

# Unraveling Deceptive Narratives: A Study of Conceptual Frameworks

Elie Alhajjar

RAND

Arlington, VA, USA

[eahajjar@rand.org](mailto:eahajjar@rand.org)

## Abstract

In the era of digital communication, the rapid spread of information presents significant challenges to society. This paper provides an in-depth examination of the existing frameworks developed to understand and address these phenomena. More precisely, this paper categorizes and compares various frameworks, including typology-based, process-oriented, impact-oriented, and actor-centric approaches. It highlights the strengths and limitations of each framework type, with a particular focus on their applicability to combat false information in diverse contexts. The paper underscores the importance of adopting a holistic and flexible approach that integrates multiple frameworks and adapts to the evolving nature of technology, particularly AI-driven false and misleading content.

**Keywords:** Information spread, frameworks, policy.

## 1. Introduction

As false information spreads more rapidly and widely than ever before, understanding the frameworks that scholars and practitioners use to analyze and address these phenomena is essential for developing effective strategies to combat its spread, protect public trust, and mitigate its societal impacts. The spread of falsehoods, both unintentional and strategically engineered, has become increasingly significant. Both aspects refer to the spread of false or misleading information, but they differ fundamentally in purpose and impact. Understanding these differences is crucial for comprehending the broader challenges they pose to society.

Some forms of false content are shared without malicious intent (Baines and Elliott, 2020). Often, individuals or organizations believe what they are sharing to be true, despite it being incorrect. This can stem from misunderstandings, insufficient verification, or the use of unreliable sources. For example, during fast-moving news events, early reports that lack confirmation are frequently circulated by well-meaning individuals trying to make sense of unfolding developments. Even when shared in good faith, such content can perpetuate errors and mislead others.

In contrast, other forms are deliberately constructed and disseminated to deceive or manipulate (Baines and Elliott, 2020). These efforts are typically intentional and often orchestrated by actors seeking specific outcomes such as shaping public opinion, influencing elections, harming reputations, or sowing discord. These campaigns are frequently sophisticated, exploiting digital infrastructure and cognitive biases to extend their reach. Their strategic and deliberate

nature makes them especially dangerous, as they erode institutional trust, deepen social divisions, and threaten the stability of democratic systems.

The complexity of detecting and responding to false information further underscores the importance of studying these phenomena. The methods used to spread false information are becoming more sophisticated, employing technologies such as bots, deepfakes, and algorithmic amplification to reach wider audiences and evade detection (Moral, 2022). This technological evolution makes it more challenging to identify and counteract misleading content effectively. In addition, the proliferation of such content raises important ethical and legal questions (Dalessandro et al., 2019). Society must navigate the delicate balance between protecting free speech and preventing harm, determining the responsibilities of platforms and governments in regulating content, and upholding the public's right to accurate information.

This paper aims to provide a comprehensive overview of the key frameworks developed to understand and address inaccurate or deceptive narratives. It starts with an introduction that sets the context, followed by an overview of historical context. The core sections present and categorize various theoretical and methodological frameworks, offering a comparative analysis to highlight their strengths, weaknesses, and applicability. The discussion synthesizes key insights, addressing the broader implications and suggesting future research directions.

## 2. Historical Context

The historical context of false and misleading content highlights the persistent nature of these

challenges and underscores the importance of understanding their evolution. The phenomena of circulating deceptive or inaccurate narratives are not new; they have existed in various forms throughout history. However, the methods and speed at which false information can spread have evolved dramatically with the advent of new technologies.

