

Review

# Mapping the Impact of Generative AI on Disinformation: Insights from a Scoping Review

Alexandre López-Borrull <sup>1,\*</sup>  and Carlos Lopezosa <sup>2</sup>

<sup>1</sup> Estudis Ciències de la Informació i la Comunicació, Universitat Oberta de Catalunya, 08035 Barcelona, Spain

<sup>2</sup> Faculty of Information and Audiovisual Media, Universitat de Barcelona, 08014 Barcelona, Spain; lopezosa@ub.edu

\* Correspondence: alopezbo@uoc.edu

## Abstract

This article presents a scoping review of the academic literature published between 2021 and 2024 on the intersection of generative artificial intelligence (AI) and disinformation. Drawing from 64 peer-reviewed studies, the review examines the current research landscape and identifies six key thematic areas: political disinformation and propaganda; scientific disinformation; fact-checking; journalism and the media; media literacy and education; and deepfakes. The findings reveal that generative AI plays a dual role: it enables the rapid creation and targeted dissemination of synthetic content but also offers new opportunities for detection, verification, and public education. Beyond summarizing research trends, this review highlights the broader societal and practical implications of generative AI in the context of information disorder. It outlines how AI tools are already reshaping journalism, challenging scientific communication, and transforming strategies for media literacy and fact-checking. The analysis also identifies key policy and governance challenges, particularly the need for coordinated responses from governments, platforms, educators, and civil society actors. By offering a structured overview of the field, the article enhances our understanding of how generative AI can both exacerbate and help mitigate disinformation, and proposes directions for research, regulation, and public engagement.



Academic Editor: Andrew Kirby

Received: 11 May 2025

Revised: 16 July 2025

Accepted: 17 July 2025

Published: 21 July 2025

**Citation:** López-Borrull, A., & Lopezosa, C. (2025). Mapping the Impact of Generative AI on Disinformation: Insights from a Scoping Review. *Publications*, *13*(3), 33. <https://doi.org/10.3390/publications13030033>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** disinformation; generative AI; ChatGPT; fact-checking; deepfakes

## 1. Introduction

Although the introduction and theoretical framework are normally two separate sections in a standard academic paper, in this case, given the methodology, this section will provide a description of the concepts related to the study topic.

### 1.1. Fake News, Disinformation, and Denial: An Old Reality Amplified by Social Media and Global Challenges

Much has been said about fake news in recent years. Clearly, the concept is not new, as shown by books such as [Correas and Kenneally \(2019\)](#) on disinformation and power, and the controlled use of information and deception to produce an effect or, as the authors put it, to change history. When examining fake news as a phenomenon associated with the media, a frequently cited example is a series of stories published in the New York newspaper *The Sun* in 1835, which described how a scientist had observed living beings on the Moon using a powerful telescope ([Salas Abad, 2019](#)).

However, the more recent phenomenon of fake news has emerged thanks to (or because of) the expansion of social media. It is a phenomenon that has had global and

media-wide resonance since the 2016 United States elections, but it has continued to gain strength and take new forms in successive electoral contests and in armed conflict, particularly in the war between Russia and Ukraine (Pierré et al., 2023) and the invasion of Gaza by Israel (Kharroub, 2023). In this context, for a more precise definition of “fake news”, we refer to Rochlin (2017), who defined it as “a deliberately false headline or story published on a website that pretends to be a news outlet.”

This distinction is important, as the term has often been used almost synonymously with lies or disinformation in general; however, in the academic context, it should be used to refer to lies disseminated through sites that pose as reliable sources of information.

Thus, despite the significance of fake news in spreading lies during key moments, such as election campaigns, there is a broader concept that encompasses it and is also used in the fields of politics and content curation on social media. This is disinformation, understood as false or inaccurate information shared on social networks and other digital platforms or environments. In this sense, disinformation refers to false content designed to cause harm but with a specific intent (Wardle, 2020). When disinformation is shared, it often becomes misinformation, which also refers to false content, but where the person sharing it is unaware that it is false. A third term, malinformation, can also be added—this refers to true information shared with the intent to cause harm. A particularly illustrative case was when Russian agents hacked emails from the Democratic National Committee and Hillary Clinton’s campaign, leaking details to the public in an effort to damage her reputation. It should be clarified that this third concept will not be addressed in this study.

