177/2056305120948190>.
- Rogers, R., & Niederer, S. (2020). The politics of social media manipulation. In R. Rogers & S. Niederer (Eds.), *The Politics of Social Media Manipulation* (pp. 19–70). Amsterdam University Press. <https://doi.org/10.2307/j.ctv1b0fvs5.3>.
- Sampor, Ž.M. (2021, July). *Monitoring Media pluralism in the Digital Era. Country report: Slovakia*. Retrieved from https://cadmus.eui.eu/bitstream/handle/1814/71961/slovakia_results_mpm_2021_cmpf.pdf?sequence=1&isAllowed=y.
- SIS.GOV.SK (2021, December). *Správa o činnosti SIS za rok 2020*. Retrieved from <https://www.sis.gov.sk/pre-vas/sprava-o-cinnosti.html#bezpecnostna-oblast>.

- Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*, 34(3), 150–160. <https://doi.org/10.1177/0266382117722446>.
- Staender, A., & Humprecht, E. (2021). Types (Disinformation). *DOCA - Database of Variables for Content Analysis*. <https://doi.org/10.34778/4e>.
- Starbird, C., Arif, A. and Wilson, T. (2017, July 1). Understanding the Structure and Dynamics of Disinformation in the Online Information Ecosystem. *Technical Report*, Retrieved at <https://apps.dtic.mil/sti/pdfs/AD1058867.pdf>.
- Syrovátka, J., Vinklová, J., Wojtula, L. and Zikmundová, A. (2020, March). Dezinformace jako Biznys. *Prague Security Studies Institute*, Retrieved at https://www.pssi.cz/download/docs/8207_751-dezinformace-jako-byznys.pdf.
- Štetka, V., and Hajek, R. (2021, July). *Monitoring Media pluralism in the Digital Era. Country report: The Czech Republic*. Retrieved from https://cadmus.eui.eu/bitstream/handle/1814/71942/the_czech_republic_results_mpm_2021_cmpf.pdf?sequence=1&isAllowed=y.
- Štětka, V., Mazák, J. and Vochocová, L. (2021). “Nobody Tells us what to Write about”: The Disinformation Media Ecosystem and its Consumers in the Czech Republic. *Javnost - The Public*, 28(1), 90-109, <https://doi.org/10.1080/13183222.2020.1841381>.
- Štětka, V. (2015). The Rise of Oligarchs as Media Owners. In J. Zielonka (Ed.). *Media and Politics in New Democracies: Europe in a Comparative Perspective* (pp. 85–98). Oxford: Oxford University Press.
- Štětka, V. and Hájek, R. (2020). Monitoring Media Pluralism in the Digital Era: Application of the Media Pluralism Monitor in the European Union, Albania and Turkey in the years 2018-2019. Country report: Czech Republic. Retrieved from https://cadmus.eui.eu/bitstream/handle/1814/67798/czech_republic_results_mpm_2020_cmpf.pdf?sequence=1&isAllowed=y.

- Šuplata, M., Nič, M. (2016, June 30). *Russia's information war in Central Europe: New trends and counter-measures*. Retrieved from https://old.infocenter.gov.ge/uploads/files/2017-09/1505817613_russias_information_war_in_central_europe.pdf.
- Trend.sk (2017, May 4). *Českí novináři: Babiš používá svoje médiá v politike. Nahrávka to len potvrdzuje*. Retrieved from <https://www.trend.sk/spravy/ceski-novinari-babis-pouziva-svoje-media-politike-nahravka-to-len-potvrdzuje>.
- UN.ORG (2020). *Department of Global Communications: UN tackles 'infodemic' of misinformation and cybercrime in COVID-19 crisis*. Retrieved from <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%98infodemic%E2%80%99-misinformation-and-cybercrime-covid-19>.
- Vojtěchovská, M. (2017). The gradual takeover of the Czech media system. In A. Schiffrin (Ed.) *In the Service of Power: Media Capture and the Threat to Democracy*. Washington, DC: Center for International Media Assistance.
- Wu, S. (2021). As Mainstream and Alternative Media Converge?: Critical Perspectives from Asia on Online Media Development. *Journalism Practice*, DOI: 10.1080/17512786.2021.1976072.

TECHNOLOGIES AND FACT-CHECKING: A SOCIOTECHNICAL MAPPING

Oscar Westlund, Rebekah Larsen, Lucas Graves, Lasha Kavtaradze
& Steen Steensen

Scholars, publishers, authorities, fact-checkers, the public, and other stakeholders have growing concerns about an “information disorder”, fueled by digital media technologies and platforms (Wardle, 2018). Social media platforms often rely on data-fueled revenue models based on selling users’ personal data and profiles as commodities to advertisers. Misinformation, as used throughout this chapter, refers to information that is false, inaccurate, or misleading, whether unintentional or produced for political and financial gain (Wardle & Derakhshan, 2017). Widespread availability of digital technologies has enabled actors, from laymen to governments, to produce misinformation, taking potential form as texts, videos, images, and/or audio. Misinformation is a multidimensional problem with numerous stakeholders. One response to concerns about information disorder has come in the form of increased attention to, resources for, and activity from fact-checking organizations. We approach fact-checking as a problem-solving practice, with fact-checkers searching for relevant and credible information that can be used to verify or debunk claims and other pieces of content. Fact-checking is also a sociotechnical practice insofar as using technology is a prerequisite for solving problems, such as verifying the authenticity of an image or the location of an event. Notwithstanding this, to date there is limited knowledge about this emerging landscape of

digital tools and systems associated with fact-checking. This chapter addresses that void through a systematic mapping of such technologies.

Organizations dedicated to what is called “external” or *ex post* fact-checking (to distinguish from internal, *a priori* verification within newsrooms) first emerged in the US, with staff members devoted to checking truth claims made by political actors (Graves, 2016, 2018). As the movement has grown to global proportions, fact-checkers have developed professional standards and mobilized into an international association called the International Fact Checking Network (IFCN), based in the US (Graves, 2016). Amid growing concerns about information disorder since 2016, the primary focus of many fact-checking organizations has shifted to target viral online misinformation; these fact-checkers often target conspiracy theories and fake media, rather than focusing on the more ‘traditional’ political claims (Graves & Mantzarlis, 2020). The fact-checks produced have recently been brought into fact-checking infrastructures managed by Poynter and Google, respectively. A recent study found that most fact-checkers contribute to Poynter, for which 43% of fact-checks originate from Facebook, compared to 16% for the Google infrastructure. There are also significant differences when it comes to the geographic regions from which the fact-checkers operate (Nissen et. al. 2022).

Concerns about misinformation—spurring such previously described changes in fact-checking practices and the field—have also given rise to a related array of digital technologies designed—or appropriated—to fight misinformation. Technology and platform companies have continuously advanced tools and systems with functionalities for solving specific tasks relating to problems of misinformation. Other actors involved in combating misinformation—journalists, fact-checkers, technologists in journalistic institutions, governmental authorities, NGOs, platform and tech companies—also use a diverse set of technologies in their efforts to identify and fight misinformation (see Graves & Anderson, 2020). Some digital technologies are open-source and can be used freely by various actors, such as fact-checkers, as they see fit. Other technologies are owned and controlled

by commercial companies requiring a subscription plan or participation in a contracted partnership. Scholars have witnessed how a diverse set of stakeholders mobilize to fight misinformation, as no social actor can single-handedly manage this challenge (e.g., Belair-Gagnon et al., 2022; Tully et al., 2021).

The development of emerging technologies has been fueled by a growing anti-misinformation industry or sector, spanning technology firms, universities, news organizations, policymakers, and other stakeholders. Platform companies such as Meta, TikTok, and Twitter have entered partnerships with fact-checking organizations, providing financial remuneration but also new technological tools and systems. Other networks of partnership and innovation include triple-helix collaborations through EU funded consortiums. In 2021 the European Commission granted more than 11 million euros to eight networked hubs of academics, fact-checkers, tech companies and other partners connected to the European Digital Media Observatory (EDMO). Each hub will work towards detecting and analyzing disinformation campaigns, organizing media literacy activities, and offering support to authorities.

Ultimately, the fight against misinformation involves both social actors and digital technologies, as well as a diverse set of institutions. Digital developments enable, and require, new approaches to the critical evaluation of sources and information. This chapter thus draws on a sociotechnical framework that enables analysis of the interrelationships between humans and technology (Lewis & Westlund, 2015). Technological infrastructures, systems and tools available may become integral for fact-checking activities. Importantly, with such great diversity in technologies this means that fact-checkers do not necessarily appropriate each distinct technology, nor do these technologies determine how they will be used.

This chapter develops a mapping of digital technologies associated with fact-checking. Utilizing a sociotechnical framework, we focus on technologies relevant for different fact-checking activities in three broad stages that

span multiple actors and audiences. Employing this lens of stages allows us to identify and discuss available digital technologies (technological actants) through which we decipher and discuss interrelationships with social actors and audiences. This chapter thus advances knowledge around the sociotechnical infrastructures *available to* and *potentially used* in fact-checking—that is, digital technologies that carry affordances for fact-checkers. We carried out a systematic mapping from fall 2020 to spring 2022, identifying the digital technologies through multiple publicly available lists, as well as from fact-checkers, journalists and so forth via our own international interview study (see method section). This chapter also integrates discussion of relevant research and industry reports with our mapping, synthesizing state-of-the-art knowledge about ongoing developments from multiple fields. In 2022, the EU-funded NORDIS project published a brief mapping of fact-checking technologies in Europe and the US, focusing mostly on visual content verification, and generating a list of technologies among which many are not included in ours (Lindén et al., 2022). That mapping, just like ours, approaches technology as an enabler and constituent part of fact-checking practices rather than being an end-to-end solution. This chapter offers further added value through its international approach to the mapping, but also through its sociotechnical approach involving both inductive and deductive analysis. Utilizing a complementary research design, we analyzed and classified these technologies both deductively and inductively (DeCarlo, 2018). The overall classification deductively draws on earlier works suggesting fact-checking consists of three main stages: I) identification, II) verification and III) distribution (Graves, 2018; Nakov et al., 2021). Graves (2018) describes the identification stage as encompassing activities such as monitoring the media and political sources, identifying factual statements, and prioritizing claims to check. He discusses the verification stage in terms of involving checking against existing fact-checks and authoritative sources. Finally, the distribution stage entails flagging repeated falsehoods, publishing fact-checks, and providing contextual data Graves (2018). These three stages serve as a lens for structuring our systematic mapping. We find a concentration of technologies focusing on the identification stage, similarly

to findings in the abovementioned recent mapping by Lindén et al. (2022) about automated technologies (e.g., artificial intelligence, natural language process, and machine learning). Within each of these stages we decided to inductively construct sub-categories based on available technologies, how these can be used, and their key affordances. For each subcategory we also describe in detail a selection of available technologies.

What follows is first a description of the sociotechnical, analytical framework employed in this systematic mapping. This framework guided and informed our inquiries and analysis of the roles played by human social actors, technological actants, and audiences for different digital technologies (e.g., Lewis & Westlund, 2015). Then follows a concise review of contemporary, relevant literature around digital technologies and fact-checking. Thereafter we turn to an analytical and case-based mapping of digital technologies associated with fact-checking, structured by the stages and subcategories briefly introduced above.

Sociotechnical Framework

Scholars in digital journalism and/or disinformation studies have situated misinformation, fake news, and online falsehoods as sociotechnical problems, involving multiple human and social processes by actors and audiences, as well as digital technologies and infrastructures (e.g., Bakir & McStay, 2018; Creech, 2020; Lien et al., 2021). While the fact-checking movement largely grew out of journalism as a cultural and epistemic practice (Graves, 2016), it has increasingly become integrated with large-scale technological systems (Himma-Kadakas & Ojamets, 2022). Throughout the 2010s, numerous (digital) journalism studies scholars have developed analyses that focus not only on human agency, but also digital technologies and their materiality (e.g., Domingo et al., 2015; Primo & Zago, 2015). We find a sociotechnical approach—explicitly examining both the human and non-human—useful in studying digital technologies, particularly in a mediascape marked by powerful platforms with large user bases. Platforms have been studied extensively with different approaches, such as within

infrastructure and platform studies (Plantin et al., 2018), and increasingly also within digital journalism studies (Nielsen & Ganter, 2018; Steensen & Westlund, 2021).

We conceive of platforms as sociotechnical infrastructures that have the power and algorithmic capacity to perform numerous activities and mediate the flows of information and communication for the public (e.g., Nieborg & Poell, 2018). Google (Alphabet) has built its platform and infrastructure on the open web, whereas Facebook (Meta) and Apple have developed proprietary ecosystems (e.g., van der Vlist & Helmond, 2021). One way platform companies exert power is by being digital intermediaries to the public, as their digital materiality (including openness vs. proprietary boundaries) sets conditions for how other institutions (e.g., publishers and fact-checkers) can function (Chua & Westlund, 2022; Lien et al., 2021; Nielsen & Ganter, 2018; Steensen & Westlund, 2021). For example, the algorithmic operations of Facebook's feed, utilizing machine learning models, directs what content is made visible to different users. Facebook feeds fact-checkers with content to check; these fact-checkers then decide, based on their own criteria, what is worth fact-checking. If fact-checks are carried out, the fact-checkers can report false claims in articles and this becomes an input into the Facebook algorithms, which in turn reduces user exposure to the original article. Moreover, Facebook can display misinformation labels and popup messages alerting users who are about to share false-rated content. Lien, Lee & Tandoc (2021) discuss how platform companies have become "prime breeding ground for online falsehoods" because of sociotechnical affordances such as the widespread use, ease of sharing, and their (relatively) limited in-house content-related gatekeepers.

Sociotechnical infrastructures consist of digital materiality carrying a set of affordances that essentially condition (or allow/permit) what is possible to do with a technology or platform. Yet these affordances do not necessarily determine how platforms or technologies are used. Affordance theory, as first introduced by Gibson (1977), stresses how a specific materiality carries both opportunities and constraints, with which users can engage in

different ways. Affordance theory has been used by scholars analyzing contemporary digital platforms and their diverging materiality when it comes to communication practices being enabled and restricted (e.g., Schrock, 2015). For example, scholars have examined how this influences the spread of conspiracy theories (Theocharis et al., 2021). The affordances approach draws attention to how technologies and platforms interrelate with humans in sociotechnical infrastructures. Molina and Sundar (2019) are among scholars addressing the role of affordances for different media and platforms in the context of misinformation and journalists. They enumerate specific affordances of technologies used in modern journalistic practice—modality, agency, interactivity, and navigability—and the concomitant cautions that journalists should take when presented with these ‘action possibilities.’

This chapter focuses on the affordances of technologies associated with fact-checking; these can be understood as non-human technological materialities potentially interwoven with human practices. Thus, an underlying question behind this research is: how do the activities and agency of humans – professional social actors as well as diverse assemblages of audiences/users – intersect with technologies associated with, and having agency for, fact-checking? The sociotechnical theoretical foundation we draw upon in answering this question (in conjunction with affordance theory) is the four A’s framework, consisting of social *actors*, technological *actants*, *audiences*, and *activities* (Lewis & Westlund, 2015). The four A’s grew out of an emerging sociotechnical research tradition in (digital) journalism studies that focuses on the interrelationships between humans and technology. The four A’s sought to recognize the role of diverse social actors (not only journalists and editors), and the emerging roles and agency of technology per se rather than treating it only as something which journalists use or not. Thus, this article first and foremost focuses on the digital technologies (i.e., technological actants), and their interrelationships with activities by way of the three general stages of fact-checking. Moreover, we analyze the myriad entanglements of such digital technologies (including platforms, tools, systems, and algorithms), numerous social actors, and diverse assemblages of audiences.

Ultimately, social actors inscribe agency and affordances into digital technologies (the technological actants) associated with fact-checking. These in turn shape sociotechnical infrastructures by setting conditions for what actors and audiences can do with the technologies, including but not limited to activities associated with misinformation and fact-checking. By deductively applying a lens of the three stages of fact-checking practices, we assess types of technologies associated with fact-checking as sociotechnical infrastructures. We then explore and analyze specific technologies, inductively classifying them into meaningful subcategories based on the activities they enable. Within this categorization, we also discuss related forms of agency between the different actors, actants, and audiences.

Research into digital technology and fact-checking

Research into the intersection of digital technologies and fact-checking remains a nascent field, especially around institutional actors such as fact-checkers, publishers and platforms, although there are recent notable exceptions (Belair-Gagnon et al., 2022; Horowitz et al., 2022; Lien et al., 2022). Collaborations and partnerships have developed between fact-checking organizations and other social actors, many of which are tech- or platform companies (Authors, in review). Following the 2016 US presidential election, Facebook rolled out a partnership programme for fact-checking. Working with independent fact-checkers across all continents, Facebook provides financial remuneration and ad hoc technological tools and systems to fact-checkers. In recent years, TikTok, Twitter, and WhatsApp have also developed partnership programmes, and Google (Alphabet) offers databases, advanced search tools, training programmes and project funding via the Google News Initiative. Graves and Anderson (2020) detail a collaboration with Google to produce a data standard, ClaimReview, which makes fact-checks legible to algorithms and has been embraced by almost all leading fact-checkers, as well as Facebook and multiple search engines.

Broadly speaking, platform and tech companies, with their technological infrastructures and tools, have arguably helped shape fact-checking as a

field. This includes influencing fact-checker practices around identification, verification, and distribution of (mis)information. For instance, one study by Guo et al. (2022) focuses on platform-influenced “misinformation detection” methods (MID) in relation to identification. While technological infrastructures can be used for advancing misinformation identification, there certainly are limits. For instance, a report by UK fact checker Full Fact discusses how fact-checkers find the tool provided by the Facebook partnership too cluttered with non-checkable content; this technical issue also speaks to larger problems around the volume and virality of misinformation (Full Fact, 2019; see also Full Fact, 2020). While there is some relevant computer science work on technologies of misinformation identification (Guo et al., 2022), there is limited research in the fields of communication and journalism studies (exceptions include Hassan et al., 2015; Teyssou et al., 2017).

Other studies have focused on different sociotechnical connections between actors, actants, activities, audiences. For example, Hassan et al. (2015) focus on how fact-checkers use technology to relate to audiences. In another example, Nakov et al. (2021) highlighted the intelligent technologies that support human experts in different fact-checking steps: 1) identifying claims for fact-checking; 2) detecting previous fact-checked claims; and 3) retrieving evidence to verify a claim (see also Karadzhev et al., 2017). Platform companies such as Facebook employ artificial intelligence to assist with human content moderation, including within partnerships with fact-checkers (Iosifidis & Nicoli, 2020). Finally, multiple studies have reported on companies working on technologies enabling automated fact-checking (AFC) (Graves, 2018; Hassan et al., 2015). Assessing the veracity of a claim can be automated at different stages, including a) identifying check-worthy factual claims, b) matching claims to the existing credible sources, and c) ascribing them veracity value (Saquete et al., 2020). As a sociotechnical phenomenon, AFC is in a nascent phase, though researchers agree that the goal of AFC is to assist actors with manual fact-checking (Thorne & Vlachos, 2018). In another example of sociotechnical connections, a pioneering study by Lien et. al (2021) assessed and analyzed public statements and press releases by a selection of platform companies (Meta, Alphabet and Twitter), discussing

these as aimed at image repair, with platforms trying to claim social responsibility via sociotechnical efforts. The platforms funded media literacy efforts and supported fact-checking programs (Lien et al., 2021). Platforms must navigate challenges of freedom of expression vis-à-vis content moderation; this is an area in which platforms have controversially claimed they are neutral content hosts, rather than publishers with responsibility toward content (e.g., Gillespie, 2018; Napoli, 2020).

Approach, method, and material

Fact-checkers can be understood as specialized professionals, sharing fundamentally similar skills and are expected to approach misinformation in similar ways. Fact-checking practices are linked to a multitude of digital technologies, including relatively general and mundane technologies (e.g., e-mail, phone, and common word processing software). For example, given significant coordination in their work (similarly to traditional journalists), fact-checkers also often use general digital collaboration tools such as Slack or Teams (Koivula et al., 2020). Around collaboration, fact-checkers also use more specialized technologies such as *AirTable*, *Chequeo Colectivo* and *Truly Media Workbench*.

While our research acknowledges a multitude of technologies, our mapping is structured around more specific technologies associated with fact-checking (including those created for other purposes but that can be used for fact-checking problems). As stated in the introduction, we draw on previous research distinguishing three key stages for fact-checking practices: 1) identification, 2) verification, and 3) distribution (Graves, 2018; Nakov et al., 2021; c.f. four stages by Vlachos & Riedel, 2014). This chapter applies these three stages as a deductive framework for the first analytical step in mapping technologies. These stages cover most types of fact-checking work, and in practice may overlap; for instance, identification of potentially false content may require initial verification work. Some technologies serve only one stage whereas others address several, in some cases without human intervention. For example, Nakov et al. (2021) discuss how the tool Squash

(developed by Duke Reporters' Lab) can capture audio, transcribe it into text to identify claims to check, and then find and compare results to existing fact-checks. The mapping includes other end-to-end solutions, such as the *InVid* platform developed by *WeVerify*. Accessed as a browser plugin, *InVID* offers multiple tools, such as a reverse image search, a metadata extractor for visual contents, a keyframe extractor for videos, and an advanced Twitter search. Other examples include: a) *InfoFinder*, which functions as a data-oriented helpdesk through which fact-checkers can get help with collecting and verifying data; and b) the so-called engagement management systems and community management systems by *Hearken*.

We carried out a systematic review of multiple data sources, both primary and secondary, collected from 2020-2022. They encompass a multitude of digital technologies made available on the market, on an open-source basis, or via partnership agreements. In the first phase of research, the mapping was systematically populated via technologies mentioned in the context of fact-checking in industry reports, trade press, databases of tools, fact-checking conferences, scholarly literature, and interviews for an ongoing research project. Two of these data sources have been especially important. First, tools were identified from primary data, obtained during an international interview study of over 50 fact-checkers, journalists, and technologists working with them. Conducted by the authors between 2020-2021, these interviews included questions around tools and resources in combating misinformation and interrelated cross-sector partnerships (Authors, in review). Second, the mapping was populated with existing lists of digital tools from open-source investigation projects. This includes lists from projects in academia and human rights (Bellingcat and the University of Cambridge's Whistle project) as well as material curated by investigative journalists such as BuzzFeed's verification expert Craig Silverman. Our mapping thus represents a current snapshot of what digital technologies are relatively available for fact-checking practices. These technologies are used/developed by fact-checking or related organizations, or developers make explicit references to fact-checkers or journalists as end-users.