Historically, such content was often spread through word of mouth, pamphlets, or print media. In ancient civilizations, this often took the form of rumors or oral traditions that were passed down through generations. These could include exaggerated tales of heroism, false reports of enemy movements during wars, or distorted accounts of natural disasters. One of the earliest documented cases of strategic falsehoods dates back to the Roman Empire (Frasier, 2020), where political rivals would spread false rumors to undermine each other's credibility.<sup>1</sup>

In the medieval period, the Catholic Church played a significant role in controlling the flow of information. The selective use or distortion of truth became a tool used to enforce religious orthodoxy and suppress dissent. For instance, during the Inquisition, accusations of heresy were often based on exaggerated information, leading to the persecution of individuals or groups who were seen as threats to the church's authority. The church also controlled the dissemination of knowledge, often suppressing scientific discoveries that contradicted religious doctrine, such as the heliocentric theory proposed by Copernicus (Hess and Allen, 2008).

The invention of the printing press in the 15th century marked a significant turning point in this context. While the printing press revolutionized the spread of information, it also made it easier to disseminate false information on a larger scale. Pamphlets, leaflets, and books could now be mass-produced and distributed widely, often without any means of verifying the accuracy of the content. This period saw the rise of propaganda as a powerful tool in religious and political conflicts, such as during the Protestant Reformation, where both sides used printed materials to spread their messages and discredit their opponents (Edwards, 1994). Martin Luther's 95 Theses, for example, were widely circulated, but so were counter-pamphlets filled with distorted portrayals of his arguments and intentions.

The 19th and early 20th centuries witnessed the growth of mass media, particularly newspapers,

which became the primary means of information dissemination. This era also saw the emergence of yellow journalism, a term used to describe sensationalist and often misleading news designed to attract readers and increase sales. Publishers like William Randolph Hearst and Joseph Pulitzer engaged in fierce competition, leading to the publication of fabricated stories (Spencer, 2007). The Spanish-American War is a prime example of how yellow journalism contributed to the spread of misleading information, as reports in American newspapers fueled public support for the war, often based on false accounts of events.

During World Wars I and II, propaganda was widely used by various governments to influence public perception and morale, often blurring the lines between true and falsified information. The Cold War era saw the rise of sophisticated information campaigns, particularly by state actors. The Soviet Union, for instance, engaged in active measures, a term used to describe a range of activities aimed at influencing global opinion and destabilizing Western governments. An infamous example was Operation INFEKTION,<sup>2</sup> which falsely claimed that the U.S. government had created the HIV/AIDS virus as a biological weapon (Selvage, 2021). These activities were meticulously planned and executed, often using forgeries, false media reports, and manipulated documents to create and spread false narratives.

In the modern era, the internet and social media have fundamentally reshaped how inaccurate and manipulative content is produced and circulated. The internet has made it easier than ever to create and spread false information on a global scale with minimal cost and effort (Vosoughi et al., 2018). Social media platforms, in particular, have become fertile ground for the rapid dissemination of all kinds of information, as content can go viral within minutes, reaching millions of people around the world. The decentralized nature of the internet also means that it is more challenging to control or regulate the spread of false information, leading to a proliferation of conspiracy theories, fake news, and other forms of misleading content.

### 3. Existing frameworks

A variety of frameworks have been developed to understand, categorize, and address how false or misleading content is created, spread, and mitigated. In this section, we explore the major existing frameworks, categorizing them based on their focus and approach.

---

<sup>1</sup> Octavian (later Augustus) waged a propaganda campaign against his rival Mark Antony by spreading rumors portraying Antony as a traitor and a puppet of Cleopatra.

---

<sup>2</sup> Also known as Operation Denver.

### 3.1 Typology-Based Frameworks

They focus on categorizing falsehoods into distinct types based on various factors such as intent, content, and dissemination method.

1) **Intent-Based Typologies:** One of the most common typological approaches is to classify false information based on the intent behind its creation and dissemination. For example, some frameworks distinguish between unintentional and intentional spread as primary categories. Further subcategories might include the deliberate spread of truthful information with the intent to cause harm, such as doxing or releasing private information (Carmi et al., 2020).