However, the renewed intensity of the disinformation and misinformation phenomenon in recent years has emerged largely due to the proliferation of social media platforms, which have accelerated the spread of unverified and emotionally charged content. As Tomassi et al. (2025) argue, digital media ecosystems amplify information disorder by fostering rapid dissemination, low editorial oversight, and algorithmic reinforcement of polarizing narratives, particularly around topics such as climate change and public health. And it has become an even greater topic of discussion since the 2016 elections in the United States.

Among the key authors who help better understand this phenomenon is McNair (2018), who offers an interesting perspective by placing fake news in the same framework as populism and the loss of prestige of elites and the media, and by extension, of journalists. This helps us understand the challenge and risk that disinformation poses, acting both as a barometer and as a tool in many of today’s social and political conflicts. At the same time, to fully grasp the threat, studies such as Vosoughi et al. (2018) show that fake news spreads faster than the truth, especially in relation to events that threaten people’s safety (such as earthquakes or terrorist attacks).

Finally, one of the most significant recent events marked by disinformation was COVID-19, undoubtedly a global phenomenon that became a new arena for the spread of false information. In fact, the World Health Organization coined the term infodemic on 2 February 2020 (World Health Organization, 2020), just days or weeks before the lockdowns introduced by European countries. Specifically, it was announced that “the 2019-nCoV outbreak and response have been accompanied by a massive ‘infodemic’—an overabundance of information, some accurate and some not, that makes it hard for people to find trustworthy sources and reliable guidance when they need it.”

It is quite significant, therefore, that even before a consensus name for COVID-19 had been reached, there was already talk of the large amount of disinformation circulating, labeled as an infodemic—a concept that accompanied the pandemic itself.

All this highlights the urgency of combating disinformation. In fact, the European Commission report tackling online disinformation (European Commission, 2018) empha-

sized the need for clear policies and a global effort to address this phenomenon. These premises serve as the foundation for our research.

### 1.2. Generative Artificial Intelligence: A New Tool with a Huge Potential to Transform the Way We Create and Consume Content

Generative artificial intelligence (generative AI) is a subfield of artificial intelligence that involves creating new content based on existing data. This technology has experienced enormous advances in recent years, enabling the generation of text, images, music, and other types of data. Tools like ChatGPT and Copilot are examples of generative AI that have reached the general public, generating debates and challenges, but also opportunities.

In this regard, generative AI has the potential to transform the way we create and consume content. However, it also raises concerns about inequality in access to technology and its potential benefits. Regarding ethical and legal aspects, there is intense debate about its implications and impact—not only on the labor market but also on intellectual property, plagiarism, and biases in the sources on which its knowledge is built (Hosseini et al., 2024). This includes the transparency of its algorithms, ensuring that users understand how content is generated. Fairness is another key aspect, as biases in the models that could lead to discrimination or the generation of harmful content must be prevented. Finally, moral concerns about the use of generative AI raise questions about the systems' ability to make decisions aligned with universal human rights and values.

Nowadays, especially with the rise of generative AI, a wide range of relevant studies have been conducted that examine AI from various perspectives. From a social sciences standpoint, there is clear evidence of the growing use of AI within the discipline (Bircan & Salah, 2022). Some studies suggest that this technology can be understood as an emerging general method of scientific invention (Bianchini et al., 2022), while others highlight AI's role both as a paradigm supporting the advancement of fundamental science (Xu et al., 2021) and as a tool for science education in particular (Almasri, 2024).

Likewise, in the last two years, we have seen a proliferation of studies on the introduction of generative AI in multiple fields, such as education (Bond et al., 2024; Chan & Hu, 2023), public relations (CIPR, 2023), and scientific publishing (Zhou, 2024). Given the risk of merely cosmetic change, there is a recognized need to reflect on and assimilate the technology as it enters our daily lives. In fact, it is a process similar to what occurred with the arrival of ICT and, in particular, search engines.

### 1.3. Uses of Generative AI to Combat Disinformation

Generative AI has the potential to become a powerful tool in the fight against disinformation, especially in today's digital era, where information spreads rapidly and often without verification. Among its possible uses is automated data verification and fact-checking, which could help develop systems capable of verifying the accuracy of online claims by comparing them against databases of verified facts, thereby improving the efficiency of fact-checkers and journalists (Cuartielles et al., 2023).