The resultant mapping comprises a total of 136 digital technologies. We charted each with descriptive information, including various affordances (e.g., functions for text and media content), languages available, and any indication of usage (e.g., information from companies about customer/user base). It also includes information on the ongoing development and provenance of the technology—its funding/ownership, the organization(s) involved, and its stated original and eventual aims. For our mapping of digital technologies relevant for fact-checking practices, we have predominantly relied on discourse and descriptions from website information, industry press and reports, and by interviewees. When possible, one or several members of the research team would engage more with a particular tool (e.g., participating in tool training, testing out interfaces) to develop more familiarity with the technical possibilities and constraints. (Future research for this project includes a more systematic interface analysis, or ‘technography’, of technologies identified in the mapping.) Obviously, there is a continuous development of emerging technologies relevant for mappings like ours, with some technologies also being discontinued. Thus, we stress that our mapping is not exhaustive on the level of individual technologies; we argue that our snapshot analysis yields meaningful findings and analytical categories for longer-term conceptualization.

In the second phase, we analyzed the identified technologies in two interrelated steps. First, we deductively relied on the three fact-checking stages as our lens for systematic mapping and analysis. We first categorized all technologies based on their affordances relevant to identification, verification, and dissemination. (Some technologies occupied multiple categories; i.e., as some technologies are end-to-end services, they can fit in each stage.) Second, we inductively thematized the material for each stage into meaningful subcategories; analysis is centered on the roles and interrelationships played by the technologies vis-à-vis actors and audiences in fact-checking activities. The sociotechnical approach, via the four A’s framework, has been integral to our inductive analysis in the second step.

Sociotechnical mapping of technologies associated with fact-checking

This section reviews technologies associated with the three stages of fact-checking practice. Each subsection sets the context and then discusses the technologies by way of inductively constructed subcategories within each stage. In connection to the 4A's framework, the three stages represent overall *activities* integral to fact-checking practices, including identifying checkable claims, verifying different types of content and sources, and producing fact-checks suitable for different kinds of distribution. Social *actors* include three main groups: fact-checkers/publishers, tech providers, and platform companies. Fact-checkers and publishers are combined as both have journalistic connections (e.g., organizational, ideological, work routines). Tech providers include non-profits, open-source efforts, NGOs, and universities, as well as commercial tech companies requiring payments/subscriptions. Some technological tools and systems are developed through triple-helix collaborations in which the university sector participates in formative stages. Platform companies include organizations providing digital platforms that offer services and functions to users, and act as digital intermediaries. (This chapter focuses on technologies for which a cross-sector partnership with the platform companies is not a prerequisite.) *Audiences*, in the context of news work, have been conceptualized as recipients, commodities, and active participants. In terms of *audiences*, the mapping focuses mainly on whether technologies carry affordances for involving audiences as active participants in the three stages of fact-checking.

The 136 *technological actants* are marked by heterogeneity; some operate autonomously whereas others function as tools operated by humans. Lewis and Westlund (2015) discuss diverse human/manual and computational modes of orientation and output. In a subsequent conceptual mapping of the human-machine divide, they explore how X form of journalism depends on Y form of technology, and outline four specific facets of technological dependence (Lewis & Westlund, 2016). Similarly yet more simply, here we differentiate technological actants into two overall categories: 1) tech-led,

and 2) human-led. The tech-led approach refers to technological actants designed to operate more autonomously, providing fact-checkers with information and/or actionable guidance. The technologies in the identification stage are predominantly tech-led, featuring sophisticated systems by platforms and tech companies that generate actionable information. The human-led approach refers to technologies requiring human actors to engage with these systems more actively and manually for specific purposes and in distinct practices. The verification stage is marked by human-led technologies, for example, image verification tools that fact-checkers can use in their efforts to verify authenticity related to an image. The four A's framework is used for the mapping of technologies in each of the three stages of the fact-checking practice.

Table 1 offers a concise visualization of our sociotechnical framework and mapping of technologies associated with fact-checking. The three stages of fact-checking on the vertical axis can guide the reader assessment, each of which is paired with the four A's on the horizontal axis.

Table 1: Sociotechnical framework for technologies associated with fact-checking

	Activities	Actors	Actants	Audiences
Identification	Identifying checkable claims	Fact-checkers, publishers	Tech-led	Content flagging
	Traffic analysis	Tech providers	Human-led	
	Prioritization	Platforms		
Verification	Source verification	Tech providers	Human-led	Tech-led
	Content verification		Tech-led	
Distribution	Fact-check publishing	Fact-checkers, publishers	Tech-led	Sharing
	Social media circulation	Platforms	Human-led	
	Live	Tech providers		
	Algorithmic visibility reduction			
	Personalization			
	Packaging			

For the *identification* stage, the most central activity involves identifying checkable claims; related activities include analysis of traffic/engagement and prioritization. While some publishers and fact-checker actors have developed proprietary tech (i.e., owned and controlled in-house), they are largely dependent on third parties such as tech and platform companies at this stage. The identification stage section is thus structured via an actor perspective (though some technologies tap into audience participation by enabling flagging of potential misinformation). The technological actants are mainly characterized as tech-led, feeding human social actors with actionable input.

Turning to *verification*, tech actors are most active in developing technologies, whereas platforms and fact-checkers are largely absent when it comes to owning, funding, and developing verification technologies. We recognize the importance of most tech for verification being owned and controlled by tech actors, but this also means that publishers, fact-checkers, and platforms have limited importance; thus, the actor category does not help differentiate verification technologies much. Moreover, verification appears to be human-led insofar that human social actors engage in significant manual work with technological actants to generate actionable outputs. Following our inductive analysis, verification will be discussed as an activity-oriented category, focusing on 1) source verification and 2) content verification.

Finally, the section focusing on *distribution* is the most concise; it mainly distinguishes between the channels and platforms proprietary to publishers/fact-checkers, and distribution associated with platforms. It is characterized as tech-led since technological actants effectively distribute fact-checking content without much human interplay as far as we are able to see into these closed proprietary systems.

Technologies for identification

The practice of identifying potential misinformation involves monitoring the mediascape via routine search and mapping of specific sources,

sites and platforms – using both mundane and specialized technologies. This monitoring may be carried out manually by human actors with the assistance of technological tools and systems for specific tasks (human-led); alternatively, technologies automatically take action, or produce information and other forms of actionable outputs (tech-led). There are both tech-led and actor-led technologies for identification, but we contend that this category is predominantly tech-led. These are mainly technological systems linked to infrastructures with data, designed to gather, analyze, and present specific information to users. Identification encompasses textual content as well as audio, images, and videos which must be assessed for indicators of misinformation. The identification stage also includes establishing whether potentially misinforming content can be fact-checked using available evidence and following established standards; for example, the IFCN Code of Principles includes commitments to nonpartisanship, fairness, and the transparency of sources, methodology, and organization. Finally, given limited resources and an abundance of potential misinformation, this stage also involves prioritization to focus fact-checking efforts on the most important, widespread, or potentially harmful material.

We find that the landscape of identification-related technologies is the most fully developed of the three stages we consider in this study, with the widest array of tools available to assist fact-checkers with different facets of the identification process. Nevertheless, no single integrated system or tool exists to continuously cast a net over the World Wide Web—much less the wider mediascape—detecting all signs of potential mis- or disinformation that surface. Barriers to establishing such a holistic system range from regulatory regimes, ethical considerations around privacy, and language diversity to the fact that platform companies maintain closed infrastructures that restrict monitoring and analysis. As a result, fact-checkers rely on a diverse mix of specialized tools for particular tasks or environments, developed by tech providers, platform companies, or in many cases by fact-checkers themselves. For example, Chequedo's in-house *Chequeabot* scans broadcast and print media outlets to identify potentially misleading

political claims, while Istinomer, a Serbian fact-checking organization, won an IFCN grant to develop a browser extension to help readers identify problematic stories. Fact-checker Lead Stories has developed a proprietary social-media monitoring platform called *Trendolizer*, which facilitates their own fact-checking work with Facebook (Meta) and which they also license for revenue. It indexes trending content, as well as source code. Through features such as the fingerprinting tool, *Trendolizer* also allows for potential identification of networks of misinformation-spreading sites.

Identifying claims is a vital, value-laden, and time-consuming activity for fact-checking organizations; it is central to how they define and execute their mission. As a result, the affordances of platforms and technologies for identification condition how fact-checkers work. We have inductively focused on two distinct subcategories of actors providing technologies for identification. The first category consists of diverse tech companies providing systems and tools, and includes technologies developed in-house by publishers and fact-checkers. Technologies for identification of claims are mostly offered by external tech providers and platform companies, and for some platforms a partnership with them becomes a prerequisite. The second category consists of technologies attained through platform partnerships. Next, we discuss the types of technologies (actants) offered by each of these actors; we explore how fact-checking activities related to misinformation identification on platforms with relatively closed infrastructures de facto necessitate a cross-sector partnership with platform companies.

Identification tech (in-house or tech companies)

Multiple tech providers have developed technologies that help fact-checkers to identify and prioritize misinformation. Some technologies focus on monitoring and analyzing information on the open web, some focus on social media monitoring, and others do both. Content moderation is a central aspect of what platform companies do (Gillespie, 2018). Amid the growth of misinformation, there has been mounting pressure on platform companies to address it. Some offer solutions for monitoring multiple sites and/or

platforms, continuously assessing information and misinformation flows. Several tech companies have developed multi-platform detection solutions. For example,

- *DataMinr* offers journalists, fact-checkers, and other actors opportunities to detect stories, by daily collecting and analyzing billions of public data inputs and sending alerts to users. *DataMinr* advertises the use of artificial intelligence to process data from numerous publicly available sources and analyze multiple types of formats such as texts, video, images, and sound.
- *InVid* is a verification plugin for web browsers, developed by the EU-funded *WeVerify* consortium of journalistic, academic, and company partners. It is free to use and gathers no data from its users, and the creators present it as a verification “Swiss army knife” for fighting disinformation. *InVid* features affordances for analyzing metadata, keyframes and forensics, and has an integration with multiple social media platforms to access and analyze content such as Facebook videos and YouTube thumbnails.
- A technology developed in-house by fact-checkers, *Trendolizer* enables fact-checkers to search and identify networks of actors using the same IP or IDs for advertising or Google analytics etc.

Other multi-platform identification tools include *WebMii* for analyzing social media profiles across platforms, and *Social Blade*, which tracks metrics of users across Facebook, Twitter, Youtube, Instagram, Twitch, etc. There are also technologies for analyses of specific social media platforms. Twitter makes most of its content publicly available, and thus it can be monitored and analyzed using technologies developed by tech providers. There are thus numerous tools for gathering Twitter data.

- *Tweetdeck* enables users to search and analyze content on just Twitter, and is categorized as a general technology given its availability to diverse actors.

- *AccountAnalysis* offers details about when and what type of content accounts publish, what accounts they interact with, and what websites they share.
- *Botometer* analyzes the activity of a Twitter account and gives it a score, where higher scores indicate more bot-like activity, which can be mobilized to create false impressions on engagement, attitudes and so forth.

Other tools, such as *InfoLeaks*, provide information on specific Twitter users by using content, image, and geographical information. *Foller.me* offers affordances for analyzing the most popular words used, whereas *Hoaxy* visualizes Twitter conversations by topic, illuminating influential accounts and the origins of hashtags.

CrowdTangle is a social media monitoring tool owned by Facebook, provided to its fact-checking partners (as well as a selection of journalists and academics). *CrowdTangle* enables its users to monitor engagement with different types of content on Facebook, Instagram, and Reddit, in public accounts and groups. For the Russian platform VK, an image and facial recognition tool enables search with a face and find any potential profile user; for Instagram, there is a “search engine” to find users without having an account (called *SearchUsers*). *MediaWise* can be used to monitor TikTok content, and *Map-snapchat.com* enables heatmap visualization of public usage over the last 48 hours. *Searchmy.nio* enables search on Instagram for descriptions.

Some platforms, especially messaging applications like WhatsApp, Signal and Telegram, facilitate closed and encrypted communication; this makes monitoring harder and tools are in shorter supply. *Telegago* is a sort of search engine that can be used to search Telegram contents. Rappler gathers tips and screenshots from the public using closed messaging apps. Aos Fatos, Maldita and others have launched WhatsApp chatbots, technological actants enabling users to offer leads and content about potential misinformation. Maldita employs a natural language response model that collects,

classifies, and analyzes content data associated with misinformation, and its chatbot is capable of detecting and responding to disinformation reports in video, audio, text and image formats. The chatbot is also capable of interpreting and fact-checking responses.

Let us turn specifically to the central activity of identifying checkable claims, for which fact-checkers regularly turn to technologies offered by tech providers. In the context of developing AI for identifying check-worthy claims, Nakov et al. (2021) discuss how such technologies should be capable of producing relevant lists of claims deemed worthy of checking, preferably containing ranking and check-worthiness scores to guide fact-checkers (who in turn should be able to give feedback on the scores to continuously calibrate the system). AI technologies are utilized to identify claims in social media (text, audio, or video), news, and websites as well as broadcast. Other areas of development include technologies for quickly checking whether specific claims have already been fact-checked, and technologies for exploring whether false claims have been translated into multiple languages (Nakov et al., 2021). Moreover, Google has worked on a significant open-source language model called the Bidirectional Encoder Representations from Transformers (BERT) to identify misinformation (e.g., on social media), and is used by fact-checkers and trained with annotations by them to improve its capacity to identify relevant claims. There are also technologies using bots to identify statements that can be fact-checked (e.g., *Transcriptions*) as well as finding claims within a text and ranking them based on plausibility and their checkworthiness (*Logically*).

Fact-checking companies have also invested in developing technologies with AI to quickly identify claims from text transcripts. Faktisk.no in Norway employs natural language processing (NLP) technologies to assist manual information identification. The technology automatically highlights in different colors the personal and organization names, geographical locations, numbers, percentages, and other relevant figures in TV transcripts. The technology also marks the potential claims in bold that might be worth paying extra attention to, helping human fact-checkers in navigating the

text faster, reading only the highlighted parts and disregarding the excess information. *Squash* from Duke University Reporters' Lab also uses audio-to-text conversion technologies before certain claims are identified, ranked, compared to previously existing fact-checks and presented on-air TV. *ClaimBuster* is the umbrella name for several IDIR Lab fact-checking projects focusing on exploring and developing automated fact-checking. The *ClaimBuster* website states that it "started as an effort to create an AI model that could automatically detect claims worth checking. Since then, it has steadily made progress towards the holy grail of automated fact-checking." *ClaimBuster* provides lists of checkable claims figuring in some form of text. Their *ClaimPortal* enables users, mostly fact-checkers, to register and track specific accounts on social media. Whenever those accounts make claims, the user will receive an alert with a curated message. *ClaimBuster* facilitates assessing if a claim has previously been fact-checked, checking material from Twitter and political statements against fact-checks published by Snopes and PolitiFact (US-based fact-checkers). Full Fact, a UK-based fact-checker, has developed its own technology for claim-spotting. Its technology searches viral content online and on social media as well as a selection of political transcripts and news media. Full Fact is also developing technology for checking tweets with claims about images. Full Fact has stated that it aims to develop technology to assist fact-checkers in knowing the most important things to fact-check, identifying when known misinformation is spread, and reducing time delays in fact-checking close to real-time (Full Fact, 2022).

In addition to academic, NGO, and fact-checking organizations, profit-oriented start-up companies also utilize and develop AI technologies for information verification. One of them is the Norway-based organization Factiveuse that has developed an AI-based editor (i.e., technological actant) that identifies text claims and sorts them based on whether they are supported or disputed in authoritative sources. When deciding upon the veracity of the claim, the final word still belongs to the human fact-checker. Ultimately, while Full Fact, ClaimBuster, and Factiveuse use automation to identify

claims across different types of materials, their technologies are feeding into human activities, enabling a sociotechnical fact-checking practice.

Finally, some fact-checkers have developed tools to enlist audiences in identifying misinformation. As previously noted, Serbia's Istinomer is developing a browser extension to make it easy for readers to report potentially problematic stories. Another example comes from Correctiv in Germany, which has developed an interface called *CrowdNewsroom* that allows community members to help identify potential misinformation. Launched in 2015, *CrowdNewsroom* is described as "...a kind of virtual editorial team which allows editors, reporters and readers to cooperate on major investigations. CrowdNewsroom provides the individual tools for any type of investigation, for use by everyone involved."

Identification tech and platform partnerships

Platform companies have used internet infrastructures to develop proprietary areas of the web. They control their technological affordances and infrastructures, including but not limited to the visibility of information and communication. Some global platform companies offer technologies customized for fact-checking practices, but only to their formal cross-sector partners. Platform companies (e.g., Meta and Alphabet) have developed technologies that allow entrusted collaborators, such as fact-checkers and publishers, to gain in-depth insights into areas such as trending topics, viral content, and audience engagements. Following the problems with mis- and disinformation associated with the 2016 US presidential election, Facebook (now known as Meta) initiated what was to become a global, cross-sector fact-checking partnership programme. In addition to financial remuneration, Facebook has developed a technological system for its partners, called the *Facebook Fact-Checking Product*. When using this product, the fact-checker logs into their personal Facebook account to access a specific interface. The fact-checking product populates a list of articles, possibly containing misinformation, ranked by relevance, and linked to parameters such as content virality and the geographical coverage area of the

fact-checker. When fact-checkers carry out identification and verification, and subsequently publish their fact-checks, they report back to the platform via the *Facebook Fact-Checking Product*. Subsequently, Facebook employs fact-checks to calibrate the algorithmic visibility of the published posts. The fact-checkers in the partnership programme access the product based on institutional affiliation in a specific country. Facebook employs artificial intelligence as well as audience flagging to identify potential misinformation, and to generate the geographically related queue to the fact-checkers for further checking. While the *Facebook Fact-Checking Product* is offered only to a selected few, the company offers technologies such as the aforementioned analytics infrastructure *CrowdTangle* to a wider set of professionals.

In terms of other platforms, TikTok has also rolled out a fact-checking program with international partners. TikTok has developed a fact-checking system via which fact-checkers access a continuously updated list of content potentially containing misinformation. The TikTok system interface lets fact-checkers swipe among a selection of content having been detected by tech-led TikTok algorithms. As of 2022 TikTok has partnered with far fewer fact-checkers than Facebook, but is continuing to expand their programme. Some platforms have also engaged their audiences as active participants in identification of misinformation, using technological actants to flag content not in line with their community principles. For example, Facebook gives users options for flagging content as well as advertisements. In flagging a post as inappropriate, the users can select from a range of categories, including nudity, violence, harassment, suicide, false information, junk mail, unallowed sales, hatred, terrorism, and other. Content flagged will be processed by content moderators, and in cases of false information it may be sent for independent review by independent fact-checkers.

Technologies for verification

Tech providers dominate the space for verification technologies, and source verification and content verification are the two central activities. The identification stage may lead to the verification stage, in which fact-checkers use

their own skills and various methods and technologies to thoroughly assess content veracity. Some technologies naturally overlap with those available for identification mentioned earlier. Numerous and diverse technologies are relevant for verification, encompassing searching for relevant information and facts, as well as identifying and potentially contacting the source of a claim. Verification includes significant efforts towards making assessment of sources and contents. Once again there are general and more mundane information-and-communication technologies offering multiple affordances for verifications, which fact-checkers can utilize in their work process. For example, common search engines offer advanced properties, such as (reverse) image searches and geographical position. There are also several technologies that have been specifically developed and designed to be used for verification. Some of these are open-source technologies whilst others are commercially driven and require a partnership or license. For example, the European-based open-source WeVerify platform/plugin, InVid, was developed with multiple partners (AFP, Deutsche Welle, ATC and universities); it houses purpose-built technologies to be used by fact-checkers and journalists engaging in verification. While open source, the platform does maintain certain advanced tools that require an institutional email address and administrative approval.

Moreover, applied research seeks to advance verification processes through automated fact-checking. One such procedure involves developing technologies that retrieve and analyze a collection of documents, advancing a verdict for assessment by humans. Researchers have problematized the challenges of automatic verification; e.g., automation is useful for specific, narrow tasks, such as identifying existing fact-checks or evidence from databases that supports or rejects claims about numerical values. This is compared to more nuanced or newer claims, e.g., that Covid-19 vaccines have been developed too quickly and thus are risky (Nakov et al., 2021). Verification can be a complicated, time – and resource – demanding practice, also challenged by the fact that techniques and technologies for verifying images and videos do

not always work. Human social actors, technological actants, and audiences all participate in verification in different ways.

Verification practices can be defined as the searching for information and sources that can help verify or dismiss a claim, a piece of media, or any form of report or indices that has been identified as important to look further into. Some technologies enable verification altogether, in assessing sources and types of content. Other technologies are essentially geared towards making specific tasks of verification easier for fact-checkers; this can include affordances to remove backgrounds from images (*remove.bg*), extract and analyze metadata (*IrfanView*, *Context Aggregation and Analysis* tool from WeVerify) or make text transcripts of media files (audio/video) to enable quicker fact-checking. Our mapping of verification technologies has yielded two main types: *source verification* (the entities behind content), and *content verification* (verifying claims within media content). Some technologies are also used for both types of verification (e.g., *Wayback Machine*, *Claimbuster*). For each type, we assess the potential role played by actors, actants, and audiences.

Source verification

Source verification refers to the practice of identifying the entities behind a piece of content or a claim; these entities can include different actors such as individuals, groups, organizations, or countries. Source verification is often a critical first step in making assessments about the likely credibility of a given claim or piece of content. Source verification entails technologies but also analyses of digital and social media, which, for instance, corresponds to verifying the actors connected to social media profiles. Thus, technologies for source verification encompass verification on social media and websites, respectively.