2) **Content-Based Typologies:** Another approach focuses on the nature of the content itself. These frameworks classify information spread based on the type of falsehood or distortion present in the content (First Draft, 2017). For example, Wardle introduced a typology that categorizes false information into seven types: satire or parody, false connection, misleading content, false context, impostor content, manipulated content, and fabricated content. Each type represents a different way in which the truth is distorted, providing a detailed map of the information landscape.

3) **Dissemination Method-Based Typologies:** Some frameworks classify falsehoods based on the methods and channels used to spread them. These might include distinctions between organic spread (e.g., via social media sharing) and coordinated campaigns (e.g., through bot networks or paid advertisements). Understanding the dissemination methods helps in identifying the mechanisms by which false information reaches and influences audiences (Tsfati et al., 2020).

### 3.2 Process-Oriented Frameworks

They focus on the lifecycle of false and misleading narratives, examining how these phenomena are created, disseminated, consumed, and ultimately affect audiences. These frameworks often draw from communication and media studies to map out the stages through which false information travels, and the factors that influence each stage.

1) **The Information Disorder Framework:** This type of frameworks identifies three key stages in the lifecycle of false information: creation, production, and distribution (Wardle and Derakhshan, 2017). It also distinguishes between three elements involved: agents (creators, producers, and distributors), messages (the content itself), and interpreters (audiences). This framework is useful for

understanding how deceptive narratives are constructed and spread across different platforms and contexts.

2) **The Information Lifecycle Model:** Another process-oriented approach is the information lifecycle model, which outlines the stages through which false or misleading content moves from its initial creation to its eventual impact on public perception. These stages typically include creation, amplification, dissemination, and correction (Puska and Pereira, 2023). This model emphasizes the role of social media algorithms, news cycles, and audience engagement in the amplification and spread of such content.

3) **The Knowledge-Based Approach:** This type of framework focuses on how individuals process and interpret information (Bode and Vraga, 2017). It examines the cognitive processes that occur when people encounter false information, including how they decide whether to believe it or share it. The model suggests interventions at different stages of information processing, such as providing corrective information or promoting critical thinking skills, to reduce the spread and impact of misleading information.

### 3.3 Impact-Oriented Frameworks

They are concerned with the consequences of falsehoods on individuals, communities, and societies. They assess the effects of false information and help identify the broader implications on public opinion. Each model below represents a cluster of frameworks rather than a single framework, considered as a many-to-one mapping.

1) **The Trust Erosion Model:** This family of frameworks explores how false information campaigns erode public trust in institutions, media, and democracy (Atele-Williams and Marsh, 2023). It posits that repeated exposure to false information, especially when it aligns with existing biases or distrust, leads to a gradual decline in trust. The model is particularly relevant for understanding the long-term societal impacts of falsehoods and the challenges in restoring trust once it has been damaged.

2) **The Public Health Impact Model:** This family of models examines the spread of health-related narratives (e.g., about vaccines or treatments) and its impact on public health, such as vaccine hesitancy or non-compliance with health guidelines (Pulido et al., 2020). The framework also considers the role of public health communication in countering such narratives and promoting accurate information.

3) **Behavioral Impact Model:** This cluster of frameworks looks at how information flows

influence individual and collective behavior (Alhajjar, 2022). It considers factors such as cognitive biases, social influence, and emotional responses that lead individuals to accept or act on false information. The framework is useful for designing interventions that address the behavioral drivers of false information spread, such as social norms campaigns or behavioral nudges.

### 3.4 Actor-Centric Frameworks

They focus on the roles and motivations of different actors involved in the creation, dissemination, and consumption of distorted narratives. These frameworks analyze the behaviors, strategies, and networks of various stakeholders, including individuals, organizations, governments, and platforms.

1) The Actor-Network Theory (ANT) : This sociological framework explores the complex relationships between different actors (both human and non-human, such as algorithms) involved in the spread of falsehoods (Latour, 2005). ANT examines how these actors form networks that facilitate the dissemination of false information and how power dynamics within these networks influence the spread and impact of information. The framework is useful for understanding the interconnectedness of different actors and the systemic nature of information ecosystems.