It can also be used to create verified content through audiovisual generation tools, as well as in fake news detection, by training models on the common patterns of such content to alert users or information managers (González et al., 2022). Additionally, it is useful in the development of media literacy materials (Uribe & Machin, 2024) and in social media monitoring, allowing real-time supervision of potentially misleading content—though it should not replace human judgment in decisions regarding its potential removal.

These applications could significantly help reduce the spread of disinformation and promote more accurate and trustworthy information. However, the existence of AI hallucinations, particularly in early versions of generative AI, must be taken into account. These hallucinations refer to the generation of content that appears coherent and factually

plausible but is actually false or fabricated by the model due to the way it predicts and constructs language based on statistical patterns rather than verified knowledge (OpenAI, 2023; Ji et al., 2023). Such behavior can undermine the reliability of outputs in fact-checking or educational contexts if not properly controlled.

## 2. Objectives

### 2.1. Main Objective

The main objective of this study is to analyze the impact of generative AI on the creation, dissemination, and mitigation of disinformation, identifying both the advantages and risks associated with the use of these technologies.

### 2.2. Secondary Objectives

The secondary objectives of this study include the following:

1. To examine how the scientific literature in social and human sciences addresses the interaction between generative AI and disinformation, highlighting optimistic and pessimistic perspectives.
2. To identify the mechanisms by which generative AI contributes to the spread of disinformation content and strategies to reduce these effects.
3. To propose lines of future research and recommendations for ethical and responsible use of generative AI in combating disinformation.

## 3. Methodology

This section is divided by subheadings. It should provide a reliable overview of results and their critical interpretation, as well as the preliminary conclusions based on previous research steps.

The scoping review is a systematic review methodology, the main objective of which is to analyze and synthesize the academic literature produced in a given scientific field. This type of review is especially useful for mapping the available evidence on a specific topic, especially when it is complex and varied (Codina, 2021).

One of the most notable characteristics of scoping reviews is their breadth. Unlike conventional systematic reviews, which usually focus on very specific research questions, scoping reviews are suitable for broad research questions that seek to explore the characteristics of an area of knowledge, rather than determine the efficacy of a specific intervention. This breadth provides an overview of the state of the art in a given field.

A scoping review generally involves several phases. First, the broad research question guiding the review is formulated. Next, a comprehensive search of the literature is carried out. The studies identified are selected using predefined inclusion and exclusion criteria, and the analysis framework is designed to systematically extract the data from the selected studies.

The methodology used in this research is shown in Table 1.

Before starting the scoping review, the most suitable database to use was agreed on, choosing Scopus, as it is widely recognized by the scientific community as one of the most prestigious and best-quality databases available. Using it also provides access to high-level research documents.

Once the database had been selected, the next step was to define the search equation. The first proposal used general keywords, like “artificial intelligence” and “disinformation”. However, it was decided to also include “ChatGPT”, as this technology has been the key catalyst for disruptive use of AI. In addition, keywords related to disinformation, such as “fake news” and “disinformation”, were also added, with the aim of covering the phenomenon more comprehensively.

**Table 1.** Application of the SALSA framework, adapted to that proposed by Lopezosa et al. (2023a, 2023b).

Application of the SALSA Framework	
Phase	Criteria
<b>Database:</b> Scopus. (see Supplementary Materials)	
<b>Search equation:</b>	
Search	<p>• “fake news” OR disinformation OR misinformation AND “Chat GPT” OR “artificial intelligence”.</p> <p><b>Years of publication:</b> 2021–2025 (10 November 2024 and updated to 21 December 2024).</p>
<p><b>Initial number of documents:</b> 224.</p> <p><b>Final number after applying the inclusion/exclusion criteria:</b> 64.</p>	
<b>Inclusion/exclusion criteria:</b>	
Evaluation	<ul style="list-style-type: none"> <li>- Omission of false positives.</li> <li>- Documents must be in English or Spanish and focus principally on ChatGPT and disinformation in the field of social sciences.</li> </ul>
Synthesis	<p>Using tables and narrative synthesis.</p> <p>For the analysis of these documents, a structured, systematic table that permits a heterogeneous synthesis is proposed.</p>
<b>Components:</b>	
Analysis	<ul style="list-style-type: none"> <li>• Object of study.</li> <li>• Objectives.</li> <li>• Questions/Problems/Hypothesis.</li> <li>• Methodology.</li> <li>• Results.</li> <li>• Labels (keywords of the article).</li> <li>• Category (central cluster or topic of the analyzed document).</li> </ul>