In 2021 alone, Meta identified and closed billions of fake Facebook accounts connected to misinformation. Despite efforts on this scale, across platforms there is still a flurry of fake social media profiles and accounts, as well as

fake web and social media traffic, often involving significant use of bots (Braun & Eklund, 2019). Ultimately there is a need to identify and verify social media profiles, accounts, and traffic. Amid this, multiple technologies enabling verification of social media accounts have emerged. These include numerous technologies used for misinformation identification that have also proved worthwhile for identification of actors.

- *WebMii* is a technology that makes it possible to verify social media profiles for different social media platforms, aggregating public information about people.
- *What's My Name* is a service for searching a username across platforms.
- *Trendolizer* makes it possible to verify sources among networks of online users by assessing IP addresses or their advertising or Google Analytics IDs. Several technologies are used to verify social media profiles as fake accounts. *AccountAnalysis* enables systematic assessment and verification of Twitter accounts, analyzing what type of content Twitter users publish, when they do so, what links they share, and with whom they interact.
- *Botometer*, formerly known as *BotOrNot*, is a website created by The Observatory on Social Media and the Network Science Institute at Indiana University; this tool checks the activity of Twitter accounts to assess the likelihood of their being bots.
- *TrulyMedia* is a technology that also monitors viral topics and sources and can be used to assess the trustworthiness of Twitter accounts.
- *TruthNest* provides a bot probability score, as well as metrics around the account's activity, network, and influence.
- *Social Blade* collects and analyzes user metrics across social media (Facebook, Instagram, Twitter, Youtube, Twitch), working to identify potential misinformation by checking viral content.

We found that fact-checking technologies enable or facilitate social media verification in a variety of ways, answering questions such as: 1) Who owns the accounts/profiles?, 2) Is there linked activity to other social media platforms?, 3) Is the source a human or a bot?, and 4) What is the virality of the content?

Information flows on websites are also being updated continuously; there are technologies associated with fact-checking that can be used to evaluate web pages over time. *Wayback Machine* is a browser plugin that enables users to access different versions of a website over time. *Website Informer* gathers and presents detailed information on websites via a widget or URL. There are a range of technologies providing information about domain ownership and IP address history (e.g., WhoIs, *DNSlytics*, *Iris*). Moreover, there are technologies to efficiently solve specific tasks, such as checking registries for domain names (e.g., *NORID* for Norwegian domain names), or analyzing whether reviews on Amazon/Yelp/TripAdvisor are authentic (*FakeSpot*).

Content verification

Fact-checkers regularly single out different types of materials to prioritize for fact-checking. As mis- and disinformation comes in many shapes and forms (Wardle, 2018), at the identification stage, some content requires the use of specific technologies. In other cases it is the verification process that simply becomes more efficient when technologies are used. To date, the fact-checking community has mostly gravitated towards fact-checking of textual claims made by known sources such as politicians that can be approached by traditional means for verification. However, the digital mediascape has encouraged visual practices. Smartphones have sophisticated cameras and access to mobile apps for editing videos and images, and social media platforms carry affordances for publishing such visual elements easily. Visual cultures have grown strong on social media platforms such as Instagram, TikTok, and Snapchat. Amid such developments, concerns

have been raised about the broad accessibility and ease of use of technologies for editing, as well as manipulating and fabricating, images and videos. Examples of ‘deep fakes’ have gone viral on social media, illustrating the sophistication with which actors can fabricate videos of people talking, and thus making it appear they make certain claims (Kietzmann et al., 2020). Accordingly, we have identified two types of content verification where the technologies come into the picture: 1) claim verification, and 2) media content verification.

Claim verification

- Verifying claims made by public figures is at the very core of fact-checking as an epistemic and occupational activity. Fact-checking organizations develop their own protocols and routines to verify claims. Though, importantly, whether it is done manually or computationally, the information verification process is almost impossible without utilizing some sort of technological actant. Apart from rather mundane tools traditionally used in media work, claim verification technologies include numerous trackers and databases offering a gateway to check for statistical information (*Statista*) or statements and promises made by politicians or institutions (e.g., *Fact-Base*, *Promise Tracker*).
- Google (Alphabet) has a wide portfolio of services; this includes the suite of “Google FactCheckTools”. Google’s “FactCheck Explorer” service enables users to find fact checks produced associated with specific keywords, themes, and people. Google has also partnered with fact-checkers, e.g., providing UK fact-checker Full Fact with funds and Google fellows to assist in developing AI tools to identify, collate, and compare political claims across a multitude of sources (Dudfield, 2021).

Both Google and Facebook have also supported ClaimReview, a tech-led digital tagging system helping search engines and platforms to find and display published fact-checks. Moreover, AI-powered claim verification technologies assist fact-checkers in comparing claims against databases or repositories

with previously circulated public statements and fact-checks (*Claimbuster*). Other technologies assist in identifying what has been published previously online on specific sites (e.g., *Wayback Machine*). Such technologies might come especially handy while covering live events, when the claim should be verified or debunked in a timely manner (*Squash*). Datasets or repositories containing existing fact-checks are especially important for training AFC algorithms. Though several training datasets are already in use (e.g., those from Politifact and Snopes), the lack of available, high quality, and up-to-date training data remains as one of the main hurdles for automating the verification stage of fact-checking (Kotonya & Tony, 2020).

Media content verification

Multiple technologies exist for media content analysis, or the determining of the authenticity of images, videos, and audio. Reverse image search is a common media verification practice; it can be conducted via search engines (e.g., Google, Yandex, Baidu) or via specifically designed technologies (*TinEye*, *RevEye*).

- Another technology, WeVerify's *InVid*, as well as the *YouTube Data Viewer*, extract data such as thumbnails from YouTube videos to facilitate reverse image searches; they also extract other data such as exact upload times to facilitate analysis of original video source vs. copies.
- *WeVerify Context Aggregation and Analysis* technology enables analysis and verification of images and videos on social media by producing reports that draw on the content and contextual data. The results might be combined with Google Maps or Google Earth, to verify surroundings, time and/or weather applications. Relatedly, *YouTube GeoFind* lets users view geographically tagged videos on a map, sorted by channel, topic and location, and can prove useful in verifying events. Moreover, there are technologies that offer more detailed image analysis.

· *Forensically* is another image forensics tool for identifying fabricated or manipulated images, with affordances for clone detection, noise analysis, and luminance gradient. The Visual Geometry Group (VGG) at Oxford University has developed an open-source project to search for faces and images, enabling classifications and analysis of altered media.

Another important area of development is audio and video forensics. WeVerify has collaborated with the German research institute Fraunhofer to develop a technology called *Digger*. It can be used to identify deep fakes during fact-checking but it is also used by forensics experts in law enforcement.

Technologies for distribution

One underlying premise for fact-checking legitimacy is effectiveness in curbing the spread of misinformation. However, reaching the right audience, in the right way, and at the right time to counteract a false message poses an enduring challenge. Fact-checking operations vary widely in their popularity, and most maintain limited direct reach online. Many outlets rely on an established media parent or partner to reach a wider audience via print or broadcast (e.g., Graves & Cherubini, 2016). However, emerging research suggests that even well-known outlets rarely reach individuals exposed to a particular message with the relevant fact-check (Guess et al., 2018). Porter, Wood and Bahador (2019) raise the concern that distributing fact-checks via fact-checkers proprietary channels risks reaching only those already paying attention to fact-checks. False information spreads “farther, faster, deeper, and more broadly than the truth” online (Vosoughi et al., 2018).

This suggests a profound need for tools and systems to help with promoting and directing fact-checkers’ work to relevant audiences. However, as our mapping analysis indicates, this is the least developed of the three stages in terms of available technologies, significantly outnumbered by tools supporting identification and verification. Distribution of fact-checks via social media presents a particular challenge, especially as platforms often use algorithms to identify content their users like the most and feed them more

similar content. One study found YouTube users watching videos with misinformation were subsequently exposed to more such videos (Hussein et al., 2020). Platform companies have publicly and repeatedly disclaimed responsibility as publishers; yet they are not mere hosts of content, but content moderators (Gillespie, 2018). Misinformation oftentimes spreads across multiple platforms simultaneously, but the conditions for fact-checking diverge. Platform companies such as Meta, Google, Twitter, and TikTok have been partnering with fact-checkers in fighting misinformation, with Meta and TikTok distributing fact-checks on their platforms, and Google supporting developments of the ClaimReview project, by providing credits for advertising to enable greater visibility for fact-checks (along with G-suite and Cloud storage). Arguably, the most important non-proprietary technology, in terms of targeting fact-checks to new occurrences of a false claim, is the open data standard ClaimReview. This initiative emerged from the fact-checking community (Graves & Anderson, 2020). ClaimReview makes it possible for search engines and social networks to identify fact-checks and match them to a search query or a social media post.

Another fact-check distribution intervention, implemented on Facebook and Twitter, comes in the form of cartoon figures aimed at reaching audiences in the right way. Called Tooties, these cartoons kindly point out incorrect claims, and then have been used to experiment with different forms of fact-check intervention. The authors of the study behind these figures found that some of these fact-check interventions had a positive impact on perceptions of disputable claims (Opgehaffen, 2022). Platform companies can utilize their algorithmic capacities for displaying fact-checks prominently, such as with Facebook using a “Related Articles” function to show fact-checks relevant to a post. Moreover, for content flagged as misinformation by independent fact-checkers and passed onto Facebook via its fact-checking product, Facebook substantially reduces the visibility of such content in the news feeds (Full Fact, 2019). This is a fundamentally important form of content moderation that Facebook refers to as “reducing the spread” of problematic content, and which we refer to as algorithmic visibility reduction.

Such content moderation means original posts remain and thus do not jeopardize freedom of expression, which Facebook has been criticized for in the past (e.g. Gillespie, 2018).

Our mapping of digital fact-checking tools pinpoints that distribution technologies can be used for the production phase of fact-checking, as well as for sharing checked information to the audiences. For instance, *Canva*, *InfoGram* (tools for making graphs) or *Adobe* or other design programs can be used for producing fact-checks in user – or social media friendly formats. Decisions regarding the packaging or formatting of fact-checks affect the overall visibility of the final product of the whole process. Naturally, the role of technologies in these decisions is paramount.

As a part of a distribution phase, technologies such as *Dega*, *Facebook Fact-Checking Product*, *Fatima*, *Squash*, and *FactStream* are used for increasing the exposure and efficient distribution of fact-checks. In this regard, we distinguish between proprietary and non-proprietary platforms and technologies for distribution. First, we have fact-checking and publisher platforms owned and controlled by fact-checkers themselves (proprietary). Fact-checkers and news agencies may also utilize proprietary digital platforms such as news sites, mobile applications, fact-checking blogs, newsletters and so forth to display fact-check articles. Just like news publishers, they turn to analytics infrastructures to generate metrics that assist them in learning how the public and more specific target groups engage with the content. Fact-checkers publish on proprietary platforms, and cross-promote across other media and digital infrastructures. Some fact-checkers have partnerships with legacy news media to distribute fact-checks via print, such as *Correctiv* distributing fact-checks on a weekly basis via free print newspapers circulated to the wider German population. Fact-checkers also use (third-party) technologies for live fact-checking operations. For example, some technologies have enabled distribution of fact-checks in real-time. This includes a second-screen app (*FactStream*) that displays relevant fact-checks pushed out by fact-checkers during live events. Also operating with the live format, *Squash* is a pop-up fact-checking tool based on the *ClaimBuster* engine

that retrieves relevant fact-checks in real time based on audio transcription. Brazilian social media users are enlightened by an automated chatbot focusing on Covid-19 related disinformation. The chatbot, titled *Fátima*, is available on Twitter and WhatsApp, and uses a database administered by media startup Aos Fatos. The *Fátima* chatbot searches for tweets containing references to debunked stories, and engages in a conversation, providing them with a link to the verified information.

Second, many fact-checkers also distribute their work directly on social media platforms where misinformation thrives. This is done in multiple ways. Fact-checkers and publishers use their editorial content management systems to redistribute their work, which may require human intervention to manually share their fact-checks on social media. Moreover, the public can be enrolled as active participants and cross-circulate fact-checks. In addition, platforms also use algorithms to personalize exposure of content, including fact-checked contents, to their users. IFCN fact-checkers have enrolled into cross-sector partnerships with platforms. The partners of the Facebook fact-checking programme must tag the fact-checks they publish as part of the programme. Articles deemed false in fact-checks, performed by independent fact-checkers, will yield much less exposure (approximately around 5-10 percent of the expected average exposure). Facebook calibrates their algorithms and incorporates the fact-checks into content potentially put on display (e.g. Authors, in review).

Importantly, published fact-checks are not merely “distributed” to audiences via fact-checkers’ proprietary and non-proprietary platforms, they are embedded into technological infrastructures with affordances for algorithmic selection and personalized exposure, sometimes also interactive human-machine communication with bots. This includes how publishers can use editorial content management systems and analytics to customize weight given to fact-check exposure; how social media platforms personalize feeds of content; and how search engines may prioritize display of fact-checks linked to sensitive topics.

Concluding discussion

This chapter sets out to map and analyze existing digital technologies associated with fact-checking; one main aim is to develop an example of a flexible sociotechnical approach for understanding this rapidly changing space. We deductively organized technologies based on three core fact-checking stages of practice, and then inductively developed meaningful subcategories based on their affordances for fact-checking guided by a sociotechnical framework and the four A's. Here we share four main, initial takeaways from this mapping and conclude with thoughts around how this work can inform future research.

First, to date a remarkable number and diversity of technologies associated with fact-checking has emerged, including numerous tools developed specifically by and for fact-checkers. While this chapter does not explicitly show how fact-checkers use technologies in their practices, it nevertheless shows the availability of relevant technologies, many of which carry affordances that are impossible for humans to replace with manual labor in order to solve the misinformation problems at hand. We contend this reinforces our sociotechnical approach, underlying the importance of approaching contemporary fact-checking as a practice in which human social actors engage with technological materiality (technological actants) in their fact-checking practice. We call for more research into fact-checkers' concrete practices with technologies in approaching and solving specific misinformation problems.

Second, the identification stage is the most populated in terms of available technologies, despite (or perhaps because of) the difficulties of monitoring vast amounts of information online. These identification technologies are predominantly owned and controlled by platform and tech companies. This includes technologies for identifying specific claims or content to check, but also a range of technologies for monitoring and analyzing online traffic, which fact-checkers rely on to prioritize their work. Importantly, such technologies are programmed with affordances in which metrics, cues

and variables are used as indicators or proxies for potential misinformation. Digital technologies help identify internet activities around content that stand out as unusual. There are tech-led systems for monitoring social media traffic and activity, generating information and/or directions that human fact-checkers can act upon. Social media monitoring is a daunting task, and whereas some platforms such as Twitter have open API enabling third party tech companies to tap into data, other platforms have erected walled gardens. For platforms with closed infrastructures, such as Facebook and TikTok, the fact-checkers essentially must enroll in cross-sector partnerships to gain access to how misinformation spreads. The identification stage is one in which audiences are enrolled as active participants, encouraged to crowdsource by flagging content that potentially contains misinformation.

Third, the verification stage is predominantly characterized by human-led technologies developed by tech providers, and to some extent in-house by fact-checkers and publishers. The human-led technologies refer to systems and tools that fact-checkers manually use to solve specific problems and tasks, such as image- or location verification, and require tacit knowledge and human action to actively use the technology to yield something meaningful. Platform companies have little presence in the verification stage, and the same applies to audience participation. There are technologies that can be used for assessing the authenticity of media content such as image or exploring whether footage has been tampered, yet there are limits to such tech and its affordances for assessing inauthentic audio- and visual materials.

Fourth, distribution largely takes place via proprietary digital platforms as well as through media partners, with some innovative efforts for mobile apps with personalization and chatbot functionality. There is some tech available for producing and packaging fact-checks in communicatively appealing ways. Social media platforms are important means for redistribution, and via platform partnership programs the *potential* effectiveness of fact-check distribution arguably increases significantly Facebook employs algorithms to feature fact-checks prominently, and claims to reduce visibility and

further spread of materials determined false. Ultimately, we contend that fact-checkers face significant challenges in distributing and targeting their fact-checks effectively, amid a platformized mediascape in which platforms predominantly use closed, commercial algorithmic infrastructures to make more widely desired content available to their audiences.

Let us extend our concluding discussion by drawing on the preceding takeaways. More generally, the technologies associated with fact-checking discussed in this chapter feed into an understanding of contemporary fact-checking as a sociotechnical endeavor. Some tools are tech-led, with degrees of automation and operated by technological actants; others are human-led and thus require social actors to engage. Recognizing that fact-checking practices are sociotechnical, we should not stay at recognizing how human fact-checkers are dependent on an array of technologies, but also recognize the power possessed by the organizations owning and controlling these technologies. Most importantly, we stress the tremendous power that platform companies have when it comes to the technologies associated with identification and distribution. This comes as no surprise as many platform companies have developed their niches by creating sophisticated platform infrastructures using algorithms, operating as commercial enterprises with limited-to-no transparency into their inner workings or widespread impacts. Fact-checkers must engage in cross-sector partnerships with the platforms, and the platforms in turn must develop efforts of their own to effectively engage in content moderation on their proprietary platforms. The effectiveness and orientation of such content moderation efforts vary over time, and among platform companies.

This pioneering study, and the above initial takeaways, certainly close some knowledge gaps around the current state of sociotechnical fact-checking systems. Yet it also sets the stage for a variety of directions for future research. For example, in connection to questions of platform power and transparency, which fact-checking activities can be performed with open-source technology only? Given the vital enabling role of the ClaimReview standard in particular, what else can be achieved by open standards developed

collaboratively by key stakeholders? Cross-sector partnership programmes obviously create opportunities, but what are the tasks and problems that cannot be managed effectively without such partnerships? Technologies for identification are closely connected to social media monitoring in other fields, yet little is known about how human fact-checkers deal with audience engagement metrics vis-à-vis the precise needs for identifying and working with checkable claims. Relatedly, more granular research could focus on significant and changing usage by fact-checkers across these technologies; even if some fact-checking activities can be conducted with open source technology only, what other factors or systems might keep fact-checkers from using them?

Given that this chapter represents a snapshot, comparative mappings in the future could also provide researchers and fact-checkers better insight into the sociotechnical factors affecting fact-checking technologies and practices over time. Table 1 can serve as both a template and a source of future research endeavors—how might the population of the table change over time, but also its categories? Will there be increased concentration—and has it already begun—in terms of what technologies are used and how? Or will the sociotechnical aspects—actors, actants, audiences, maybe even activities—proliferate? Are there historical comparisons we can draw, using this approach, in terms of new knowledge creation practices? This could in turn provide more insight into other questions around actor influence and responsibilities when it comes to combating misinformation. For example, this mapping, combined with discourse analysis such as Creech (2020) around platform responsibilities, could provide a more comprehensive sociotechnical picture of the accountability of social media platforms around misinformation. We also call for future research into how fact-checkers appropriate and use diverse technologies in their fact-checking practices. Future research should also include more in-depth case studies and technographies of particularly significant technologies being used by the fact-checking community, in order to generate deeper and potentially more actionable insights into sociotechnical factors in fact-checking practices.

Acknowledgements

The authors are thankful to the members of the OsloMet Digital Journalism Research Group for offering constructive feedback on an earlier version of this chapter at a seminar in June 2022. Particular thanks to Karoline Ihlbeck and Tine Ustad Figenschou. We also thank the external opponent Henrik Örnebring for his invaluable feedback at this seminar.