2) The Political Economy Framework: This approach focuses on the economic and political motivations behind false information campaigns (Dobson and Hunsinger, 2016). It examines how state and non-state actors use information as a tool for political gain, financial profit, or social influence. The framework also considers the role of media ownership, advertising revenue models, and regulatory environments in shaping the spread of misleading information. Understanding these motivations is crucial for designing policies and interventions that address the root causes of falsified information.

3) The Platform Responsibility Framework: With the rise of social media and digital platforms, this framework addresses the responsibilities of these platforms in managing falsehoods (Ramdas, 2022). It examines the role of algorithms, content moderation policies, and platform governance in either exacerbating or mitigating the spread of false information. The framework also explores the ethical and legal implications of platform actions, such as content removal or algorithmic transparency.

## 4. Categorization of frameworks

The diverse array of frameworks for understanding false and misleading content can be overwhelming due to the various perspectives and methodologies they encompass. In this section, we divide the existing frameworks into thematic, methodological, and geographical or cultural considerations. Practically, when frameworks are categorized by themes, methods, or regional contexts, the focus is on tailoring research approaches and interventions to specific settings or problems. In contrast, when frameworks are grouped into actor, process, impact, and typology categories, they are examined based on what aspect of the problem they target.

### 4.1 Thematic Categorization

It involves grouping frameworks based on the primary themes or issues they address. This approach helps illuminate the specific domains each model concentrates on, whether it be political, social, health-related, or technological (Tsfati et al., 2020; Pulido et al., 2020).

1) Political Frameworks: Frameworks in this category focus on the role of false narratives in political contexts. They examine how false information is used to influence elections, shape public opinion, and destabilize political systems (Howard et al., 2018; Wardle & Derakhshan, 2017). For example, frameworks that analyze information campaigns during elections or state-sponsored propaganda efforts fall into this category. These frameworks often emphasize the strategic use of deceptive arguments by political actors to achieve specific goals, such as voter manipulation or undermining opponents.

2) Social Frameworks: Socially oriented frameworks explore how information can affect social dynamics and relationships. They may focus on how false information spreads within communities, influences social norms, or exacerbates societal divisions. Frameworks in this category often address issues like the role of social media in amplifying falsehoods, the formation of echo chambers, and the impact of misleading information on social cohesion (Sunstein, 2017; Vosoughi et al., 2018). These frameworks are particularly relevant for understanding how information contributes to polarization and the fragmentation of public discourse.

3) Health-Related Frameworks: Given the significant impact of false information on public health, several frameworks specifically address the spread and effects of health-related falsehoods (Pulido et al., 2020; Bode & Vraga,

2017). Health-related frameworks often emphasize the need for accurate communication, the dangers in undermining public health efforts, and strategies for combating inaccurate health narratives through education and public awareness campaigns.

4) **Technological Frameworks:** Technological frameworks focus on the role of digital platforms, algorithms, and artificial intelligence (AI) in the spread of information. They explore how technology facilitates the rapid dissemination of false information, the role of social media algorithms in promoting sensationalist content, and the potential for automated tools like bots and deepfakes to spread information (Gradoń et al., 2021; Gillespie, 2018). These frameworks often address the challenges of regulating digital platforms and the ethical implications of technological interventions designed to counter misleading information.

## 4.2 Methodological Approaches

They group frameworks based on the research methods they employ. This categorization highlights the diversity of techniques used to study misleading content, ranging from qualitative analyses to quantitative data-driven models (Lazer et al., 2018).

1) **Qualitative Frameworks:** Qualitative frameworks often involve case studies, interviews, content analysis, and other non-numerical methods to explore false and misleading content (Wardle & Derakhshan, 2017). These frameworks are valuable for understanding the nuanced and contextual factors that influence how false information is created, spread, and received. For example, qualitative studies may examine the narratives used in information campaigns, the motivations of actors involved in spreading false information, or the experiences of individuals who encounter false and misleading content.