Next, the period of the sample for analysis was defined. The research team decided to include the years 2021 and 2025, with the aim of obtaining background from before November 2022, the month in which ChatGPT was launched, and thus assess where there was a significant increase in research on AI and disinformation.

Once the three above-mentioned aspects had been defined, the next step was to establish the inclusion and exclusion criteria for selecting documents. It was determined that the documents had to meet certain requirements to ensure the relevance of the sample; specifically, the keywords in the equation had to appear in the titles of the documents, thereby ensuring that the analyses centered on ChatGPT, generative AI, and disinformation. All duplicates, i.e., documents indexed in more than one database, were eliminated and those with errors or which could not be retrieved were rejected. These criteria ensured that the sample focused on the impact of generative AI, in particular, ChatGPT and disinformation.

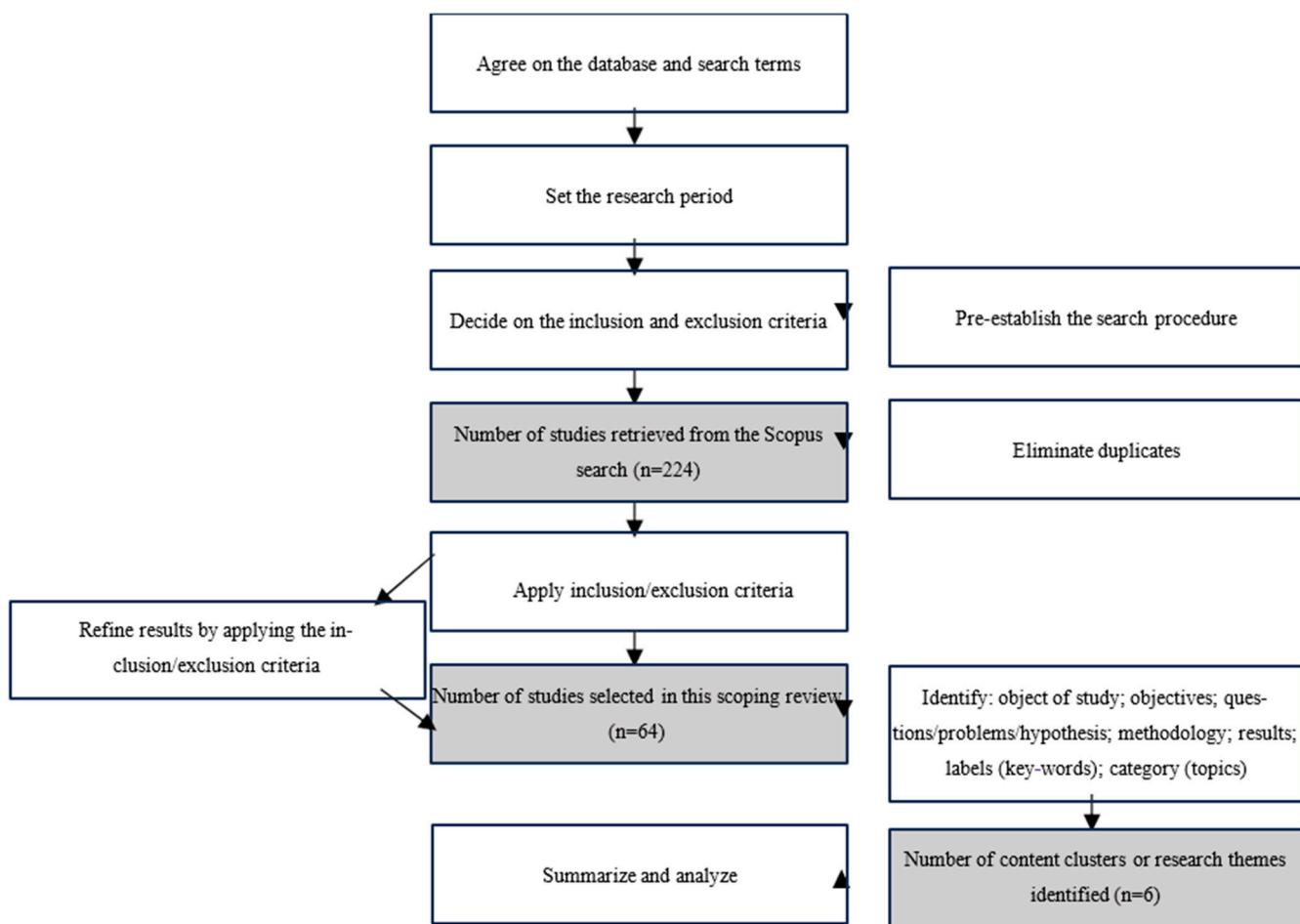
In general, only documents written in English and Spanish were included for methodological and practical reasons. English, as the dominant language in international scientific communication, ensures access to the main global contributions on artificial intelligence and disinformation. The inclusion of Spanish reflects both the research team’s linguistic competence—as Spanish is their native language—and their interest in incorporating studies that represent the realities and perspectives of the Spanish-speaking world.