References

- Bakir, V., & McStay, A. (2018). Fake News and The Economy of Emotions. *Digital Journalism*, 6(2), 154–175. <https://doi.org/10.1080/21670811.2017.1345645>
- Belair-Gagnon, V., Graves, L., Kalsnes, B., Steensen, S., & Westlund, O. (2022). Considering Interinstitutional Visibilities in Combating Misinformation. *Digital Journalism*, 10(5), 669–678, DOI: 10.1080/21670811.2022.2072923
- Braun, J., & Eklund, J. (2019). Fake News, Real Money: Ad Tech Platforms, Profit-Driven Hoaxes, and the Business of Journalism. *Digital Journalism*. <https://doi.org/10.1080/21670811.2018.1556314>
- Chua, S., & Westlund, O. (2022). Platform Configuration: A Longitudinal Study and Conceptualization of a Legacy News Publisher’s Platform-Related Innovation Practices. *Online Media and Global Communication*, 1(1), 60–89. <https://doi.org/10.1515/omgc-2022.-0003>
- Creech, B. (2020). Fake news and the discursive construction of technology companies’ social power. *Media, Culture & Society*, 42(6), 952–968. <https://doi.org/10.1177/0163443719899801>
- DeCarlo, M. (2018). *Scientific Inquiry in Social Work*. Open Social Work Education. <https://scientificinquiryinsocialwork.pressbooks.com/>
- Domingo, D., Masip, P., & Costera Meijer, I. (2015). Tracing Digital News Networks. *Digital Journalism*, 3(1), 53–67. <https://doi.org/10.1080/21670811.2014.927996>

- Dudfield, A. (2021, April 2). How AI helps us detect 100,000 potential claims a day [NGO]. *Full Fact*. <https://fullfact.org/blog/2021/apr/ai-google-100000--claims-day/>
- Full Fact. (2019). *Report on the Facebook Third Party Fact Checking programme*. Full Fact. <https://fullfact.org/media/uploads/tpfc-q1q2-2019.pdf>
- Full Fact. (2020). *Fighting the causes and consequences of bad information (The Full Fact Report)* [Annual]. Full Fact. <https://fullfact.org/media/uploads/fullfactreport2020.pdf>
- Full Fact. (2022). *Automated Fact Checking*. Full Fact. <https://fullfact.org/about/automated/>
- Gibson, J. J. (1977). The Theory of Affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, Acting, and Knowing: Toward an Ecological Psychology* (pp. 67–82). Lawrence Erlbaum Associates. https://monoskop.org/images/c/c6/Gibson_James_J_1977_1979_The_Theory_of_Affordances.pdf
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (Illustrated edition). Yale University Press.
- Graves, L. (2016). *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*. Columbia University Press.
- Graves, L. (2018). *Understanding the Promise and Limits of Automated Fact-Checking* (p. 8) [Factsheet]. Reuters Institute. <https://reutersinstitute.politics.ox.ac.uk/our-research/understanding-promise-and-limits-automated-fact-checking>
- Graves, L., & Anderson, C. (2020). Discipline and promote: Building infrastructure and managing algorithms in a “structured journalism” project by professional fact-checking groups. *New Media & Society*, 22(2), 342–360. <https://doi.org/10.1177/1461444819856916>
- Graves, L., & Cherubini, F. (2016). *The Rise of Fact-Checking Sites in Europe*. Reuters Institute for the Study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/our-research/rise-fact-checking-sites-europe>

- Graves, L., & Mantzarlis, A. (2020). Amid Political Spin and Online Misinformation, Fact Checking Adapts. *The Political Quarterly*, 91(3), 585–591. <https://doi.org/10.1111/1467-923X.12896>
- Guess, A., Nyhan, B., & Reifler, J. (2018). *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign*. <https://www.semanticscholar.org/paper/Selective-exposure-to-misinformation%3A-Evidence-from-Guess-Nyhan/a795b451b3d38ca1d22a6075dbb0be4fc94b4000>
- Guo, Z., Schlichtkrull, M., & Vlachos, A. (2022). A Survey on Automated Fact-Checking. *Transactions of the Association for Computational Linguistics*, 10, 178–206. https://doi.org/10.1162/tacl_a_00454
- Hassan, N., Adair, B., Hamilton, J. T., Li, C., Tremayne, M., Yang, J., & Yu, C. (2015, October). *The Quest to Automate Fact-Checking*. Computation + Journalism Symposium, New York, USA. <http://cj2015.brown.columbia.edu/papers/automate-fact-checking.pdf>
- Himma-Kadakas, M., & Ojamets, I. (2022). Debunking False Information: Investigating Journalists' Fact-Checking Skills. *Digital Journalism*, 10:5, 866-887, DOI: 10.1080/21670811.2022.2043173
- Horowitz, M., Cushion, S., Dragomir, M., Gutiérrez Manjón, S., & Pantti, M. (2022). A Framework for Assessing the Role of Public Service Media Organizations in Countering Disinformation. *Digital Journalism*, 10:5, 843-865, DOI: 10.1080/21670811.2021.1987948
- Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 48:1-48:27. <https://doi.org/10.1145/3392854>
- Iosifidis, P., & Nicoli, N. (2020). The battle to end fake news: A qualitative content analysis of Facebook announcements on how it combats disinformation. *International Communication Gazette*, 82(1), 60–81. <https://doi.org/10.1177/1748048519880729>
- Karadzhev, G., Nakov, P., Marquez, L., Barron-Cedeno, A., & Koychev, I. (2017). Fully Automated Fact Checking Using External Sources. *ArXiv:1710.00341 [Cs]*. <http://arxiv.org/abs/1710.00341>

- Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135–146. <https://doi.org/10.1016/j.bushor.2019.11.006>
- Koivula, M., Villi, M., & Sivunen, A. (2020). Creativity and Innovation in Technology-Mediated Journalistic Work: Mapping out Enablers and Constraints. *Digital Journalism*, 0(0), 1–18. <https://doi.org/10.1080/21670811.2020.1788962>
- Lewis, S. C., & Westlund, O. (2015). Actors, Actants, Audiences, and Activities in Cross-Media News Work. *Digital Journalism*, 3(1), 19–37. <https://doi.org/10.1080/21670811.2014.927986>
- Lewis, S. C., & Westlund, O. (2016). Mapping the Human–Machine Divide in Journalism. In *The SAGE Handbook of Digital Journalism* (pp. 341–353). SAGE Publications Ltd. <https://doi.org/10.4135/9781473957909>
- Lien, C. H., Lee, J., & Tandoc, E. C. (2022). Facing Fakes: Understanding Tech Platforms’ Responses to Online Falsehoods. *Digital Journalism*, 10:5, 761-780, DOI: 10.1080/21670811.2021.1982398
- Lindén, C.-G., Dang-Nguyen, D.-T., Salas-Gulliksen, C., Khan, S. A., Amelie, M., & Dierickx, L. (2022). *State of the Art in Fact-checking Technology*. NORDIS. https://datalab.au.dk/fileadmin/Datalab/NORDIS_reports/Report_task_4.2_-_The_State_of_the_Art_of_Fact-Checking_Tools.pdf
- Molina, M. D., & Sundar, S. S. (2019). Technological Affordances Can Promote Misinformation. In J. E. Katz & K. K. Mays (Eds.), *Journalism and Truth in an Age of Social Media*. Oxford University Press.
- Nakov, P., Corney, D., Hasanain, M., Alam, F., Elsayed, T., Barrón-Cedeño, A., Papotti, P., Shaar, S., & Martino, G. D. S. (2021). Automated Fact-Checking for Assisting Human Fact-Checkers. *ArXiv:2103.07769 [Cs]*. <http://arxiv.org/abs/2103.07769>
- Napoli, P. M. (2020). Connecting Journalism and Public Policy: New Concerns and Continuing Challenges. *Digital Journalism*, 8(6), 691–703. <https://doi.org/10.1080/21670811.2020.1775104>

- Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), 4275–4292. <https://doi.org/10.1177/1461444818769694>
- Nielsen, R. K., & Ganter, S. A. (2018). Dealing with digital intermediaries: A case study of the relations between publishers and platforms. *New Media & Society*, 20(4), 1600–1617. <https://doi.org/10.1177/1461444817701318>
- Oppenhaffen, M. (2022). Fact-Checking Interventions on Social Media Using Cartoon Figures: Lessons Learned from “the Tooties.” *Digital Journalism*, 10:5, 888-911, DOI: 10.1080/21670811.2021.2011758
- Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293–310. <https://doi.org/10.1177/1461444816661553>
- Porter, E., Wood, T. J., & Bahador, B. (2019). Can presidential misinformation on climate change be corrected? Evidence from Internet and phone experiments. *Research & Politics*, 6(3), 2053168019864784. <https://doi.org/10.1177/2053168019864784>
- Primo, A., & Zago, G. (2015). Who And What Do Journalism? *Digital Journalism*, 3(1), 38–52. <https://doi.org/10.1080/21670811.2014.927987>
- Saquete, E., Tomás, D., Moreda, P., Martínez-Barco, P., & Palomar, M. (2020). Fighting post-truth using natural language processing: A review and open challenges. *Expert Systems with Applications: An International Journal*, 141(C). <https://doi.org/10.1016/j.eswa.2019.112943>
- Schrock, A. R. (2015). Communicative Affordances of Mobile Media: Portability, Availability, Locatability, and Multimediality. *International Journal of Communication*, 9(0), 18.
- Steensen, S., & Westlund, O. (2021). *What is Digital Journalism Studies?* Routledge. <https://www.routledge.com/What-is-Digital-Journalism-Studies/Steensen-Westlund/p/book/9780367200909>

- Teyssou, D., Leung, J.-M., Apostolidis, E., Apostolidis, K., Papadopoulos, S., Zampoglou, M., Papadopoulou, O., & Mezaris, V. (2017). The InVID Plug-in: Web Video Verification on the Browser. *Proceedings of the First International Workshop on Multimedia Verification*, 23–30. <https://doi.org/10.1145/3132384.3132387>
- Theocharis, Y., Cardenal, A., Jin, S., Aalberg, T., Hopmann, D. N., Strömbäck, J., Castro, L., Esser, F., Van Aelst, P., de Vreese, C., Corbu, N., Koc-Michalska, K., Matthes, J., Schemer, C., Sheaffer, T., Splendore, S., Stanyer, J., Stepińska, A., & Štětka, V. (2021). Does the platform matter? Social media and COVID-19 conspiracy theory beliefs in 17 countries. *New Media & Society*, 14614448211045666. <https://doi.org/10.1177/14614448211045666>
- Thorne, J., & Vlachos, A. (2018). Automated Fact Checking: Task Formulations, Methods and Future Directions. *Proceedings of the 27th International Conference on Computational Linguistics*, 3346–3359. <https://aclanthology.org/C18-1283>
- Tully, M., Madrid-Morales, D., Wasserman, H., Gondwe, G., & Ireri, K. (2022). Who is Responsible for Stopping the Spread of Misinformation? Examining Audience Perceptions of Responsibilities and Responses in Six Sub-Saharan African Countries. *Digital Journalism*, 10:5, 679-697, DOI: 10.1080/21670811.2021.1965491
- van der Vlist, F. N., & Helmond, A. (2021). How partners mediate platform power: Mapping business and data partnerships in the social media ecosystem. *Big Data & Society*, 8(1), 20539517211025060. <https://doi.org/10.1177/20539517211025061>
- Vlachos, A., & Riedel, S. (2014). Fact Checking: Task definition and dataset construction. *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, 18–22. <https://doi.org/10.3115/v1/W14-2508>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science (New York, N.Y.)*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>

- Wardle, C. (2018). The Need for Smarter Definitions and Practical, Timely Empirical Research on Information Disorder. *Digital Journalism*, 6(8), 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*. <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>

THE ROLE OF FACT-CHECKING IN FIGHTING THE 'INFODEMIC' OF DISINFORMATION ON COVID-19: A CASE STUDY OF POLÍGRAFO

Marina Ferreira & Inês Amaral

Introduction

The Covid-19 pandemic created an unprecedented scenario in terms of public health, with the determination of unprecedented social isolation measures and massive information sharing on social networks – classified by the WHO as an “infodemic”. However, in the context of political communication on the pandemic, several mistakes were made in an early pandemic phase, mainly due to the scarcity of information and scientific evidence on the new disease. In addition, several leaders used the pandemic caused by the new coronavirus as a weapon for political combat, disseminating false information according to their governmental needs and personal beliefs. In response to the alarming amount of misinformation about SARS-CoV-2 and the dangers it poses to public health, widespread mobilisation of fact-checking platforms has been observed in an attempt to identify and correct false or misleading information.

Although its origins date back to previous decades, the journalistic activity of fact-checking began to become widespread in the USA in the 1990s and early 21st century. The first targets of the scrutiny method were mainly figures in the North American political spectrum, with election seasons standing out as periods of excellence for verifying potentially misinformative statements made by media political figures.

The massive use of online spaces and social media have completely transformed the way individuals access and interact with information, leaving the role of mere passive spectators and becoming an active part of the information process until then managed almost exclusively by traditional media.

This new scenario, defined as ‘post-truth’, in which most individuals acquire information through online media not registered as media, brings about complex consequences. The first of them is a growing discrediting of traditional media, which often end up being associated with a role of complicity and partiality concerning political elites, who are also successively seen as targets of much of the information circulating online.

In so-called common scenarios, this indiscriminate circulation of information carries a severe risk of hurting democratic institutions. In that case, this threat intensifies during electoral processes and in scenarios of attempted arrival to power of extreme right-wing movements. For example, the 2016 US elections, the Brexit referendum in the UK that took place in the same year and the Brazilian elections in 2018 were electoral periods that left an indelible mark on online information distribution. On these occasions, misinformation or the presentation of biased facts were used as a means to acquire electoral advantage, for example, through the creation of bots that sowed alternative facts with thousands of users and that influenced, in an unprecedented way, the results of elections, to the point of victory for some around the world.

In the first months of 2020, a new disinformative focus emerged: the new virus detected in the Chinese city of Wuhan, which was named Covid-19. The disease spread worldwide with alarming speed and severity for public health. At the same pace, information sharing about the virus spread on social media, which became a stage for fake news and conspiracy theories about the disease, its treatment methods and the measures taken to contain it. Most newspapers and international fact-checking platforms then directed their attention to fact-checking related to the pandemic to minimise the effects of the rapid “infodemic” of misinformation about the new coronavirus.

Through the analysis of the specific case of the Polígrafo online fact-checker, this chapter aims to observe which are the main focuses of disinformation on the topic in Portugal and their origins.

A brief history of fact-checking

The term fact-checking can have two distinct meanings within journalism (Ireton & Posetti, 2018), a fact relevant to analysing its roots. Graves (2013) points out that references to “reviewers” integrated into North American newsrooms date back to the 18th century. These professionals had the function of checking and ensuring the accuracy of the information in journalistic pieces before being published, i.e., they operated as a second sieve of the facts presented by the journalists’ authors. American weekly magazines such as TIME in the 1920s were the first publications to integrate this type of professionals in their staff (Ireton & Posetti, 2018). However, with the beginning of the 21st century and the adjacent economic and financial revolution that the vast majority of media organisations have faced (and continue to face), these internal fact-checking departments have been reduced or eliminated. Nowadays, few media outlets retain full-time positions of editors as fact-checkers. Among the holdouts are the US magazine The New Yorker and the German weekly publication Der Spiegel (Bloyd-Peshkin & Sivek, 2017).

Modern fact-checking has its origins in twentieth-century America (Amazeen, 2020). However, at that time, this journalism category presented itself in different formats and with different goals from its current concept. The so-called *muckrakers*, presented by Amazeen (2020) as the precursors of today’s fact-checkers, were American writers who provided detailed journalistic accounts of the political and economic corruption of large companies. For example, journalists verified and exposed manufacturers’ claims in medicinal patenting as false. The controversial and revealing articles by muckrakers such as Samuel Hopkins Adams and Upton Sinclair paved the way for the passage of regulations guaranteeing

consumer protection and the implementation of a reform in the American public health system against the unfair practices these companies imposed (Cassedy, 1964).

Dobbs (2012) argues that the beginning of the modern fact-checking movement can be identified during the candidacy of Ronald Reagan, who became, in 1981, the 40th US president. As a presidential candidate, Reagan made several peculiar statements that captured attention for their level of falsity, for example, his famous claim that trees caused four times as much pollution as cars and industrial chimneys combined, continuing in succession to refer to trees as a threat to the environment during and after the election campaign. After he occupied the White House, several journalists began trying to verify his claims during press conferences and televised speeches in an attempt to correct factual errors, but, as Dobbs (2012) points out, this kind of activity was not accepted by much of the public, generating a wave of criticism and hostility towards the verifiers, mainly from the targets of scrutiny.

Thus, and as described by Amazeen (2020), the disinformation transmitted in the North American political scene only began to generate more attention as of the 1990s, more specifically during the period of the 1998 North American presidential elections in which George H. W. Bush was elected. A new and already more defined type of fact-checking then emerged that focused essentially on the statements of political programmes and advertisements, as well as on the speeches and debates that took place during the campaigns. However, it was not until the 21st century that the first projects fully dedicated to political fact-checking began to emerge (Amazeen, 2020), such as Factcheck.org, launched in 2003, PolitiFact.com and the Washington Post's Fact Checker in 2007. Before that, in 1994, a non-political fact-checking website had already been created that was dedicated to investigating urban legends and myths and publicising fraudulent schemes: snopes.com, and this is the oldest and largest online fact-checker in the US. In healthcare, the fact-checking platform HealthNewsReview.org was born in 2004.

The new century brought what Graves and Cherubini (2016) define as a new democratic institution: independent fact-checkers, especially political ones. Although it was in the US that they were born and developed in a more numerous and specialised way, in Europe and the rest of the world, several projects dedicated to fact-checking began to emerge, mainly from 2010. In addition to the analysis of statements by politicians and public figures, which gave the motto to the creation of the first fact-checkers, a new online scenario has emerged with new agents of information dissemination that, without media literacy and possessing, in most cases, financial and/or political-ideological interests, generate misinformative content that, on a large scale, has the potential to hurt democratic institutions, especially during periods of elections and political crisis (Amaral & Santos, 2019).

The “post-truth” and the new gatekeeping agents

The massification of misinformation and disinformation, as well as its circulation is necessarily associated with the digital ecosystem and social networks, media where confronted with new agents that start to circulate content and, in some cases, even replacing traditional gatekeepers (Singer, 2014).

With the emergence in the first instance of blogs and later of social media as networks of content (Amaral, 2016), the audience now takes an active role in selecting and sharing content, a task that was almost entirely reserved to journalists, editors and producers of the media. Before, these publics were limited to the function of passive receivers of news through their online presence. In that case, they can now interact with the already mentioned traditional gatekeepers (Singer, 2014).

Singer (2014) states that the public begins from the 90s onwards to have an active voice on new media. However, unlike what happens with journalism professionals, the new actors “are not affected by professional codes of conduct, tending rather have personal preferences and emotions as criteria for the evaluation and selection of the content to (re)publish” (Amaral & Santos, 2019, p. 65).

As Bakyr and McStay (2017) theorise, the contemporary misinformation and disinformation phenomena directly relates to several features of the modern landscape of media organisations. Thus, the two authors highlight five preponderant factors for the creation of a perfect storm that has massified the dissemination of false information (Bakyr & McStay, 2017, p. 5): “the financial decline of traditional media organisations; the growing immediacy of the news cycle; the rapid circulation of disinformation created by social media users and propagandists; the increasingly emotive nature of online discourse; and the increasing number of people financially capitalising on the algorithms used by social media platforms and search engines”.

The issue of emotionality mentioned shows the role of selective exposure, i.e., a theory that proves that the human mind tends to choose the information that is aligned with its beliefs, attitudes and behaviours, and rejects what is contradictory (Manjoo & Porter, 2008, p. 763). It is in this scenario that each individual manages, more or less consciously, to personalise their informational environment (Spinelli & Santos, 2018), researching and even elaborating content that meets their interests and beliefs, these being validated and often disseminated within the digital spheres in which their ideological peers move.

Bakir and McStay (2017) highlight the critical role of algorithms and online cognitive systems in the formation of so-called echo chambers – which can be defined as closed informational environments where ideas or beliefs are amplified in a defined system in which contrary and dissenting opinions have no place and shared information is biased and partial. The authors note that as a consequence of these echo chambers are created that are formed through the action of algorithms that are based on data collected about the user, such as their connections and interactions with other users on social media, their searches and online activity.

“By keeping each user’s digital footprint recorded, algorithms are able to manipulate this data – called big data – which exists on a large scale and is very complex for specific purposes that are defined and programmed”

(Amaral & Santos, 2019, p. 75). It is precisely this “datification” of personal information and online activity that, by allowing the creation of the aforementioned echo chambers (Bakir & McStay, 2017), enhances the “bubbles” of information that allow “social distortion through narratives of ‘alternative facts’” (Amaral & Santos, 2019, p. 76).

As van Dijck (2014) suggests, despite having been adopted as a technique of knowledge by companies, governments and researchers who take as a basis this meta-data to analyse and predict human behaviour, “datification” has also become what she define an “exercise of faith” in the entities that collect the data. In this way, truth and independence “become are controversial notions in an ecosystem of connectivity in which all online platforms are inevitably interconnected” (van Dijck, 2014, p. 204).

This indiscriminate sharing of personal data and its use to manipulate information in the online environment constitutes one of the main reasons for the need for the intervention of fact-checkers, especially in times that are decisive for democratic systems, such as electoral processes, as mentioned above. In the next section, we analyse some of the events that determined a definitive transformation in the online information landscape, imposing the need for the use of fact-checking as a means to contain alternative facts.

The role of the 2016 US election, Brexit and the 2018 Brazilian election in definitively changing the disinformation landscape

Since the beginning of the 21st century, social networks have played a predominant role in distributing information. As Allcott and Gentzkow describe, the content that started to circulate “can be shared among users without significant third-party filtering, fact-checking or an editorial evaluation” and, thus, “an individual user, with no background or reputation, can, in some cases, reach as many readers as Fox News, CNN or the New York Times” (2017, p. 211). Social media then became a medium for the indiscriminate dissemination of information. After the 2016 US elections, which pitted Hillary Clinton and Donald Trump head-to-head, the impact that

so-called fake news can have during electoral processes and political and social instability scenarios was perceived concretely. One of the main conclusions reached in the multiple analyses of fake news related to the 2016 US presidential elections candidates is that the most popular fake news was shared with a much more significant reach when compared to the main news from reliable media (Silverman, 2016).

Moreover, it is noteworthy that the main fake stories circulating in the months preceding this election had a clear bias favouring Donald Trump most of the time. As a result, he would be elected over his opponent, Hillary Clinton (Silverman, 2016; Allcott & Gentzkow, 2017). For example, among the fake news that achieved the most significant reach on Facebook, measured through a total number of shares, reactions and comments, were the alleged news that Pope Francis had released a statement showing his support for Trump's candidacy or the various fake news stories that associated Hillary Clinton, through the alleged release of emails, with the sale of arms to Middle Eastern terrorist groups (Silverman, 2016). As Silverman and Singer-Vine (2016) conclude, many of those who encountered the fake news effectively believed them, so there is a clear relationship between their dissemination and the achievement of a specific electoral result, Donald Trump's victory.

Only later, in 2018, the scandal that linked the Cambridge Analytica (CA) company to the 2016 US election broke. It was made public that the company had collected personal information from millions of registered Facebook users without their consent. A team of researchers developed an application that included a personality and social network activity questionnaire whose results were shared with the information technology company (Rehman, 2019).

In the same year that this personal data disclosure scandal came to light, in 2018, Jair Bolsonaro entered the presidential election race in Brazil. Presenting himself as a conservative candidate, he focused much of his campaign on social media. In addition to Facebook, WhatsApp was one of

the communication channels with the most significant influence on the Brazilian electorate (Ituassu et al., 2018). An article in *El País*, published at the end of September 2018, discloses the results of monitoring carried out on the public groups of Brazilian voters on WhatsApp for three weeks. The analysis concludes that there were, at that time, at least 100 groups supporting the candidate, in which there was a predominance of fake news that appeared to be reliable, videos that presented themselves as demystifying the negative information circulating about Bolsonaro, as well as fake support from national and international celebrities for the candidacy of the then federal deputy (Benites, 2018).

Maranhão Filho and colleagues highlight the fake news circulating about the “gay kit”, so pejoratively nicknamed, gender ideology. Two days before the second round of the presidential elections, the Superior Electoral Court (TSE) ordered the president to eliminate videos on Youtube and Facebook in which he associated a book entitled “Aparelho Sexual e Cia”. Bolsonaro argued that both this book and the whole project and discussion of such topics with children “would stimulate not only early sexual interest in children but also paedophilia – which obviously intended to provoke in his audience a feeling of social dread and demonisation of the enemy to be combated: the book falsely linked to the ‘gay kit’” (Maranhão Filho et al., 2019, p. 74). In 2018, Brazil was the third country globally with the most users on Facebook and the sixth on Twitter. In total, more than 110 million users were available to receive unverified information coming from a wide variety of channels (Ituassu et al., 2018). According to Arnaudo (2018), even before the 2018 presidential elections, online electoral propaganda, through bots and algorithms, already played a relevant role in the Brazilian political system.