2) **Quantitative Frameworks:** Quantitative frameworks rely on numerical data and statistical analysis to study false and misleading content. These frameworks often involve large-scale data collection, such as social media analytics, survey data, or experiments designed to measure the effects of false and misleading content (Vosoughi et al., 2018). Quantitative approaches are useful for identifying patterns in the spread of false and misleading content, assessing the prevalence of false information, and evaluating the effectiveness of interventions.

3) **Mixed-Methods Frameworks:** Some frameworks combine qualitative and quantitative approaches to offer a more comprehensive understanding of false and

misleading content. Mixed-methods frameworks might use qualitative research to explore the context and motivations behind false and misleading content, followed by quantitative analysis to measure the scale and impact of these phenomena (Tsfati et al., 2020). This approach allows for a more holistic view, capturing both the detailed, context-specific elements and the broader trends in false and misleading content spread and impact.

## 4.3 Computational Frameworks

With the rise of big data and machine learning, computational frameworks have become increasingly important in the study of false and misleading content. These frameworks use algorithms, network analysis, and other computational tools to model the spread of false and misleading content, detect false information, and simulate the effects of different interventions (Gradoń et al., 2021; Lazer et al., 2018).

1) **Geographical and Cultural Considerations** involve categorizing frameworks based on the regions or cultural contexts in which they are applied. False and misleading content does not operate in a vacuum; it is deeply influenced by the social, cultural, and political environments in which it spreads (Tsfati et al., 2020).

2) **Regional Frameworks:** Some frameworks are designed to address false and misleading content in specific geographical regions, such as North America, Europe, Asia, or Africa. These frameworks consider the unique political, social, and media landscapes of each region, which influences how false and misleading content spreads and is perceived (Howard et al., 2018). For example, frameworks developed for Western democracies might focus on the role of free speech and the media, while those for authoritarian regimes might emphasize state control and censorship.

3) **Cultural Frameworks:** Cultural frameworks examine how cultural factors, such as language, values, and traditions, shape the creation and spread of false and misleading content. These frameworks recognize that false and misleading content is often tailored to resonate with specific cultural beliefs or biases, making it more effective in certain communities. For instance, information campaigns may exploit cultural tensions or stereotypes to create division or mistrust.

4) **Cross-Cultural Frameworks:** Cross-cultural frameworks compare the spread and impact of false and misleading content across different cultural contexts. These frameworks are useful for identifying universal patterns in false and misleading content spread, as well as context-specific factors that influence how false and

misleading content is received and acted upon (Pennycook and Rand, 2021). Cross-cultural studies can reveal how different societies respond to false and misleading content and what lessons can be learned from various approaches to combating false information.

## 5. Discussion

In this section, we undertake a comparative analysis of the existing frameworks for understanding false and misleading content. By systematically comparing these frameworks, we aim to identify their strengths, weaknesses, and areas of overlap or divergence. This analysis will help clarify which frameworks are most effective in addressing specific aspects of false and misleading content based on scope and effectiveness and summarize key insights from the study and implications for future research in this field.

### 5.1 Comparative Analysis

Typology-based frameworks are designed to offer a comprehensive categorization of false and misleading content, attempting to classify all forms of false information into a structured taxonomy. By creating categories based on criteria such as intent, content, and medium, typology frameworks allow for a systematic analysis of the different types of false and misleading content that exist. This broad classification system is advantageous because it provides a high-level overview that can help in identifying patterns and trends in the spread of false and misleading content. However, the very breadth of typology frameworks can also be a limitation as they may not delve deeply into the specific processes that lead to the creation and dissemination of false and misleading content. Process-oriented frameworks, in contrast, focus on the lifecycle of false and misleading content. This narrower focus allows for detailed insights into the stages of false and misleading content spread, identifying critical points where interventions could be most effective. By understanding these processes, stakeholders can develop targeted strategies to disrupt the spread of false and misleading content at key stages. However, the focus on processes can limit the ability of these frameworks to account for the broader social, political, or cultural contexts that influence the spread of false and misleading content. While they provide valuable insights into the mechanics of false and misleading content dissemination, process frameworks may not fully capture the external factors that shape the environment in which false and misleading content thrives.

Impact-oriented frameworks take a different approach by concentrating on the consequences of false and misleading content, rather than its classification or lifecycle. These frameworks are particularly effective in highlighting the tangible effects of false and misleading content by linking false information to specific outcomes. By focusing on the measurable consequences of false and misleading content, impact frameworks provide critical insights into the harm caused by false information and the importance of addressing it. However, the reliance on measurable outcomes can be both a strength and a limitation. While impact frameworks excel in demonstrating the immediate and direct effects of false and misleading content, they may struggle to capture the full range of impacts, particularly those that are long-term, indirect, or difficult to quantify.

Finally, actor-centric frameworks offer a broad scope by considering the wide range of players involved in the creation, dissemination, and consumption of false and misleading content, as well as the complex relationships between them. By focusing on the motivations and behaviors of key actors, actor-centric frameworks can reveal the underlying drivers of false and misleading content. However, the inherent complexity of actor-centric frameworks can make them challenging to apply. The interactions between various actors are often intricate and not easily discernible, especially when motivations are hidden or intentionally obscured. This complexity requires significant resources and expertise to untangle, making actor-centric frameworks more difficult to implement effectively compared to other frameworks that focus on more straightforward aspects of false and misleading content.

### 5.2 Key Insights

Each framework brings unique strengths to the table, contributing valuable perspectives on how false information is generated, disseminated, and impacts society. However, the analysis also highlights the limitations of each approach, suggesting that a multifaceted strategy combining elements from multiple frameworks may be the most effective way to combat false and misleading content. One of the most significant insights is that no single framework can fully address the complexity of false and misleading content. This suggests that relying on one framework alone may lead to an incomplete understanding of the problem and potentially ineffective interventions. Another important insight is the critical role that context plays in the effectiveness of different frameworks. Effectiveness in this context is

measured by how well a framework achieves its intended purpose, which can vary depending on the framework's focus. False and misleading narratives are deeply influenced by social, political, and cultural factors in different environments. In typology-based frameworks, effectiveness is measured by how well the framework categorizes different types of false and misleading content based on key factors like intent, content, or dissemination method. A typology framework is considered effective if it provides a clear, comprehensive, and useful classification system that helps researchers and practitioners distinguish between various forms of false information. For process-oriented frameworks, effectiveness is determined by their ability to map the lifecycle of false and misleading content, identify critical intervention points, and develop strategies to disrupt its dissemination. Impact-oriented frameworks are judged by how accurately they assess the consequences of false and misleading content, such as changes in public opinion or behavior, while actor-centric frameworks are evaluated based on their capacity to reveal the motivations and behaviors of those involved in spreading false and misleading content.

False and misleading content that resonates in one cultural setting may not have the same impact in another, and the strategies used to combat it must be tailored accordingly. The analysis also highlights that false and misleading content is not solely a communication issue but also intersects with other disciplines, bringing diverse methodologies and insights to the discussion. By combining these perspectives, a more robust and comprehensive understanding of false and misleading content can be developed. Moreover, the analysis reveals that the rapid evolution of digital technologies necessitates continuous adaptation of existing frameworks. False and misleading narratives are increasingly spread through new and evolving platforms, such as social media, where traditional approaches may no longer be sufficient. This dynamic environment requires frameworks that are not only comprehensive but also flexible and adaptable to change, and that can keep pace with technological advancements as well as the changing nature of information dissemination.

### 5.3 Implications for Future Research

One of the primary implications for future research is the need for more integrative approaches that combine the strengths of multiple frameworks. For instance, combining typology frameworks with process frameworks could provide a more comprehensive

understanding of both the classification of false and misleading content and the mechanisms by which it spreads. Similarly, integrating actor frameworks with impact frameworks could help elucidate how the motivations of key players influence the tangible outcomes of false and misleading content. Future research should prioritize developing hybrid frameworks that draw on the strengths of existing models while addressing their respective shortcomings.