In addition to the two electoral processes already mentioned, an event that definitively marked the evolution of information distribution and verification – Brexit – also stands out. The United Kingdom’s exit process from the European Union (EU) included a referendum held at the end of June 2016. More than 17 million Britons, i.e. 52% of voters, voted in favour of the country’s exit from countries’ economic and political union.

In the days leading up to the referendum, social media trends – number of posts, shares and reactions – favoured a vote in favour of Brexit, “which embodied a narrative contrary to reputable polls and traditional media outlets” (DiFranzo & Gloria-Garcia, 2017, p. 2). Many of the social media accounts, created with the exclusive purpose of disseminating information convenient to promoting the country’s exit from the EU, disappeared immediately after the English went to the polls in 2016 (Bastos & Mercea, 2019).

Current state and future of fact-checking

According to the latest count of fact-checkers carried out by Duke Reporters’ Lab¹, in December 2021, there were 341 fact-checking projects active worldwide. These fact-checkers are identified in at least 102 countries², i.e. more than half of them in total. The research centre has been providing annual reports since 2014, and since that year, the upward trend in fact-checking activity worldwide has been visible.

In 2019, the peak of this record was reached. In just one year, 72 new fact-checkers were launched. However, in 2020, growth slowed significantly, with only 36 new projects emerging worldwide, the lowest annual record since these censuses have been conducted. The Reporter’s Lab (Stencel & Luther, 2021) presents as a potential justification for this slowdown of the pandemic situation, which, while further accentuating the need for fact-checking activity, has also hindered the development of projects, given the successive isolation situations that have occurred in most countries. In this report, Stencel and Luther (2021) refer to the intensive coverage of the pandemic caused by the new coronavirus. This topic dominated the content of most fact-checking platforms around the world. At least five fact-checkers were already identified who were dedicated to verifying claims related

1. The Duke Reporters’ Lab is a research centre at Duke University’s Sanford School of Public Policy dedicated to analysing the world of fact-checking and other areas of journalism. Available at <https://reporterslab.org/>

2. Reporters’ Lab runs an interactive database of fact-checking websites worldwide. Available at: <https://reporterslab.org/fact-checking/>

to health and medicine. However, as described in the report, “the pandemic transformed virtually every fact-checking operation into a team of health journalists” (Stencel & Luther, 2021, p. 1).

According to Arnold (2020), the automation of the fact-checking process is observed with scepticism by several practitioners in the field: “Many believe that the concept of ‘automated fact-checking’ is intended to automate a process that requires human interpretation – such as weighing the credibility of facts or recognising satire” (p. 7). Many practitioners and researchers understand that verifying a given fact requires intuition and creativity that a purely automated process cannot achieve (Graves, 2018). In this way, and according to Nakov et al. (2021), a double challenge is set to Artificial Intelligence development professionals for this area: “First, to develop practical tools to solve problems that fact-checkers face and second to demonstrate their usefulness for the daily work of these professionals.” (p. 2).

The role of fact-checking in combating disinformation on the Covid-19 pandemic

In Portugal, the misinformation phenomenon about Covid-19 started weeks after confirming the first infected person in the country. The report “Information and Disinformation about the coronavirus in Portugal”, carried out by researchers from ISCTE’s Media Lab project, analyses how, in an initial phase, the Portuguese reacted to the outbreak on the social networks and in search engines.

One of the study’s main conclusions is that, at the beginning of the proliferation of the disease in Portugal, there was “a moment of strong spread of disinformation via WhatsApp, mostly based on audios” (Moreno et al., 2020, p. 40). Voice messages in Portuguese in which the author presented himself as a doctor, nurse, other health professionals, or even a close relative of one of these professionals became viral. The information transmitted was related to the supposed concealment of the actual number of disease cases, reports of crises in hospital units and even the confirmation of the

first fatalities of the disease in Portugal (Moreno et al., 2020). Thus, in the study promoted by ISCTE, it is found that this first disinformative event and its viral dimension can be justified through the “credibility that health professions have in Portuguese society”, since, in this case, “for the ordinary citizen, a message shared on WhatsApp by a health professional assumes the same weight of authority as the one that occurs in a consultation or face-to-face interaction” (Moreno et al., 2020, p. 40).

The absence of a cure through scientifically proven therapeutic methods or a vaccine against Covid-19 has turned social distancing into the most effective weapon against the spread of the disease. On social media, the consumption and production of content were promoted by these periods of social isolation that millions of individuals faced over several weeks (Rodrigues, 2020). At first, at least in Portugal, communities and groups were created, mainly on Facebook, which intended, with good intentions, to share valid information about the pandemic and create spaces to clarify doubts, even counting on the participation of health professionals (Moreno et al., 2020). However, it did not take long before social media became a stage for all and any misinformative and often conspiratorial content about Covid-19.

Fact-checkers as a means of moderating and stabilising information on Covid-19

Luengo and García-Marín (2020) contend that independent fact-checkers have, in the current pandemic context, a mediating role between public authorities and the public “by separating facts from inaccurate information and classifying pieces of information as false, misleading, distorted or decontextualised” (p. 425). Moreover, in a context of widespread mistrust and doubt, such as that experienced at the beginning of the pandemic and which has continued, information verification and evaluation by fact-checkers are ultimately able to delay rumours and viral conspiracies and mitigate their effects, and may even become important ‘symbols of truth’ once they are shared massively, as is easily the case with disinformation (Luengo & García-Marín, 2020).

Case study: Analysis of fact-checks on Covid-19 published by Polígrafo

Polígrafo fact-checker published its first Covid-19-related article on 27 January 2020, more than a month before the first case of the disease was confirmed in Portugal. At this time, the main focuses of the coming torrent of disinformation were the origin of the still unknown virus and the emergence of multiple conspiracy theories that crossed political-economic interests with the development of vaccines and treatments against the outbreak of the new coronavirus that started in the city of Wuhan in central China.

Inspired by pioneer North American fact-checkers, such as Politifact, Polígrafo was, since its creation, mainly focused on producing content related to verifying statements made by politicians/public figures and publications on social networks to the Portuguese political and socioeconomic panorama. Although international and health topics have always been on the radar of Polígrafo, they did not comprise, before February 2020, such a large slice of the newspaper's production as that seen over the last few months.

Next, we analyse the fact-checking articles published on the website of the Polígrafo between September and January 2020, due to its relevance in changing the disinformation landscape with the evolution of the pandemic situation in Portugal and the approval of the first vaccine against Covid-19.

Methods

This study aims to analyse the fact-checks performed by the fact-checking online newspaper Polígrafo, specifically the cases in which fact-checking is directly related to the Covid-19 pandemic. By being anchored on a qualitative-quantitative content analysis methodological approach, the study sought to answer the following research question:

RQ: How did the evolution of the pandemic in Portugal influence the fact-checks on Covid-19?

We then created a database of all articles directly associated with the new coronavirus, which were published on the website of the fact-checking Polígrafo between 1 September 2020, and 31 December 2020. Thus, the absolute number of fact-checks carried out in the first stage in each of the four months referred (N= 463) was counted, regardless of the topic treated. Subsequently, the number of articles that addressed topics unequivocally related to the Covid-19 pandemic was calculated relatively (N= 160), encompassing national and international topics. Next, the multiple origins of the information that set the tone for the verification were considered. The allegations that Polígrafo analysed during this period came from social media, public figures (primarily political figures) and suggestions from readers. Therefore, the data collected and processed shows the predominance of the origin of the facts subject to verification.

The research is also about the ratings assigned to the fact-checks on Covid-19, according to the Polígrafo scale that comprises seven levels. In addition, monthly and total analysis is made of the number and percentage of articles assessed with each classification. These data allow us to observe the evolution of the veracity of the content assessed by Polígrafo in the last four months of 2020.

The diversity of sub-themes that can be highlighted within the universe of the pandemic caused by SARS-CoV-2 and the information circulating on it is vast. Thus, it deserves a concrete analysis that allows establishing some of the disinformative trends. Therefore, based on qualitative content analysis (Bardin, 2008), in this study, seven categories were created from the reading and interpretation of each of the 160 fact-check articles: statistical data on Covid-19 (1), therapeutic methods (2), tests to detect the disease (3), measures to contain the pandemic (4), means of contagion (5), masks (6) and vaccines (7).

The data was collected between 10 and 12 September 2021. The collection was performed manually on the back office platform of the newspaper that acts as an archive of all articles published on the website.

Results

Firstly, the total number of fact-checking articles carried out and published by Polígrafo in September, October, November and December 2020 was analysed. The numbers vary between 114 and 117 fact-checks per month. During these four months, and as it occurred since the end of January 2020, the Polígrafo developed a significant amount of fact-checks to verify information about Covid-19 and other aspects and topics related to the disease the pandemic situation experienced in the country and the world.

Thus, and as shown in Figure 1, in September 2020, 35 fact-checks about the new coronavirus were published on the newspaper's website. There was an upward trend in the two following months, with 40 articles in October and 45 in November 2020. However, there was a slight decrease in December 2020 compared to the previous month – 40 articles related to the pandemic were posted.

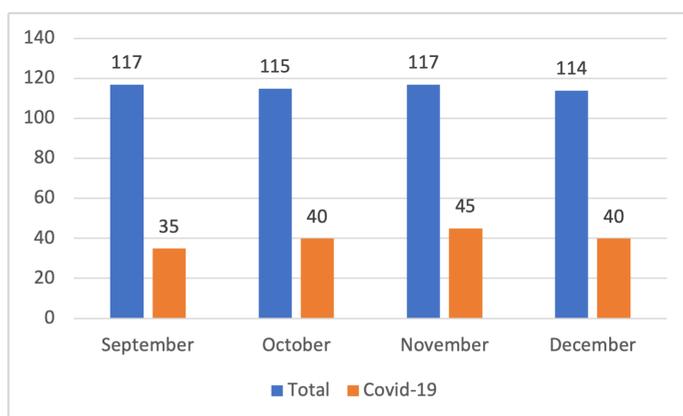


Figure 1 - Total number of fact-checks and relative number of fact-checks on Covid-19 between 1 September and 31 December 2020

In September 2020, 30% of the checks carried out by Polígrafo had SARS-CoV-2 as a theme. In October 2020, of the 117 articles published on the website, 35% belonged to this category, and in November 2020, the

percentage increased to 38%. Finally, in December 2020, 35% of fact-checks were again published on Covid-19. The percentages shown can be visualised in figure 2.

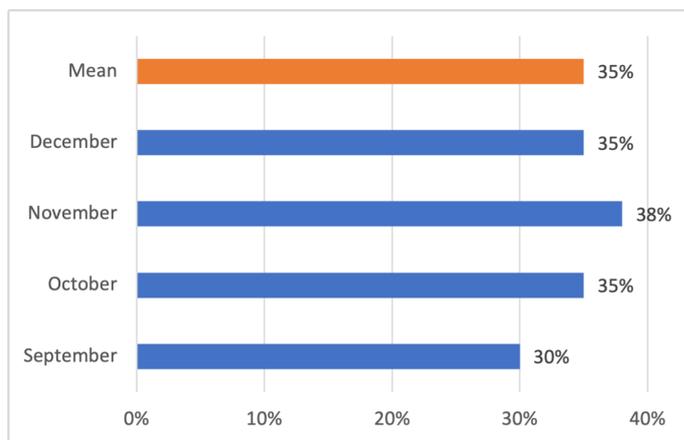


Figure 2 - Percentage of fact-check articles on Covid-19

Also noteworthy is the monthly average of fact-check pieces on Covid-19, which was, during these four months, 40 articles. Also, on average, this type of content accounted for 35% of the total number of fact-checking articles published on the Polígrafo website.

It was possible to distinguish five origins of the information on Covid-19 verified during these four months: Facebook, Twitter, WhatsApp, statements by public figures and, finally, suggestions from readers. Facebook was the space from which much of this content arrived. It is evident, observing figure 3, the increase in the number of articles based on information present in this social network over the four months. If this growth was slight and gradual between September and November 2020, registering an increase from 21 to 28 articles. In December 2020, almost all the checks (93%) were carried out from publications flagged on Facebook, a fact that can be explained by the wave of misinformation about the vaccines.

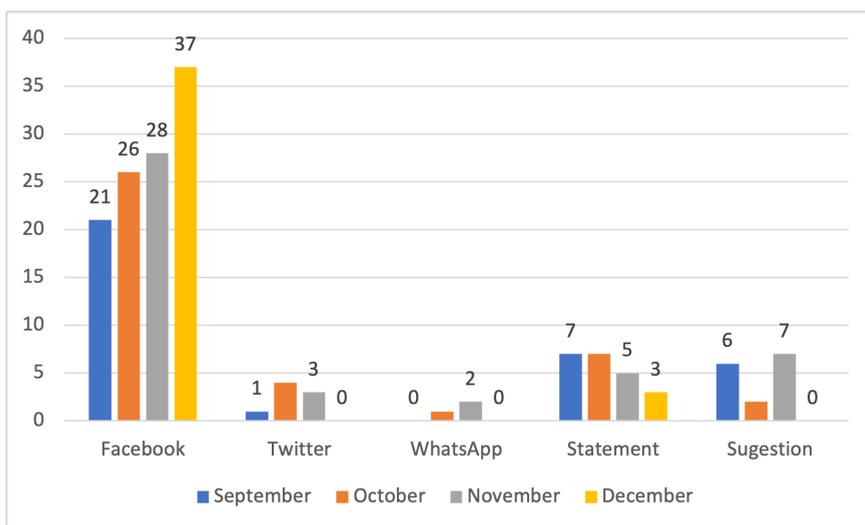


Figure 3 - Number and percentage of fact-checking articles on Covid-19

Both Twitter (8 articles) and WhatsApp (3 articles) were residual origins for the information analysed by Polígrafo. In fact, over the four months, only 11 articles, out of a total of 160, were based on posts on these social platforms.

Regarding the verification of statements by public figures about Covid-19 (22 articles), there was a decreasing trend over the four months, from 20% in September to 8% in December. A contrary trend can be observed between the increase in verifications originating from Facebook and these statements.

In September and November, reader suggestions (15 articles) had a higher number and percentage of articles based on them. In October, two articles were published based on the information disclosed through this type of proposal, which constitutes 5% of the total number of verified facts. However, this source of information was not represented in December, as there were no articles with this source during this period.



Figure 4 - Number and percentage of fact-check ratings on Covid-19 for September, October, November and December 2020

Fig. 4 shows the number and percentage of articles assessed with each classification. In all months analysed, the evaluation with the highest expression was “false”, always recording a percentage equal to or above 40%. The months of October and December stand out in which, coincidentally, the percentage of articles whose analysis was concluded as false reached 65%.

Regarding the attribution of the “true” seal, which is limited to unequivocally correct information, in each month, it never exceeded 30%. Of note was December, when only 5% of the content checks were classified as entirely accurate. Intermediate ratings, which respond to more complex factual issues, where not all information is false or true or, for example, is only incomplete or out of context, showed only slight variations during the period under study.

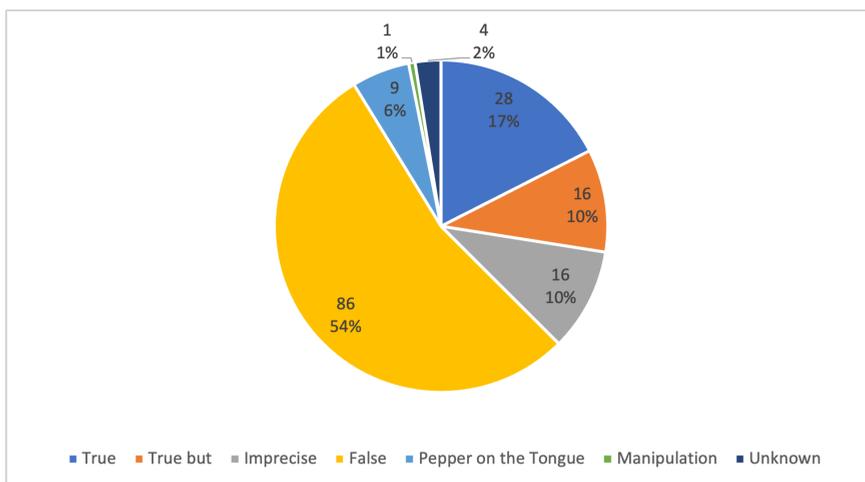


Figure 5 - Number and percentage of ratings assigned to the total fact-checks on Covid-19 published between 1 September and 31 December 2020

Evaluating the data of the fact-check ratings under analysis (N=160) in an absolute way, in the total of the four months, which can be seen in figure 5, 86 fact-checks were rated with the “false” stamp, which corresponds to 54% of the total content produced on Covid-19 in this period, or more than half. The ‘pepper on the tongue’ rating, which represents the last degree of falsehood on the Polígrafo scale and which is, of necessity, used sparingly as it only applies “when the evaluated information is outrageously false”, was applied in only nine articles during these four months. A total of 28 articles, which corresponds to 17% of the total, were developed from information considered true. As for the remaining assessments, 16 ‘true, but...’ and 16 ‘inaccurate’, i.e. each of them represents 10% of the total number of fact-checking articles on the new coronavirus. Only five articles were assessed with the labels ‘decontextualised’ and ‘manipulated’, which may be since the introduction of these levels in the assessment scale is recent. Also, to the level of specificity of the assessment since this only applies to multimedia content in the case of the ‘manipulated’ classification.

As previously mentioned, seven thematic categories were developed that include content related to statistical data on Covid-19 (1), therapeutic methods (2), disease detection tests (3), pandemic containment measures (4), means of contagion (5), masks and vaccines (7). The “other” category groups all the articles that, due to their specificity, do not fit in any of the categories mentioned.

Topic	September		October		November		December	
Statistical data	4	11%	4	10%	6	13%	4	10%
Therapeutic methods	2	6%	3	8%	5	11%	3	8%
Detection Tests	5	14%	3	8%	2	4%	3	8%
Pandemic containment measures	4	11%	10	25%	13	29%	3	8%
Means of contagion	10	29%	1	3%	4	9%	1	3%
Masks and vaccines	4	11%	8	20%	5	11%	1	3%
Others	1	3%	3	8%	5	11%	23	58%

Table 1 - Distribution of fact-checks on Covid-19 carried out between 1 September and 31 December 2020 by topic

Table 1 shows the number and percentage of articles distributed by the respective topics in each month that are part of this analysis. In September, the Covid-19 category most worked on by the Polígrafo was the means of contagion, which filled 29% of the production. In the two following months, October and November, the percentage of checks carried out that focused on pandemic containment measures stand out, at 25% and 29%, respectively. In December, a very significant percentage of fact-checks had vaccines against the new coronavirus as their main topic, around 58%, more than half of the total for the month. The production of content related to this last topic evolved increasingly throughout the four months. In September, only 3% of check pieces had vaccines as their theme, a percentage that increased to 8% in October and 11% in November, having escalated substantially in the last month of the year.

As for the topics of statistical analysis and therapeutic methods, there is a consistency in the number and percentage of checks in which they figure

as the main content to be reviewed. Over the four months, the number of pieces based on these topics ranged from two to six. The categories ‘contagion’ and ‘masks’ showed a downward trend, albeit with slight variations between months, both registering the percentage of 3% of the total articles in December.

	September		October		November		December		Total	
Experts	11	31%	15	38%	22	49%	23	58%	71	44%
Official Information	20	57%	15	38%	15	33%	15	38%	65	41%
Scientific Articles	20	57%	28	70%	28	62%	28	70%	104	65%

Table 2 – Percentage and number of the sources articles in which the different verification methods were used in September, October, November and December 2020

Consultation of articles and other pieces of scientific research was, during the four months analysed, the most frequently used method in the checks. In total, it was recorded in 104 articles, about 65% of the total checks on Covid-19 (N=160). That is, it is understood that, in most cases. However, fact-checks could rely on other sources. They were based on scientific evidence present in studies or other types of research pieces that solidify the checking process to the journalist and works as documentary evidence that can be consulted and analysed by the recipients of the articles. Regarding the contact and subsequent citation of health experts and, more specifically, Covid-19, an upward evolution is shown in Table 2. In September, the number and percentage of articles in which the Polígrafo interviewed experts to obtain clarifications was 31%. In October, it was 38%, and in November, it increased to 49%. Finally, in December, the use of this type of source reached 58%. Regarding the last month of 2020, it is important to relate these contacts with experts and the predominant verification of the topic of vaccines observed in the previous section.

The consultation of official information and/or the contact with health authorities, which includes, for example, the analysis of statistics on the pandemic evolution, requests for clarification from the Ministry of Health or the search for directives and guidelines from national and global health

entities, was carried out in 41% of the articles, i.e. 65 of them. Therefore, the evolution of the use of this source can be characterised as fluctuating, with ups and downs throughout the four months.

Conclusions

The results achieved with this case study present multiple conclusions and verify several misinformative trends. Firstly, regarding the overall analysis of the fact-checks on Covid-19 published by Polígrafo between September and December 2020, there is a progressive increase in the number of verification articles dedicated to Covid-19, reaching a peak in November and registering a slight decrease in December. This growing trend of disinformation analysis by Polígrafo is directly proportional to the evolution of the pandemic situation in the country, namely the increase in the number of newly detected cases of the disease, as can be seen by consulting the data made available by the Directorate General of Health. On 16 November 2020, a total of 8,371 new daily cases were registered, this being the peak of the four months under analysis in this study. In November, the highest number of fact-checks on the new coronavirus carried out and published by the Polígrafo was recorded: 45 fact-checks which, in that month, constituted 38% of the newspaper's total production. We conclude then that there is a direct relationship between the worsening pandemic situation and the verification of information by the Polígrafo.

Regarding the source, a significant amount of information to be analysed and evaluated comes from Facebook, registering the prevalence of this origin of fake news in all months. Firstly, this is justified by Polígrafo's partnership with Facebook since 2019. Furthermore, the easy access to contents that were previously flagged as suspicious of containing false or misleading information by the social network means that, automatically, the fact-checking platform has privileged access to them. In addition, replicating false information by the various social networks occurs regularly. Therefore, in the Polígrafo article, the verification of the content on Facebook is more clearly marked and does not exclude the possibility of information

being circulated in other social networks. Thus, justifying the low number of analysed information that arrived from Twitter and WhatsApp.

In the case of suggestions from readers, it is not possible to establish a pattern of the evolution of this source of information. September and November were the months in which there was a more significant number and percentage of articles based on the proposals of those who read the *Polígrafo*. These suggestions for the newspaper's production depend on their relevance and editorial options. The fact that there is no record of checks made based on suggestions in December may be related to the fact that much of the content suggested to the *Polígrafo* by users is already previously identified by Facebook, as observed in the data presented. This registered a decrease over the four months regarding verifying statements from public figures about Covid-19. As already mentioned, there is a decrease proportional to the increase in verifications originating from Facebook, i.e., one justification for this can only be the editorial choices made, prioritising the analysis of misinformation on social networks, another is the moderation of public communication about the pandemic situation in the country.

On the other hand, the analysis of the evaluations attributed to the fact-checking pieces included in this study also allows us to reach several conclusions. The first is that the 'false' classification was the most used every month. It is natural that, since the genesis of *Polígrafo* and fact-checkers, in general, is the intention of "keeping lies out of the public space", the majority of articles necessarily deal with verifications that end up being concluded as false or with a certain level of falsehood. If we look to the totality of the articles under analysis, only 17% were classified as entirely true and 10% as being 'true, but'.

In summary, it is concluded that there is an evident influence of the Covid-19 pandemic evolution in Portugal on the fact-checking articles carried out by *Polígrafo* in the last four months of 2020. The number of fact-checks, emphasising those classified as false, about the new coronavirus increased as the number of people infected by the disease in the country increased.

Furthermore, the arrival and advance of the winter season accentuated the risk of hospital overcrowding and the need to implement measures to contain the disease. Thus, the anger of many towards the rules that were imposed boosted the sharing of misinformation about the virus on social media. Moreover, there was a sudden change in the Covid-19 disinformation landscape in December. Most of the fact-checking pieces analysed having Covid-19 vaccines as the object of verification, a definite shift towards what would become one of the main focuses of disinformation about the disease.

In a mediatic ecosystem where the risk of sharing and receiving false information is high, concrete solutions are needed to extinguish or mitigate the effects of this constant exposure. In the long term, an actual effort on media literacy and its integration into educational programmes, offering the youngest the tools to enable them to be their fact-checkers, seems to be the most definitive solution and the least immediate. Therefore, the role of media and fact-checking platforms is considered crucial, despite their limitations, in the action of verifying disinformation, not only from a primary point of view, through the analysis and integration of information in a rating scale that assesses its level of veracity, but also by making available to the public the verification methods and tools used by fact-checkers which, in many cases, are accessible to ordinary citizens. In other words, it is necessary to demonstrate that just as there is tremendous ease in sharing and accessing information, it is also relatively simple for each individual to act as their verifier, basing their shares on the news from reliable media outlets, using free image verification programs or consulting scientific articles.

From the creation of the first projects in the 1990s and early 21st century until now, fact-checking journals have developed in number and size and the ability to verify information on and off social media. The Covid-19 pandemic has, however, posed new challenges to this activity. The emergence of a new disease, about which little was known and little scientific literature existed, meant that the information circulating was inaccurate and diffuse from very early on. Fact-checkers had to adapt quickly and practically to a new disinformative reality based on health and science, and areas not

always worked on with the most remarkable regularity. Using infectious disease specialists, pulmonologists, and public health doctors as sources and consulting information issued by governments and health authorities have become a daily routine for fact-checkers worldwide. Through the empirical study carried out based on the verification articles conducted by Polígrafo in the last four months of 2020, it is possible to observe the disinformative flows and how they related to the evolution of the pandemic outbreak in Portugal and the world. It was found that the vast majority of the verified information was classified with a certain degree of falsity and that, at least in the case of Polígrafo, the primary medium where the analysed disinformation circulated was Facebook.

From the data analysed, we contend that it will be essential to invest in the development of automated tools and programs that assist the fact-checkers activity, always bearing in mind the limitations that these processes may have. Nevertheless, more urgent is the need to integrate into schools curricula subjects on media literacy and, for example, considering the specific case, health literacy, in both formal and informal learning.

References

- Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, 31 (2): 211-36. <https://doi.org/10.1257/jep.31.2.211>
- Amaral, I. (2016). *Redes sociais na internet: sociabilidades emergentes*. Covilhã: Editora LabCom.
- Amaral, I., & Santos, S. J. (2019). Algoritmos e redes sociais: a propagação de fake news na era da pós-verdade. In J. Figueira & S. C. Santos (Eds.), *As fake news e a nova ordem (des)informativa na era da pós-verdade: manipulação, polarização, filter bubbles* (pp. 63-85). Coimbra: Imprensa da Universidade de Coimbra.
- Amazeen, M. A. (2020). Journalistic interventions: The structural factors affecting the global emergence of fact-checking. *Journalism*, 21(1), 95–111. <https://doi.org/10.1177/1464884917730217>

- Arnaudo, D. (2017). Computational Propaganda in Brazil: Social Bots during Elections. *Computational Propaganda Research Project*, 8, 1–39.
- Arnold, P. (2020). *The challenges of online fact checking*. Published by Full Fact, November 2020. Published under the Creative Commons Attribution-ShareAlike 4.0 International License. Retrieved from <https://fullfact.org/media/uploads/coof-2020.pdf>
- Bakir, V., & McStay, A. (2017). Fake News and The Economy of Emotions: Problems, causes, solutions. *Digital Journalism*, 6(2), 154–175. <https://doi.org/10.1080/21670811.2017.1345645>
- Bastos, M. T., & Mercea, D. (2019). *The Brexit Botnet and User-Generated Hyperpartisan News*. *Social Science Compute Review*, 37(1), 38–54. <https://doi.org/10.1177%2F0894439317734157>
- Benites, A. (2018) *A máquina de 'fake news' nos grupos a favor de Bolsonaro no WhatsApp*. El País Brasil. Retrieved from https://brasil.elpais.com/brasil/2018/09/26/politica/1537997311_859341.html
- Bloyd-Peshkin, S. & Sivek, S. (2017). Magazines find there's little time to fact-check online. *Columbia Journalism Review*. Retrieved from <https://www.cjr.org/b-roll/magazine-fact-checking-online.php>
- Cassedy, J. H. (1964). Muckraking and Medicine: Samuel Hopkins Adams. *American Quarterly*, 16(1), 85–99. <https://doi.org/10.2307/2710829>
- DiFranzo, D., & Gloria-Garcia, K. (2017). *Filter bubbles and fake news*. *XRDS: Crossroads, The ACM Magazine for Students*, 23(3), 32–35. <https://doi.org/10.1145/3055153>
- Dobbs, M. (2012). The rise of political fact-checking. *Media Policy Initiative: Research Paper*. Retrieved from <https://www.issuelab.org/resources/15318/15318.pdf>
- Graves, L. (2013). *Deciding What's True: Fact-Checking Journalism and the New Ecology of news*. Columbia University. Retrieved from <https://academiccommons.columbia.edu/doi/10.7916/D8XG9Z7C>

- Graves, L. (2018). *Understanding the Promise and Limits of Automated Fact-Checking. Technical report*. Reuters Institute, University of Oxford. Retrieved from Understanding the Promise and Limits of Automated Fact-Checking | Reuters Institute for the Study of Journalism (ox.ac.uk)
- Graves, L., & Cherubini, F. (2016). *The Rise of Fact-Checking Sites in Europe* (Reuters Institute Digital News Report). Reuters Institute for the Study of Journalism. Retrieved from <https://ora.ox.ac.uk/objects/uuid:d55ef650-e351-4526-b942-6c9e00129ad7>
- Ireton, C., & Posetti, J. (2018). *Journalism, “fake news” & disinformation: handbook for journalism education and training*. UNESCO. Retrieved from <https://digitallibrary.un.org/record/1641987>
- Ituassu, A., Lifschitz, S., Capone, L., Vaz, M. B., & Mannheimer, V. (2018). Compartilhamento de mídia e preferência eleitoral no Twitter: uma análise de opinião pública durante as eleições de 2014 no Brasil. *Palavra Chave*, 21(3), 860-884. <https://doi.org/10.5294/pacla.2018.21.3.9>
- Luengo, M., & García-Marín, D. (2020). The performance of truth: politicians, fact-checking journalism, and the struggle to tackle COVID-19 misinformation. *American Journal of Cultural Sociology* 8, 405-427. <https://doi.org/10.1057/s41290-020-00115-w>
- Manjoo, F., & Porter, R. (2008). *True enough*. Blackstone Audiobooks.
- Maranhão Filho, E. M. de A., Coelho, F. M. F., & Dias, T. B. (2019). Fake news acima de tudo, fake news acima de todos”: Bolsonaro e o “kit gay”, “ideologia de gênero” e fim da “família tradicional. *Correlatio*, 17(2), 65-90. <https://doi.org/10.15603/1677-2644/correlatio.v17n2p65-90>
- Moreno, J., Cardoso, G., Narciso, I., Palma, N., Pinto-Martinho, A., Crespo, M. et al. (2020). *Informação e desinformação sobre o Coronavírus nas notícias e nas redes sociais em Portugal*. Retrieved from <https://ciencia.iscte-iul.pt/publications/informacao-e-desinformacao-sobre-o-coronavirus-nas-noticias-e-nas-redes-sociais-em-portugal/70893>

- Nakov, P., Corney, D., Hasanain, M., Alam, F., Elsayed, T., Barrón-Cedeño, A., Papotti, P., Shaar, S., & Da San Martino, G. (2021). *Automated Fact-Checking for Assisting Human Fact-Checkers*, 1, 4551–4558. Retrieved from [2103.07769] Automated Fact-Checking for Assisting Human Fact-Checkers (arxiv.org)
- Rehman, I. (2019). Facebook-Cambridge Analytica data harvesting: What you need to know. *Library Philosophy and Practice*, 1-11.
- Rodrigues, D. (2020). *Pandemia de desinformação: uma análise dos tipos de notícias falsas sobre o novo coronavírus e como agências de fact-checking tratam a desinformação no Brasil e em Portugal*. Retrieved from <https://ciencia.iscte-iul.pt/publications/informacao-e-desinformacao-sobre-o-coronavirus-em-portugal--whatsapp-facebook-e-pesquisas/70894>
- Silverman, C. (2016). *This Analysis Shows how Fake Election News Stories Outperformed Real News on Facebook*. BuzzFeed News. Retrieved from <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook>
- Silverman, C. & Singer-Vine, J. (2016). *Most Americans Who See Fake News Believe It, New Survey Says*. BuzzFeed News. Retrieved from <https://www.buzzfeednews.com/article/craigsilverman/fake-news-survey>
- Singer, J. B. (2014). User-generated visibility: Secondary gatekeeping in a shared media space. *New Media & Society*, 16(1), 55-73. <https://doi.org/10.1177%2F1461444813477833>
- Spinelli, E. M., & Santos, J. D. A. (2018). Jornalismo na era da pós-verdade: fact-checking como ferramenta de combate às fake news. *Revista Observatório*, 4(3), 759-782. <https://doi.org/10.20873/uft.2447-4266.2018v4n3p759>
- Stencel M. & Luther J. (2021). *Annual census finds nearly 300 fact-checking projects around the World*. Retrieved from <https://reporterslab.org/annual-census-finds-nearly-300-fact-checking-projects-around-the-world/>
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance and Society*, 12(2), 197-208. <https://doi.org/10.24908/ss.v12i2.4776>

FRAMING COVID-19: HOW FACT-CHECKING CIRCULATE ON THE FACEBOOK FAR-RIGHT

Raquel Recuero, Taiane Volcan, Felipe Soares,
Otávio Vinhas & Luiz Ricardo Huttner

Introduction

Since the beginning of the Covid-19 pandemic, the world has faced a surge of disinformation about the virus on social media. This issue has not only captured the attention of governments and institutions in different countries but also brought to light the debate over the responsibility of social media companies for allowing health disinformation to circulate on their platforms. Some of these companies, such as Facebook, for example, took action by partnering with fact-checking agencies and creating ways for its users to flag disinformation. Boosting fact-checking circulation and using it to debunk falsehood were common actions from platforms to try to mitigate the disinformation problem.

Fact-checking connects the investigation of the accuracy of the content to the debunking of problematic posts. The number of specialized agencies and outlets exploded during the last years, particularly because of political disinformation (Graves, 2016). Even so, the extent of its effectiveness is often discussed. Some studies have shown, for example, that fact-checking is no match for disinformation regarding the velocity of spread and scale (Vinhas & Bastos, 2022). Others showed that fact-checking content is often not able to reach beyond partisan identification and thus, circulates more

on groups that agree with it (Shao et al., 2018). Partisanship and political discourse seem to be an important part of the health disinformation phenomenon, as we will argue in this text.

Covid-19 disinformation was potentialized by its alignment with political discourse in social media platforms, particularly, by far-right groups (Rogers & Niederer, 2020). In many countries, far-right populists' governments and politicians acted to spread disinformation by negating the gravity of the pandemic and publicly distrusting vaccines. In these cases, their discourses frequently aligned conspiracy theories about the pandemic and their political agendas (many of which associated the pandemic with a leftist conspiracy). These discursive connections associated Covid-19 mitigation strategies and vaccines with populism and political ideologies, which fueled negationists' postures and vaccination hesitancy (Calvillo et al., 2021, Recuero & Soares, 2020; Soares et al. 2021).

Given this context, we present a case study of how fact-checking links were shared by groups and pages that also shared disinformation links in Brazil. Brazil currently has a far-right President, Jair Bolsonaro, who was also involved in sharing disinformation about the pandemic and the virus (Soares et al., 2021). Bolsonaro also has a strong presence on social media platforms, with thousands of supporters who were also involved with sharing disinformation about the pandemic (Ricard & Medeiros, 2020). These groups are also associated with political extremism and populism because of their views on several subjects, particularly, on the Covid-19 pandemic (Medeiros & Silva, 2021). These characteristics provide an important situation for our analysis, as the Brazilian government's negationist posture increased the political polarization in the country and framed the Covid-19 pandemic as a political issue and not a public health one (Recuero & Soares, 2020). This alignment allowed an increased circulation and legitimation of Covid-19 disinformation (Soares et al., 2021). Therefore, in contexts like this, it is key to understand how effective social media platforms' strategies used to mitigate disinformation can be. Our research questions, thus, are as follows:

RQ 1: How does fact-checking links about the Covid-19 pandemic circulate among far-right groups and pages compared to other groups that also shared disinformation on Facebook? Do they contribute to debunking disinformation?

RQ 2: Are these fact-checking links framed by posts in any way? If so, how?

We hypothesize that, due to political extremism, fact-checking that circulates on these groups is framed to support disinformation, instead of challenging it. To test this hypothesis, we gathered data from Facebook using CrowdTangle. We crawled disinformation and fact-checking links in Portuguese about Covid-19 which were shared by Facebook's public groups and pages during 2020 and further selected those made by pages/groups that shared both. This original dataset of links was provided by Poynter/IFCN. Our final dataset was composed of 860 posts with 411 unique fact-checking links.

Disinformation, Fact-Checking and Political Discourse on Social Media

Social media platforms have a key role in the spread of disinformation. Their affordances, such as the capacity to help content to spread further and farther in the social network; the easy replicability (Boyd, 2010); the possibility to find like-minded people who will be more willing to share types of content; and the availability of artificial strategies such as botnets and click farms (Bastos & Mercea, 2019) provide the perfect environment to spread all types of content, including problematic ones.

Social media platforms often rely on algorithms to select content to show their users. These algorithms, combined with users' actions to select content may help create an effect called an "echo chamber" (Cinelli et al., 2021). Echo chambers are structures of conversation on social media that mostly circulate homogeneous content. That means, people, select to share only content they agree with, which tends to be reinforced by homophily. This collective action of filtering content together with platform algorithms may

create groups where people become more exposed to similar content that confirm their ideological views rather than challenges them (Westerwick, Johnson & Knobloch-Westerwick, 2017; Workman, 2018). This phenomenon has been associated with the increase of political polarization and extremism, particularly, the far-right (Rogers & Niederer, 2020; Pariser, 2011). In these cases, extremists tend to create very clustered groups where content is filtered to agree with the groups' political views (Barberá et al., 2015). The more politically biased content that circulates, the more extreme the group becomes.

Because of this context, social media platforms have also been appropriated as means of propaganda by political extremists from the far-right (Rogers & Niederer, 2020), which has also boosted the disinformation spread (Tucker et al. 2018). Disinformation, in these cases, is used to reinforce political ideas. Since the Covid-19 pandemic also happened amidst political discussions and polarization, political disinformation was often also connected to health disinformation (Recuero & Soares, 2020). Particularly, the far-right discourse may have fueled disinformation about the Covid-19 pandemic in different countries (Calvillo et al., 2020; Allcott et al. 2020). Far-right leaders and politicians have also used social media platforms to amplify their ideas, often through disinformation (Kallil et al., 2021; Galhardi et al. 2020) and as an information guerrilla weapon (Soares et al. 2021, Ricard & Medeiros, 2020). Social media platforms have affordances that help these discourses to spread e being legitimated more quickly and broadly. Moreover, this intersection between political disinformation and health disinformation was often marked by populists' discourses (Recuero & Soares, 2022), which is why it is important to further investigate these connections.

Scholarship on populism is vast. However, many authors see a new wave of "far-right" populism that has emerged among traditional democracies, particularly in western countries, since the 70s (Jagers and Walgrave, 2007). This new populism is often based on a "deeply conservative discourse" that constitutes its core identity, that focuses on (1) anti-globalism and nationalism, often relying on theories that claim conspiracies from the

“global elites” (Guimarães & Silva, 2021). Another important characteristic is the (2) anti-foe construction, which is built upon the otherness process, meaning, separating “us” (the good, virtuous people) and the others (the corrupted, the foes). In these cases, in-group values are superior and more virtuous than other values (outsiders) (Staszak, 2009). This relation of “us” x “others” creates a perception of being part of a group that shares the same characteristics (homogeneity/purity). Discourses that operate upon this strategy usually legitimate other processes such as exclusion, xenophobia, racism, etc (which Wodak, 2015, claims, is a “politics of fear”). Beeze (2020) points out that this process of “otherness” also creates a common enemy, where populist discourse can create a sense of urgency, crisis, and denunciation to justify the actions that are taken. Another important characteristic of populist discourses is (3) the idea that the leaders are representants of “the people” to fight against the “corrupt elites” and the “rotten” establishment (Gil de Zúñiga, Michalska & Römmele, 2020; Roudjin, 2019). Part of populist discourse, particularly in this scenario, is often also identified with authoritarianism, which means, discourses where the leader is strong and claims that his decisions are legitimated by “the people” (Mestres, 2021). These characteristics often align populist discourse with simplistic, yet powerful ideas that can increase people’s hesitancy to collaborate with public health measures, particularly in situations that most people never experienced, such as the Covid-19 pandemic.

Populist strategies produce complicity with some groups while discrediting the ones who disagree. Barrera et al (2022) explain that far-right populists’ leaders use “alternative facts” as counter-narrative strategies. These narratives, which often are built upon disinformation, are very persuasive. Thus, populists’ governments have also been connected to the spread of disinformation about Covid-19 (Stecula & Pickup, 2021), and the far-right political views have also been connected to a higher tendency to consume disinformation content (Baptista et al., 2021). In this scenario, populism seems to be deeply connected to the spread of disinformation, which is something we intend to explore in this research.

This is also the current political context of Brazil. The Brazilian president, Jair Bolsonaro, is considered by many political scholars a far-right populist representative (Mendonça & Caetano, 2020; Watmough, 2021). Bolsonaro's discourse was also frequently aligned with populists' arguments, such as authoritarianism, otherization, nationalism, distrust of the elites, and the idea that his government represents "the people" (Mendonça & Caetano, 2020; Watmough, 2021). Like many far-right leaders, Bolsonaro and his supporters frequently used social media to legitimize disinformation about the gravity of the pandemic, the lethality of the Covid-19 virus, and the vaccines (Soares et al, 2021; Kallil et al., 2021; Galhardi et al. 2020). He was a strong supporter of the idea that the pandemic mitigation strategies could not interfere with the economy, and that the media and health experts were creating panic. He also never used masks in public and defended the usage of ivermectin and hydroxychloroquine as the cure for Covid-19 and the solution for the pandemic (Ricard & Medeiros, 2020; Alcantara & Ferreira, 2020). Brazil, is thus, one of the cases where social media users played an important part in spreading and legitimizing disinformation about the Covid-19 pandemic through a polarized political context and, especially, through populism.

Finally, we need to examine fact-checking's potential to mitigate disinformation in such scenarios. Fact-checking, as we explained, is currently posited as one of the most popular strategies to fight disinformation. Initially viewed as a tool to hold politicians to account by enforcing journalistic truth-seeking practices (Graves, 2016), fact-checking has lately extended its scope to include verifying and correcting viral disinformation on social media platforms (Graves & Mantzarlis, 2020). This led fact-checking to grow increasingly popular after the US 2016 election, reaching 342 in 102 different countries according to Duke's Reporter's Lab (Stencel & Luther, 2021). As the Covid-19 pandemic frenzied, more and more initiatives have not only been devoted efforts to verify potential false statements on political claims but also verifying health claims that can potentially cause harm to large populations. Thus, fact-checking has established a key role in combating

Covid-19 disinformation in many countries by integrating journalistic procedures, truth-seeking ideals, international institutions, and their worldwide collaborative network.