On the level of contextualization of false and misleading content, future research should focus on comparative studies that examine how false and misleading content operates across diverse contexts, including non-Western societies that are often underrepresented in the literature. This would not only broaden the understanding of false and misleading content globally but also inform the development of context-specific interventions that are more likely to be effective in diverse environments.

In addition to the plethora of false and misleading content instances, a new phenomenon has emerged in the last couple of years: AI-enabled false and misleading content (Newsguard, 2025). AI-driven technologies, which include everything from automated news outlets that produce content with minimal or no human intervention to sophisticated AI image generators that create convincing but entirely fabricated visuals, have opened new avenues for the production and dissemination of misleading information (Alhajjar and Lee, 2022). AI's capabilities to generate large volumes of content quickly and convincingly, false and misleading content purveyors now have powerful tools at their disposal to create and spread false narratives on an unprecedented scale. This development poses serious challenges to the integrity of information ecosystems. The line between genuine and fabricated content increasingly blurs, making it harder for the public to distinguish truth from falsehood. The ease with which these tools can be used to produce deceptive content underscores the urgent need for robust strategies to detect and counteract AI-generated false and misleading content.

Finally, there is a pressing need for interdisciplinary research that brings together scholars from various fields to tackle the complex problem of false and misleading content. Future research should encourage collaboration across these fields to develop more comprehensive and multidimensional frameworks. This interdisciplinary approach would facilitate a deeper understanding of the psychological, social, and technological factors that drive false and misleading content, leading to more effective strategies for prevention and intervention.

## 6. Conclusion

The challenges posed by false and misleading content are among the most pressing issues facing societies today. The frameworks discussed in this paper provide valuable tools for analyzing the spread of false and misleading content and pinpoint the limitations of any single approach.

The selection and improvement of frameworks for combating false and misleading content requires a nuanced understanding of the problem's complexity. Each framework, whether typology-based, process-oriented, impact-focused, or actor-centric, offers distinct advantages and limitations. By choosing the right framework for the specific context and continuously improving upon existing models, researchers and practitioners can develop more effective strategies to counter false and misleading content. The key to addressing new challenges, including AI-enabled false and misleading content, lies in embracing a holistic and flexible approach. By integrating multiple frameworks, adapting strategies to specific contexts, and fostering interdisciplinary collaboration, more effective methods can be developed for combating false and misleading content.

As technology continues to evolve, it is essential to remain vigilant and proactive, continuously updating and refining our frameworks to keep pace with new developments. This ongoing effort requires not only innovation but also a commitment to transparency, accountability, and public trust. By staying ahead of emerging threats and fostering a culture of critical thinking, we can build a more resilient information ecosystem for the future.

## 7. Acknowledgments

The author would like to thank Bridget Chan at New America, DC for her support during the project. Her comments made major improvements to the clarity of the document.

## 8. Bibliographical References

- Alhajjar, E. (2022). Alternate reality: The use of disinformation to normalize extremism. In *The Great Power Competition Volume 3: Cyberspace: The Fifth Domain* (pp. 157-165). Springer International Publishing.
- Alhajjar, E., & Lee, K. (2022). The U.S. cyber threat landscape. In *European Conference on Cyber Warfare and Security* (Vol. 21, No. 1, pp. 18-24).