Social media platforms have relied strongly on fact-checking as the main approach to challenge this type of content. Facebook, for example, has a specific program to support third-party fact-checking in its platform¹. The purported intent is to boost the circulation of reliable and verified content while diminishing the relevance of posts classified as misleading or false by the fact-checkers. These programs are not devoid of dissent, as distinct technical, institutional, and epistemological issues take place over narratives around “facts” (Vinhas & Bastos, 2022). Furthermore, Cotter et al. (2022) argue that, by implementing these programs, social media platforms consolidate the idea that what is true should be ultimately determined by their users, downsizing the role of journalists, experts, and authoritative actors in promoting reliable information. Either way, authors have claimed that fact-checking programs are legitimate ways for platforms to enforce content moderation measures, which could overall help mitigate disinformation (Gillespie, 2020).

Despite showing some promising results globally (Porter & Wood, 2022), studies have demonstrated that fact-checking often may not be as effective, particularly among politically radicalized groups (Barrera et al., 2020). In addition, findings by Carey et al. (2022) show that fact-checking’s positive effects against misconceptions are often undermined by contexts in which corrections are ephemeral in comparison to the constant flow of falsehoods. Authors like Shin & Thompson (2017) claim that fact-checking circulates with a political bias, which means, circulates more within groups that already agree with their content. This may implicate that politically radicalized groups may filter or frame fact-checking to align with their ideologies, similarly to what Shao et al. (2018) argue.

1. <https://www.facebook.com/journalismproject/programs/third-party-fact-checking>

Methods

For this work, our main objective is to discuss how fact-checking links circulated among groups that also shared disinformation, with a focus on far-right groups. We want to investigate how fact-checking links circulate among these groups to see if they can challenge health disinformation, especially on extremized political groups, such as the far-right. This is particularly important since they became the main strategy used by social media platforms to reduce the circulation of problematic content, as we argued in the previous section. We chose far-right groups to understand if this content can reach politically extremist groups if it can break through echo chambers, as other studies have suggested they can't.

For this study, we chose to focus on two main points: (1) how fact-checking links circulate among groups that also shared disinformation on Facebook, particularly, the far-right political groups; and (2) if and how these links are framed by posts. We hypothesize that fact-checking that circulates in these polarized groups may be framed to increase disinformation, compared to other groups.

To collect data for this discussion, we relied on a dataset provided by Poynter/IFCN that comprises links to both disinformations about Covid-19 and the correspondent fact-checking from associated groups all over the world during 2020. We used CrowdTangle to collect posts that contained fact-checking and disinformation links in Portuguese from public groups/pages on Facebook. With these posts, we selected the ones from the groups/pages that shared both (disinformation and fact-checking). Through these steps, we were able to collect 860 posts that contained fact-checking links and that were posted in groups/pages that also shared disinformation. Based on this sample, we examined the fact-checking posts (N=860) and unique fact-checking links (N=411, some links were shared several times) that circulated on these groups/pages. These posts were shared by 420 pages/groups in this dataset. From these 860 posts, 270 had an explanatory text framing the link.

To analyze the data, we worked with a three-step mixed-methods combination. First, to classify posts from the far-right and discuss disinformation framing, we used Content Analysis (Krippendorff, 2013). Three independent coders visited every single post and classified (1) how the fact-checking link was framed (if it was framed as disinformation) and (2) if the group/page was aligned with far-right views. The far-right political classification was based on names that included politicians, political parties, political ideologies, and/or references to the far-right and conservative ideologies present in Brazil, as we explained in the previous section. Most of them, in this dataset, were connected to president Jair Bolsonaro or his supporters. To discuss if the fact-checking was framed as disinformation, coders observed how the link was posted (text, other links, etc.). In these cases, coders observed the association of fact-checking to leftist conspiracies and anti-globalism, far-right populist discursive characteristics, or the framing of the fact-checking as misleading content by the text in the post that contained the link. These posts were read and examined by all coders.

Cohen's Kappa (Cohen, 1960) was 0.71 for framing (95.5% interpair agreement) and 0.86 for political alignment (93.7% interpair agreement). As the entire dataset (860 posts) was coded by three independent coders, the final classification was reached based on the agreement between at least two coders.

For the next step, we used qualitative analysis. We used discourse analysis (Fairclough, 2003) to understand if and how they articulated this far-right populism to legitimate health disinformation. For this, we looked for the characteristics of populists' discourses as explained in the previous section (anti-foe/otherness construction, nationalism or globalism, and the arguments against fact-checkers as corrupt elites). This part of the analysis was done over with 74 posts that included text to frame fact-checking as disinformation.

Further on, on the third step, we used Social Network Analysis (Wasserman & Faust, 1994) to create a bipartite graph (nodes were pages/groups and fact-checking links) to understand connections between these pages and the links they shared. In this step, we wanted to observe patterns of sharing links between far-right pages/groups versus others. We used indegree to find out the most shared links by far-right pages and outdegree to find out which pages were the most active in sharing fact-checking links. Also, we examined clusters of far-right pages around fact-checking links. These links were further collected, and we analyzed their titles, as they are the main thing that circulates on Facebook posts. We wanted to understand the patterns of sharing fact-checking within these groups. Because we found some different patterns among far-right groups (compared to other groups that also shared disinformation), we decided to investigate further. So, we also examined the five most shared links among these far-right pages (which were shared by at least five different pages/groups) to understand their discourse and how it possibly was aligned to the groups' ideology and populist characteristics. This analysis was also qualitatively and was done through discourse analysis (Fairclough, 2003).

Results and Discussion

The results of our analysis are presented in this section. We organized our findings into main ones, as we will explain.

1. Fact-checking does circulate among far-right political groups, however, it is often framed as disinformation

To understand if and how posts were framed, we compared fact-checking links shared in other pages in the same dataset to the ones shared in far-right ones. We are trying to compare, here, pages/groups that share disinformation and fact-checking to far-right groups. Table 1 summarizes these results based on coders' classification.

	Number of Posts	Posts framed as disinformation	Unique links
Far-right pages/groups	295 (34.3%)	74 (25.1%)	32
Other pages/groups	565 (65.6%)	10 (1.8%)	379
Total	860	84 (9.8%)	411

Table 1: Data from pages/groups and links.

For this dataset, we observed that the majority of fact-checking posts on all pages weren't framed as false. The large majority circulated either with confirmation framing on their posts (such as, "Check this" or "Rumors") or without comments. Our results show that 776 of the examined posts (90.2%) fact-checking was not framed as disinformation by the post. Another 84 posts (9.8%), thus, were framed as disinformation by the pages/groups that shared them. These posts were largely published on far-right pages/groups.

Most posts within far-right groups/pages did not frame fact-checking as disinformation nor included a text to deny fact-checking content. Nevertheless, compared to other groups/pages, far-right accounts were almost 14 times more likely to frame fact-checking as disinformation. In the next section, we discuss some of the main strategies used by those actors to frame fact-checking content. While we had other politically themed pages (N=225, 53.6% of the dataset), fact-checking framed as disinformation was much more common among the far-right than other pages/groups of the political spectrum. This framing was done either by a comment subverting the information or by circulating fact-checking aligned with the views of the group, often through comments.

2. Implicit framing: Far-right pages and groups also tend to cluster around fact-checking links that agree with their ideological views without the need to frame them explicitly

Many posts within far-right groups/pages did not contain any text along with the fact-checking links. Nevertheless, given the general contexts of Covid-related disinformation in Brazil, some of the fact-checking content

might have been used to “prove their points” (that the virus was not dangerous, for example) without any explicit framing. Based on this perception, we decided to investigate further. To better understand the circulation of fact-checking links among far-right groups, we examined the structure links posted on far-right groups through social network analysis. We focused on two types of nodes: pages/groups and links. Far-right pages were colored red (other pages were colored blue) and links were colored gray.

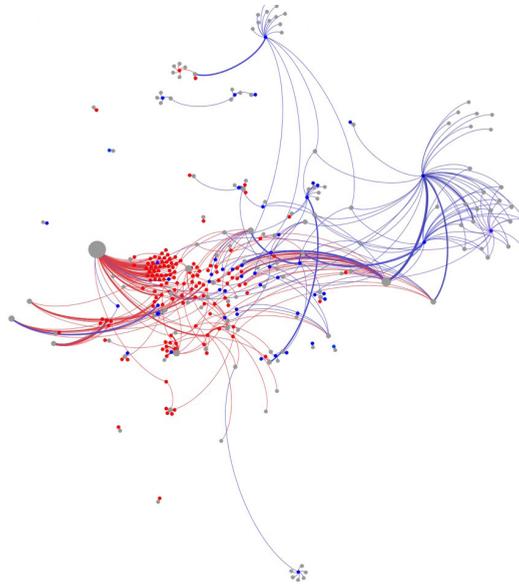


Figure 1: Red nodes are far-right groups/pages, blue nodes are others and gray nodes are links. Size of node is given by the number of times it shared the links.

Figure 1 shows this network of pages/groups and links. Node size is defined by outdegree (the higher the outdegree, the bigger is the node). Outdegree is connected to the number of times each page shares each link. We see, in this picture, two patterns: (1) Big blue nodes that share several different fact-checking links and (2) a small cluster of red nodes that share the same fact-checking links. We can observe that red nodes cluster together, which means that these far-right nodes tend to share the same fact-checking links, clustering around fewer links than other that were shared by other pages.

We also see that most pages (from the blue nodes) share several different fact-checking links (and thus, these nodes are bigger).

This structure suggests that while fact-checking circulates in far-right groups, it appears that certain links circulate more, and others are ignored. So why is this happening? The following image highlights the most shared links by using the indegree to adjust the size of nodes (Figure 2). The indegree is the number of connections each node receives. It is expected that more influential nodes would receive higher visibility from the network. Different from the blue cluster, the gray nodes from the red cluster are much bigger, thus more shared by several far-right nodes. This structure also suggests that far-right pages and groups have preferences to post certain fact-checking posts (several pages post the same link), whereas in the rest of the network, the structure is the opposite: few pages post several different links.

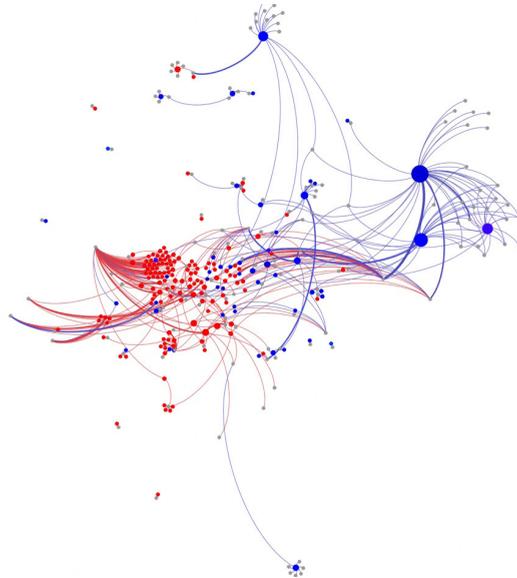


Figure 2: Gray nodes are links, blue nodes and red nodes (far-right) are pages/groups. Size of node is given by the number of times it was shared.

To investigate what was happening, we decided to focus on these far-right most-shared links, which have several nodes clustered around them. We further examined the five most popular fact-checking links that had at least five shares by these groups. The table below (Table 2) shows the headline of the most shared fact-checking within far-right groups.

	Story	Link	Indegree
1	“Picture that shows a large number of coffins is falsely attributed to Covid-19 deaths - it is in fact from Italy in 2013”	http://piaui.folha.uol.com.br/lupa/2020/03/18/coronavirus-caixoes-italia	63
2	“Empty coffin picture and video are old and have no connection to Amazon’s Covid-19 funeral”	http://piaui.folha.uol.com.br/lupa/2020/04/29/verificamos-foto-caixao-vazio-enterros-covid-19-amazonas	26
3	“It is false that the Brazilian media silenced after Lula said ‘I’m happy that nature created this Covid-19 monster’”	https://piaui.folha.uol.com.br/lupa/2020/05/20/verificamos-imprensa-se-calou-lula-ainda-bem-monstro-coronavirus/	15
4	“It is false that elderly people who disrespect Covid-19 isolation will have their retirement suspended”	http://piaui.folha.uol.com.br/lupa/2020/03/20/verificamos-idosos-desrespeitarem-isolamento-covid-19-aposentadoria-suspensa	11
5	“The video of Dr. Drauzio Varella minimizing Covid-19 isolation is old”	https://piaui.folha.uol.com.br/lupa/2020/03/22/drauzio-salles-coronavirus/	7

Table 2: Most shared stories in the far-right network and number of times.

It is interesting to notice that, taken out of context, these fact-checking stories help dismiss the gravity of the pandemic. Story 1 implies that false images are used to create panic about the Covid-19 pandemic. While this story was shared mostly without a frame, in a few groups we found a confirming framing, such as “See how evil people are. They are spreading terrorism. The picture is from 2013 and has been shared as from 2020.” Another text was “Leftist fake!”. In this case, even though the story is shared as a real fact-checking link, it contributes to the discourse of the far-right that argues that the pandemic wasn’t that serious, as Bolsonaro has frequently argued (Soares et al., 2021).

Story 2 is about a mistake when someone was buried, and it is used to fight the criticism of politicians on the state of Amazonas as the deaths sharply increased and happened. In this case, very few links used textual framing, also, with confirming framing. Examples such as “This is a circus and will only end when the people that share this kind of content are punished! This didn’t happen in Amazon!”. Another one is “Picture is from 2017 in Sao Paolo, not Amazon.”. These texts are also not questioning the veracity of the fact-checking, but they imply that pictures like these are used to negatively frame how the Amazon state government (which is from a party that openly supported Bolsonaro, PSC – Social Christian Party) was in chaos, dealing with the sharp increase of deaths.

Story 3 attacks a leftist leader (Lula) with a misleading title, as it appears to fact-check only the first part of the sentence – the media silenced. In this case, most of these links circulated also without framing. The ones that circulate with framing often used criticism of Lula. Lula is an important leftist leader, who is also an ex-president of Brazil. This is also another link that, while confirming the fact-checking as truthful, underlined the fact that the left was “happy” about the pandemic.

Story 4 is used to dismiss “terrorist” media about the lockdown measures (people won’t be punished for breaking the lockdown). Also, while this content is true, it reinforces the idea that people could break the lockdown. Story 5 is used to put in a bad light a doctor that protested for more Covid-19 mitigation measures and often criticized the federal government about the lack of action during the pandemic.

These stories were shared on pages and groups that have interpretative far-right contexts, mostly guided by Bolsonaro and his supporters (Soares et al, 2021; Kallil et al., 2021; Galhardi et al. 2020). In these cases, these titles also provoke more distrust in political elites and specialists, as well as the left and the media, which are also ways for far-right populists’ discourses to reinforce themselves (Roudjin, 2019). Thus, it is also very likely that these links weren’t framed because they already contributed to confirm, and not

challenge, the far-right populist discourse present in these groups. As we explained, these discourses play on characteristics of populism and the legitimization by the far-right ideology.

This data suggests that not only far-right groups shared the same fact-checking links as they also shared links whose titles would agree with their political views. We describe this movement as “implicit framing”. It suggests that fact-checking links are selected based on how their stories agree or not with the discourse of the group which brings the interpretative context for the discussion. In these cases, the fact-checking wasn’t debunked, which means, fact-checking is biased towards the context where it is shared, similarly to what Vinhas & Bastos (2021) argue.

3. Both implicit and explicit frames use populist discourse strategies

As we explained, far-right groups would largely use populists’ discursive strategies, particularly, otherness. Most fact-checking posts published by these pages would be shared with explicit framing that would oppose fact-checkers and the page audience (74 posts). Some examples are: “We need to unmask this bunch of liars!” or “Face ‘good’ says this is a lie. Do you believe?”. In both these phrases, there is an opposition between “us” (the good people) and them (the bad people), the virtuous and the bad ones (Staszak, 2009; Guimarães & Silva, 2021). The disbelief in the traditional media and platforms is also connected to this framing, as elites that should be questioned. Nationalism was found in three of these posts (“This is a national scandal! We can’t allow this in our country”). In these cases, the fact-checking would be connected to leftists’ conspiracies and attempts to destabilize the economy of the country and the “good, virtuous” government of Jair Bolsonaro.

One example from these populist strategies was framing fact-checking agencies as outlets supported by the left. One case was the phrase “Look at the sickle lie! The true which ‘honest journalists’ of Brazil are denying!” which framed a fact-checking link that claimed some disinformation shared

by the far-right was false. In this case, the post highlights the supposed existence of a political alignment behind the fact-checking agencies here called “sickle”, a reference to the agency’s supposed alignment with communism. Another framing strategy was to associate fact-checking with media manipulation. One example was the post “Do not trust everything that the Lupa agency² claims to be a lie or truth because it makes mistakes and ends up manipulating opinions. Research the truth yourself.” In these cases, we observe the ideas of conspiracies from corrupt elites as ways to discredit fact-checking.

When we analyzed the clusters of fact-checking links shared by the far-right pages. We found that these links had characteristics of populist discourses as well, as the links shared agreed with the ideological views of these groups/pages. These movements, of implicit and explicit framing, suggest, thus, that fact-checking that circulates on far-right groups tend to be framed as disinformation. Although these are not problematic contents per se, they are a way to increase the general discourse built by far-right disinformation on health. The usage of populist strategies also helps by creating distrust of elites and general health information, even the ones shared by fact-checking agencies.

These posts would share and legitimize the idea that the pandemic wasn’t serious, and people were being manipulated by corrupt leaders and elites to take action that would harm themselves (for example, using masks could provoke suffocation). These strategies would frame fact-checking as something misleading, used to manipulate a product from these corrupt elites (Gil de Zúñiga, Michalska & Römmele, 2020; Roudjin, 2019).

These processes can be associated with echo chambers (Cinelli et al., 2021). Because of the polarized political context, these pages and groups can be filtering certain types of fact-checking links, strengthen their political narrative about the pandemic. There are also some “anti-foe” or “othering” alignments in the interpretation of these titles. Many of these fact-checking

2. Lupa is a fact-checking agency well known in Brazil. <https://piaui.folha.uol.com.br/lupa/>

content could be read as criticism to the left and institutions and elites such as traditional media and scientists/leftist. In this case, the context for the far-right is usually to align with Bolsonaro's views about the pandemic, which could explain these clusters of pages sharing fact-checking stories that corroborate with their views. Fact-checking is, thus, also subjected to polarized effects from group actions (Vinhas & Bastos, 2021) and our data suggest that, on politically radicalized Facebook groups and pages, it may not be effective (Barrera et al., 2020).

This data supports findings from other research that show that the far-right political affiliation may be strongly connected to receiving and sharing health disinformation during the Covid-19 pandemic.

Conclusion and implications

Our results suggest that far-right groups are much more likely to resist fact-checking and align these contents according to their ideology than other groups/pages. We found that, while fact-checking does circulate on politically radicalized groups for the far-right, they are often implicitly or explicitly framed as disinformation.

Populist arguments also play an important part in this framing. Comments that discredit the press and especially fact-checking agencies, alleging a supposed "leftist conspiracy", or "otherness process" have a strategic effect to maintain the relevance and circulation of disinformation and reduce the effects of fact-checking. Thus, discrediting fact-checking is an important way to reassure the populist discourses that are frequent in these groups. Results also suggest that in other not so extremist political groups, fact-checking may have better effects as it circulates without framings that distort their content.

These results suggest that platforms need to go further than boosting fact-checking to challenge disinformation. Different strategies are needed to deconstruct the different frames used to share fact-checking in politically

extremized groups. This study has several limitations such as the size of the sample and the language. However, we believe that it shows a contribution to the studies of the disinformation ecosystem, discursive strategies, and far-right disinformation.

References

- Alcantara, J., & Ferreira, R. (2020). A infodemia da “gripezinha”: uma análise sobre desinformação e coronavírus no Brasil. *Chasqui. Revista Latinoamericana de Comunicación*, 1(145), 137-162. <http://doi.org/10.16921/chasqui.v1i145.4315>.
- Allcott, H.; Boxell, L.; Conway, J.; Gentozow, M.; Thaler, M.; Yang, D. Y. (2020) Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic (Working Paper No. 26946; Working Paper Series). *National Bureau of Economic Research*. <http://doi.org/10.3386/w26946>.
- Baptista, J.P.; Correia, E.; Gradim, A. & Piñero-Naval, V. (2021) The Influence of Political Ideology on Fake News Belief: The Portuguese Case. *Publications 2021*, 9, 23. <http://doi.org/10.3390/publications9020023>
- Bastos, M. T., & Mercea, D. (2019). The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review*, 37(1), 38–54. <http://doi.org/10.1177/0894439317734157>
- Barberá, P.; Jost, J. T; Nagler, J.; Tucker, J. A.; Bonneau, R. (2015) Tweeting from Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, v.26, n. 10, 1531–1542, 2015. <http://doi.org/10.1177/0956797615594620>.
- Barrera, O.; Guriev, S.; Henry, E & Zhuravskaya, E. (2020) Facts, alternative facts, and fact checking in times of post-truth politics, *Journal of Public Economics*, Volume 182, 104123. <http://doi.org/10.1016/j.jpubeco.2019.104123>.
- Breeze, R. (2020) Exploring populist styles of political discourse in Twitter. *World Englishes*. 39: 550– 567. <http://doi.org/10.1111/weng.12496>

- Boyd, d. (2010) "Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications." In *Networked Self: Identity, Community, and Culture on Social Network Sites* (ed. Zizi Papacharissi), 2010, pp.39-58.
- Calvillo, D. P., Ross, B. J., Garcia, R. J. B., Smelter, T. J., & Rutchick, A. M. (2020). Political Ideology Predicts Perceptions of the Threat of COVID-19 (and Susceptibility to Fake News About It). *Social Psychological and Personality Science*, 11(8), 1119–1128. <http://doi.org/10.1177/1948550620940539>
- Calvillo, D.; Garcia, R.; Bertrand, K; Meyers, T. (2021) Personality factors and self-reported political news consumption predict susceptibility to political fake news, *Personality and Individual Differences*, V. 174, 110666. <http://doi.org/10.1016/j.paid.2021.110666>.
- Carey, J.M., Guess, A.M., Loewen, P.J. et al. (2022) The ephemeral effects of fact-checks on COVID-19 misperceptions in the United States, Great Britain and Canada. *Nat Hum Behav* 6, 236–243. <http://doi.org/10.1038/s41562-021-01278-3>
- Cinelli, M. & M., Gianmarco; Galeazzi, A., Walter, Q. & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*. 118. e2023301118. <http://doi.org/10.1073/pnas.2023301118>.
- Cohen, J. A. (1960). Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1): 37-46.
- Cotter, K., Decook, J. R. & Kanthawala, S. (2022) Fact-Checking the Crisis: COVID-19, Infodemics, and the Platformization of Truth. *Social Media + Society*. <http://doi.org/10.1177/20563051211069048>
- Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. Routledge.
- Galhardi, C. P. et al. (2020) Fato ou Fake? Uma análise da desinformação frente à pandemia da Covid-19 no Brasil. *Ciência & Saúde Coletiva*. v. 25, suppl 2, pp. 4201-4210. <http://doi.org/10.1590/1413-812320202510.2.28922020>.

- Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Big Data & Society*. <http://doi.org/10.1177/2053951720943234>
- Gil de Zúñiga, H., Koc Michalska, K., & Römmele, A. (2020). Populism in the era of Twitter: How social media contextualized new insights into an old phenomenon. *New Media & Society*, 22(4), 585–594. <https://doi.org/10.1177/1461444819893978>
- Graves, L. (2016). *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*. Columbia University Press.
- Graves, L. & A. Mantzarlis (2020). Amid Political Spin and Online Misinformation, Fact Checking Adapts. *The Political Quarterly* 91: 585-591. <http://doi.org/10.1111/1467-923X.12896>
- Guimaraes, F. & Oliveira, E & Silva, I. Far-right populism and foreign policy identity: Jair Bolsonaro's ultra-conservatism and the new politics of alignment. *International Affairs* 97: 2, 2021
- Jagers, J. & Walgrave, S. Populism as political communication style. *European Journal of Political Research*, v. 46, n. 3, p. 319-345, 2007.
- Kalil, I.; Silveira, S., Pinheiro, W.; Kalil, A.; Pereira, J.V.; Azarias, W. & Amparo, A. (2021) Politics of fear in Brazil: Far-right conspiracy theories on COVID-19, Global Discourse. *Global Discourse: An interdisciplinary journal of current affairs*. Volume 11, Number 3, 1–17, <http://doi.org/10.1332/204378921X16193452552605>
- Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). Sage Publications, Inc.
- Mendonça, R. F., & Caetano R. D. (2021). Populism as Parody: The Visual Self-Presentation of Jair Bolsonaro on Instagram. *The International Journal of Press/Politics*, 26(1), 210–235. <http://doi.org/10.1177/1940161220970118>
- Mestres, L. (2021) The Rise of Right-Wing Populism, Authoritarianism & Fascism, *Socialism and Democracy*, 35:1, 142-165. <http://doi.org/10.1080/08854300.2021.1949778>
- Pariser, E. *The Filter Bubble*. The Penguin Press, 2011.

- Porter, E. & Wood, T. (2021). "The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom." *Proc Natl Acad Sci U S A* 118(37).
- Recuero, R., & Soares, F. (2022). #VACHINA: How Politicians Help to Spread Disinformation About COVID-19 Vaccines. *Journal of Digital Social Research*, 4(1), 73-97. <http://doi.org/10.33621/jdsr.v4i1.112>
- Ricard, J., & Medeiros, J. (2020). Using misinformation as a political weapon: COVID-19 and Bolsonaro in Brazil. *Harvard Kennedy School (HKS) Misinformation Review*. <http://doi.org/10.37016/mr-2020-013>
- Rogers, R. & Niederer, S. (2020) *The Politics of Social Media Manipulation*. Amsterdam University Press, 292p.
- Shao, C., Hui, P., Wang, L., Jiang, X., Flaminni, A., Menczer, F., & Campaglia, G. L. (2018). Anatomy of an online misinformation network. *PLoS One* 13(4): e0196087. <http://doi.org/10.1371/journal.pone.0196087>
- Shin, J. & Thorson, K. (2017). Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media. *Journal of Communication* 67(2), 233-255. <http://doi.org/10.1111/jcom.12284>
- Soares, F. B., Recuero, R., Volcan, T., Fagundes, G., & Sodré, G. (2021). Research note: Bolsonaro's firehose: How Covid-19 disinformation on WhatsApp was used to fight a government political crisis in Brazil. *Harvard Kennedy School (HKS) Misinformation Review*. <http://doi.org/10.37016/mr-2020-54>
- Stecula, D. A., & Pickup, M. (2021). How populism and conservative media fuel conspiracy beliefs about COVID-19 and what it means for COVID-19 behaviors. *Research & Politics*, 8(1). <http://doi.org/10.1177/2053168021993979>.
- Staszak, J.F. (2009) Other/Otherness, Editor(s): Rob Kitchin, Nigel Thrift, *International Encyclopedia of Human Geography*, Elsevier, Pages 43-47. <http://doi.org/10.1016/B978-008044910-4.00980-9>.
- Stencel, M. & Luther, J. (2021). Fact-checking census shows slower growth, *Duke Reporter's Lab*. <https://reporterslab.org/fact-checking-census-shows-slower-growth/>

- Tucker, J. A.; Guess, A.; Barberá, P.; Vaccari C.; Siegel, A.; Sanovitch, S.; Stukal, D. & Nyhan, B. Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. Hewlett Foundation, 2018. <http://doi.org/10.2139/ssrn.3144139>.
- Vinhas, O. & Bastos, M. (2022) Fact-Checking Misinformation: Eight Notes on Consensus Reality, *Journalism Studies*. <http://doi.org/10.1080/1461670X.2022.2031259>
- Wasserman, S. & Faust, K. Social Network Analysis. *Methods and Applications*. Cambridge University Press, 1994
- Watmough, S. P (2021). 'Jair Bolsonaro: Far-Right Firebrand and Cheerleader for Dictatorship', *Populism Studies*. <https://www.populismstudies.org/jair-bolsonaro-far-right-firebrand-and-cheerleader-for-dictatorship>
- Westerick, A., Johnson, B. K., & Knobloch-Westerick, S. (2017). Confirmation biases in selective exposure to political online information: Source bias versus content bias. *Communication Monographs*, 84, 343-364. <http://doi.org/10.1080/03637751.2016.1272761>
- Wodak, R. (2015). *The Politics of Fear: What Right-Wing Populist Discourses Mean* (1st ed.). SAGE Publications Ltd.
- Workman, M. (2018) An empirical study of social media exchanges about a controversial topic: Confirmation bias and participant characteristics. *The Journal of Social Media in Society*. V.7(1):381–400.

Authors

Alyssa M. Brumis

Doctoral Candidate in the interdisciplinary Language, Literacy, and Culture Program at the University of Maryland Baltimore County (UMBC). In addition to serving as a Graduate Assistant, Alyssa teaches media and communication courses at Howard Community College, the College of Southern Maryland, and UMBC. Alyssa is currently completing her dissertation, “Misinformation & Misbeliefs: Reinforcing False Beliefs Related to COVID-19 in the Age of ‘Alternative Facts,’” a qualitative study exploring the different factors that influence belief change relating to the coronavirus pandemic. Alyssa’s research interests span across disciplines such as communication and media studies, sociology, and social psychology, as she utilizes an interdisciplinary approach to study the effects and consequences of misinformation across mass media. Alyssa’s book review on Tero Karppi’s work, *Disconnect: Facebook’s Affective Bonds*, was published in *Rhizomes: Cultural Studies for Emerging Knowledge* and her article, “The Right to Privacy in a Digital Age: Reinterpreting the Concept of Personal Privacy,” can be found in *Inquires Journal*.
Email: alyssa13@umbc.edu

Allison McDowell-Smith

Deputy Director, American Counterterrorism Targeting and Resilience Institute (ACTRI). Allison is a CT and mis- and disinformation researcher, the Associate Dean of the Graduate School of Liberal Arts & Sciences, Director of the Graduate Counterterrorism Program, Chair of the Undergraduate Criminal Justice Programs, and Associate Professor of Criminal Justice and Counterterrorism at Nichols College.
Email: Allison.McDowell-Smith@americanctri.org

Alessandra Massa

Research Assistant in the Department of Communication and Social Research at Sapienza University of Rome. Her research interests focus mainly on the intersection between international political communication and online platforms. She authored several books and articles on political narratives in digital environments, international journalism, disinformation, and migrations. Her most recent work is (with G. Anzera) *Media digitali e relazioni internazionali. Tecnologia, potere e conflitti nell'era delle piattaforme online (Digital Media and International Relations. Technology, Power, and Conflict in the Online Platforms Age)* (Guerini & Associati).

Email: alessandra.massa@uniroma1.it

Amanda Garry

Senior Research Fellow, American Counterterrorism Targeting and Resilience Institute (ACTRI), and Cyber Defense Technologist within Raytheon Missiles and Defense.

Email: amanda.garry@americanctri.org

Ardian Shajkovci

Director, American Counterterrorism Targeting and Resilience Institute (ACTRI). Ardian is a lecturer and a CT and mis-and disinformation researcher with field work experience in Syria, Ukraine, Somalia, Iraq, Jordan, Kenya, Central Asia, Western Europe, and the Balkans. He has written for, and his work has been quoted by, *The New York Times*, *The Hill*, *Homeland Security Today*, *New York Post*, *The Washington Times*, *Euronews*, *The Daily Beast*, *Le Figaro*, *Washington Examiner*, *AFP*, *Daily Caller*, *Fox News*, and others.

Email: ardianshajkovci@americanctri.org

Asli Altinbay

Business Development Manager at Storyzy. Asli works as a Business international projects and partnerships and leads Storyzy's international projects and partnerships.

Email: asli.altinbay@storyzy.com

Giuseppe Anzera

Full Professor in the Department of Communication and Social Research at Sapienza University of Rome. His research interests concern various issues, including international relations, media diplomacy, nation branding, geopolitics, and conflict analysis. He authored several publications on radicalization and deradicalization techniques, international migrations, online platforms, and security issues. Among his works: *Flussi di armamenti e politica internazionale (Flows of Arms and International Politics)* (Guerini & Associati), (with A. Massa) *Media diplomacy e narrazioni strategiche. Autorappresentazione dello Stato e attuazione della politica estera in rete (Media Diplomacy and Strategic Narratives. Self-representation of the State and the Implementation of Foreign Policy on the Internet)* (Bonanno).

Email: giuseppe.anzera@uniroma1.it

Felipe Soares

Researcher at MIDIARS Research Lab. He holds a PhD in Communication and Information (UFRGS - Federal University of Rio Grande do Sul, Brazil). His research interests include social network analysis, political communication, and disinformation.

Email: fbonowsoares@gmail.com

Inês Amaral

Inês Amaral is an Associate Professor at the Faculty of Arts and Humanities of the University of Coimbra. She serves as coordinator of the Communication Section of the DFCI, and director of the 1st cycle of studies in Journalism

and Communication. PhD in Communication Sciences (Interactive Media) from the University of Minho, where she is an integrated researcher at the Communication and Society Research Centre. She is also an associate researcher at the Center for Social Studies at the University of Coimbra.
Email: ines.amaral@uc.pt

Ján Hacek

Associate Professor at the Department of Journalism of the Faculty of Arts at Comenius University in Bratislava (Slovakia) and currently Head of Department of Journalism. He focuses mainly on Data Driven Journalism, Open sources for Journalists and Media content analysis. He leads lectures on data journalism, media literacy, media research and theory. He is a member of multinational research network EU Kids Online and a member of Expert Group: The Media Pluralism Monitor – Centre for Media Pluralism and Media Freedom. He collaborated on the international research project: People between the lines: Building skills for quality migration coverage. Hacek is an author of publication Open sources and Data for Journalists in Slovakia and he has also published scientific paper about journalist Ján Kuciak's last article – published after his assassination: Open sources in the last article by Ján Kuciak. Hacek is a Journalism prize board member in Slovakia and collaborator with non-profit organizations (People in Need Slovakia, Open Society Foundation, European dialogue).
Email: jan.hacek@uniba.sk

Lasha Kavtaradze

MA, PhD student at Kristiania University College and the University of Bergen. Kavtaradze holds a master's degree in Digital Media and Society from Uppsala University. He worked as a journalist, communications manager, fact-checker, and media critic. Kavtaradze is studying the use of emerging technologies for information verification, specifically the use of AI and automatization for fact-checking, and its impact on journalism.

He is the author of the handbook for journalists to cover violent extremism and radicalisation published by the Georgian Charter of Journalistic Ethics and the Georgian Center for Strategy and Development (GCSD). He has published academic and opinion articles about media polarization and the influence of political radicalization on the Georgian media landscape on various platforms.

Email: Lasha.Kavtaradze@kristiania.no

Lucia Virostková

Assistant Professor at the Department of Journalism of the Faculty of Arts at Comenius University in Bratislava. She has joined the department after 20 years of journalist practise: she worked for Euobserver.com in Brussels and reported on EU affairs for several Slovak media. After returning to Slovakia in 2009, she worked as Radio news reporter, live political debates moderator, editor and eventually joined the investigative reporting programme at Slovak public broadcaster. As a scholar, she specializes on investigative journalism, foreign affairs and EU reporting. She is a member of research project investigating the celebritization of Marian Kocner, the businessman facing criminal court proceeding on charges of ordering the murder of journalist Jan Kuciak.

Email: lucia.virostkova@uniba.sk

Lucas Graves

PhD, Associate Professor in the School of Journalism and Mass Communication at the University of Wisconsin–Madison. He has been at the forefront of research on the global fact-checking movement and is the author of *Deciding What's True: The Rise of Political Fact-Checking in American Journalism* (2016, Columbia University Press). His work has appeared in the *New York Times*, the *Columbia Journalism Review*, the *Nieman Journalism Lab*, *Wired*, and many scholarly journals. From 2017

to 2019 he was Senior Research Fellow and then acting Research Director at the Reuters Institute for the Study of Journalism at Oxford University.

Email: lucas.graves@wisc.edu

Luísa Torre

PhD student in Sciences Communication at University of Beira Interior with a thesis related to fake news, journalism and social media. She is graduated from the Universidade Federal do Espírito Santo (UFES) in Social Communication with a focus on Journalism. To get her bachelor degree, she researched the semiotics of crime journalism in the local media. She worked as a news reporter for 10 years in Vitória, writing news articles about health, education, culture, lifestyle and economics, using a data-driven, multimedia approach. She was also a columnist and an editor during her time at A Gazeta newspaper. She also contributed with Agência Pública. During these years, Luísa won several local and national prizes for her news coverage on human rights, health and economic issues. In 2018, Luísa moved to Portugal, where she took a Master degree at University of Porto, in the city of Porto. Her research aimed to understand how the fake news phenomenon, and news distribution through social media, affected journalism, more specifically, the discourse of journalism.

Email: luisagtorre@gmail.com

Luiz Ricardo Huttner

PhD student in Communication at Federal University of Rio Grande do Sul (UFRGS), Brazil and a researcher for MIDIARS Research Lab. His research interests include studies in journalism, social network analysis and disinformation.

Email: luizricardohuttner@gmail.com

Mar García-Gordillo

Associate Professor at University of Seville. Bachelor in Journalism at the University of Seville (1994) and PhD in Communication from the same university (1998). She worked in several press offices and the international area of the newspaper *Diario de Sevilla*. Since 1998, Prof. García-Gordillo teaches at the University of Seville in different disciplines, being specialized in foreign affairs news and content analysis. She held several positions as Associate Dean or Director of the Secretariat for Traineeships. Prof. García-Gordillo is currently Director of Communication of the University of Seville, assuming a leading role in the starting up of the Ulysseus European University, which is coordinated by the own University of Seville. She is also professor of the Master in Institutional Communication and Politics and has supervised six PhD dissertations. She has published in journals such as *Profesional de la Información*, *Revista Latina de Comunicación Social* or *Communication and Society*, among others. Her research interests focus on quality journalism and foreign policy news, as well as gender studies and ethnic minorities. Member of the research group Communication & Social Sciences (SEJ-619) and the Department of Journalism II (University of Seville).

Email: marggordillo@us.es

Marina Ferreira

Degree in Law from the University of Coimbra and a postgraduate degree in Social Communication Law from the same institution. During her studies journalism has been present through the university's newspaper. In 2021 she completed her master degree in Journalism and Communication from the University of Coimbra and did an internship at Poligrafo, a portuguese fact-checking newspaper. This is where she is currently workin full time, separating the lie from the truth.

Email: marinaff2195@gmail.com

Matteo Gregori

Research Fellow, American Counterterrorism Targeting and Resilience Institute (ACTRI). Matteo Gregori is an OSINT analysis with expertise in all-source analysis, threat analyses, social media monitoring, and computer programming. He is a graduate of St. John's University with an M.A in Homeland Security and Criminal Justice. He is currently a Project Coordinator at NextGen5.0, having also worked for Horizon Intelligence, Talking about Terrorism, and International Institute for Counter Terrorism. Email: matteo.gregori@americanctri.org / matteo.gregori212@gmail.com

Otávio Vinhas

Ph.D. student in Information and Communication Studies at University College Dublin (UCD). Researcher at the MIDIARS Research Lab. Email: otavio.vinhas@gmail.com

Ramon Ruti

Co-founder & CTO Storyzy. Ramón has a computer engineering background and has more than 15 years of experience in leading, directing, and managing international projects. Email: ramon.ruti@storyzy.com

Raquel Recuero

Professor and researcher for Federal University of Pelotas and Federal University of Rio Grande do Sul. She is the leader of MIDIARS Research Lab (Media, Discourse and Social Networks) and a researcher for the National Research Council in Brazil (CNPq). Recuero has several published books and papers on social media and social networks. Her research currently focuses on disinformation and online discourse and digital methods. E-mail: raquelrecuero@ufpel.edu.br

Oscar Westlund

PhD, Professor at Oslo Metropolitan University, where he co-leads the OsloMet Digital Journalism Research Group. Westlund also has a tenured position at the University of Gothenburg. He has been the Editor-in-Chief of Digital Journalism since 2018, and serves on the editorial boards for more than a dozen journals. His award-winning research includes specialization in digital journalism, fact-checking, platforms, media management, news consumption, and mobile media. Relating to misinformation and journalism, Westlund chairs the advisory board for the EU funded NORDIS project, and works with nationally funded research projects in Norway and Sweden. He has previously worked as an academic expert for public inquiries for Sweden and the EU commission. Among his most recent co-authored books are *What is Digital Journalism Studies?* (with Steen Steensen) and *Critical Incidents in Journalism* (with Edson Tandoc Jr., Joy Jenkins and Ryan Thomas).

Email: oscarw@oslomet.no

Rebekah Larsen

PhD, Postdoctoral Fellow at Oslo Metropolitan University and a Research Associate at Cambridge University's Centre of Governance and Human Rights. She holds a PhD in Sociology and an MPhil in Technology Policy from Cambridge University. Her previous work has focused on the intersection of digital rights and the media, about which she has published academic research and consulted for the European Commission. Previously, she also helped found an academic startup at Cambridge University (*The Whistle*) and worked with a digital rights NGO in the USA. Her current research focuses on alternative media ecosystems in rural and other marginalized communities, as well as reflexive methodological innovation in media studies.

Email: beahlarsen@gmail.com

Rubén Rivas-de-Roca

PhD in Communication from the interuniversity program of the universities of Cadiz, Huelva, Malaga, and Seville. Bachelor in Journalism (2015, National End of Degree Award) and Master's Degree in European Studies (2016, Graduation Prize for Exceptional Achievement) both from University of Seville. He also holds a Master's Degree in Communication, Culture, Society, and Politics from National University of Distance Education-UNED (2019). He was a visiting scholar at Universität Leipzig (Germany), Cardiff University (United Kingdom) and Universidade da Beira Interior (Portugal), having professional experience at the European Commission and Fundación Galicia Europa in Brussels (Belgium). Besides that, Rivas-de-Roca has been granted twice by the German Academic Exchange Service (DAAD). He held a PhD candidate position by Universidad de Sevilla (VI PPIT-US), teaching on Political and Economic Journalism. His PhD project focused on local digital media across different European countries, analyzing their quality and practices. In general, his research addresses the role of journalistic quality and local media, as well as political communication and the EU, in relation to the communication policies of the EU and their impact on the media. Member of the research group Communication & Social Sciences (SEJ-619) and the Department of Journalism II (University of Seville).

Email: rrivasderoca@us.es

Steen Steensen

PhD, Professor at Oslo Metropolitan University, where he leads the PhD program in journalism and media studies. He has previously been the Head of the Department of Journalism and Media Studies at Oslo Metropolitan University (2016-2020), and he holds a visiting professorship at Bournemouth university, UK (2021-2024). His award-winning research specializes in digital journalism, epistemologies of journalism, human-interest journalism, fact-checking, and disinformation, among other topics. Steensen has been the Editor-in-chief of the Norwegian journal of media

studies and associate editor of *Journalism Practice*, and he serves on the editorial board of several journals. He is the PI of the international research project *Source Criticism and Mediated Disinformation (SCAM)*, funded by the Research Council of Norway for the period 2020-2024. His most recent books are *Journalistikkens filosofi [The Philosophy of Journalism]* (2021, co-authored with Harald Hornmoen) and *What is Digital Journalism Studies?* (2020, co-authored with Oscar Westlund).

Email: steen.steensen@oslomet.no

Taiane Volcan

Postdoctoral researcher at MIDIARS Research Lab. She has a PhD in Languages from the Federal University of Pelotas (UFPEL). She works with social media analysis, discourse, politics, humor and disinformation on social media platforms.

Email: taianevolcan@gmail.com

João Carlos Correia

Associate Professor at the Department of Communication, Philosophy and Politics at the University of Beira Interior, Covilhã, Portugal. Email: jcorreia@ubi.pt

Pedro Jerónimo

Researcher at LabCom - Communication and Arts, research unit at the University of Beira Interior, Covilhã, Portugal. Email: pj@ubi.pt

Inês Amaral

Associate Professor at the Department of Philosophy, Communication and Information at the University of Coimbra, Coimbra, Portugal. Email: ines.amaral@uc.pt

This book is part of MediaTrust.Lab - Local Media Lab for Civic Trust and Literacy, a project funded by the Foundation for Science and Technology (PTDC/COM-JOR/3866/2020), Portugal.

LabCom Books

www.labcom.ubi.pt