- Atele-Williams, T., & Marsh, S. (2023). Information trust model. *Cognitive Systems Research*, 80, 50-70.
- Baines, D., & Elliott, R. J. R. (2020). Defining misinformation, disinformation and malinformation: An urgent need for clarity during the COVID-19 infodemic. *Discussion Papers*, 20(06), 20-06.
- Bode, L., & Vraga, E. K. (2017). See something, say something: Correction of global health misinformation on social media. *Health Communication*, 33(9), 1131-1140. <https://doi.org/10.1080/10410236.2017.1331312>
- Carmi, E., Yates, S. J., Lockley, E., & Pawluczuk, A. (2020). Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation. *Internet Policy Review*, 9(2), 1-22.
- Dalessandro, R. C., Guimarães, J. A. C., & Campbell, D. G. (2019). Fake news as an emergent subject domain: Conceptual and ethical perspectives for the development of a critical knowledge organization. In *The Human Position in an Artificial World: Creativity, Ethics and AI in Knowledge Organization* (pp. 208-217). Ergon-Verlag.
- Dobson, K., & Hunsinger, J. (2016). The political economy of WikiLeaks: Transparency and accountability through digital and alternative media. *Interactions: Studies in Communication & Culture*, 7(2), 217-233.
- Edwards Jr, M. U. (1994). *Printing, Propaganda, and Martin Luther*. Fortress Press.
- First Draft (2017). Fake news. It's complicated. <https://www.firstdraftnews.org/articles/fake-news-complicated>
- Fraser, M. (2020). *In truth: a history of lies from ancient Rome to modern America*. Rowman & Littlefield.
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven: Yale University Press. <https://doi.org/10.12987/9780300235029>
- Gradoń, K. T., Hołyst, J. A., Moy, W. R., Sienkiewicz, J., & Suhecki, K. (2021). Countering misinformation: A multidisciplinary approach. *Big Data & Society*, 8(1).
- Hess, P. M., & Allen, P. L. (2008). *Catholicism and science*.
- Howard, P. N., Ganesh, B., Liotsiou, D., Kelly, J., & François, C. (2018). *The IRA, social media and political polarization in the United States, 2012-2018*. Computational Propaganda Research Project, University of Oxford.

- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
- Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094-1096.
- Moral, P. (2022). The challenge of disinformation for national security. In *Security and Defence: Ethical and Legal Challenges in the Face of Current Conflicts* (pp. 103-119). Springer International Publishing.
- Newsguard (2025). AI Tracking Center. <https://www.newsguardtech.com/special-reports/ai-tracking-center/>
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in cognitive sciences*, 25(5), 388-402.
- Pidot, J. R. (2019). Environmental nihilism. *Arizona Journal of Environmental Law & Policy*, 10(1).
- Pulido, C. M., Ruiz-Eugenio, L., Redondo-Sama, G., & Villarejo-Carballido, B. (2020). A new application of social impact in social media for overcoming fake news in health. *International Journal of Environmental Research and Public Health*, 17(7), 2430.
- Puska, A. A., & Pereira, R. (2023). Exploring digital misinformation as a sociotechnical phenomenon: Insights from a small-scale study. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems* (pp. 1-12).
- Ramdas, V. (2022). Identifying an actionable algorithmic transparency framework: A comparative analysis of initiatives to enhance accountability of social media platforms. *National Law University Delhi Student Law Journal*, 4, 74.
- Selvage, D. (2021). Operation "Denver" The East German Ministry for State Security and the KGB's AIDS Disinformation Campaign, 1986–1989 (Part 2). *Journal of Cold War Studies*, 23(3), 4-80.
- Spencer, D. R. (2007). *The yellow journalism: The press and America's emergence as a world power*. Northwestern University Press.
- Sunstein, C. R. (2018). *#Republic: Divided Democracy in the Age of Social Media* (New edition). Princeton University Press. <https://doi.org/10.2307/j.ctv8xnhtd>
- Tsfati, Y., Boomgaarden, H. G., Strömbäck, J., Vliegenthart, R., Damstra, A., & Lindgren, E. (2020). Causes and consequences of mainstream media dissemination of fake news: Literature review and synthesis. *Annals of the International Communication Association*, 44(2), 157-173.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *science*, 359(6380), 1146-1151.
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking* (Vol. 27). Council of Europe.