For the definition and validation of the final categories and clusters, each selected article was carefully read and individually thematized. Based on this initial coding, the research team reached a consensus to group emerging topics into broader categories, aiming

to structure the results coherently and facilitate a thematic synthesis representative of the analyzed corpus.

Although a systematic and transparent protocol was applied, certain limitations are acknowledged throughout the study. The exclusion of gray literature may have led to the omission of some potentially relevant sector-specific studies on generative AI. Additionally, we recognize that only one database was used. However, these limitations do not compromise the validity of the study, as quality and reliability were prioritized by including only peer-reviewed literature and using Scopus, one of the most prestigious databases worldwide. Furthermore, triangulation within the research team helped reduce the risk of bias in the categorization and interpretation of the data.

The following flow diagram (see Figure 1) illustrates the process of the scoping review described above, along with the final part, to make the entire process of obtaining the results more visual.



**Figure 1.** Flow diagram illustrating the complete scoping review process.

## 4. Results and Discussion

### 4.1. Classification of the Articles by Topic and Perspective on the Impact of Generative AI

After choosing the articles, the next step was to classify them into the topics they cover and the perspective adopted in their approach. The thematic categories used to classify the 64 articles emerged inductively from a careful and repeated reading of the corpus, guided by the main objective of the study. While the process did not follow formal coding techniques typical of grounded theory, it was structured around recurring patterns in the literature—such as deepfakes, fact-checking, and scientific disinformation—that appeared consistently

across diverse contexts. These categories are not mutually exclusive, and in some cases, articles were assigned to multiple clusters. The scheme aims to provide a coherent and functional overview of the domains in which generative AI intersects with disinformation, while acknowledging its exploratory and interpretive basis. In addition to the thematic classification of articles, several recurring concerns and transversal insights emerged during the analysis. These are presented in Sections 4.4–4.10, which synthesize key patterns that cut across multiple categories and contribute to a more nuanced understanding of how generative AI interacts with disinformation.

#### 4.1.1. Political Disinformation and Propaganda

Political disinformation and propaganda refer to the use of false or deceitful information to influence public opinion and manipulate political processes. This sub-topic includes the analysis of how generative AI technologies, including deepfakes, can be used to create false content affecting elections, the perception of political leaders, and trust in democratic institutions. Thus, a number of articles specifically on deepfakes appear under this broad topic.

#### 4.1.2. Scientific Disinformation

This sub-topic addresses the impact of generative AI on the production and distribution of scientific knowledge, with particular attention to the dynamics of scholarly communication. It includes the use of generative AI to create research articles, abstracts, and other academic documents, and the risks associated with manipulating data and creating fake studies that could mislead the academic community.

#### 4.1.3. Fact-Checking

Fact-checking focuses on the use of generative AI to improve accuracy and efficiency in detecting fake news and verifying information. This sub-topic includes the analysis of how generative AI could automate the process, reduce human bias, and provide detailed answers to debunk false claims.

#### 4.1.4. Journalism and the Media

This sub-topic explores how generative AI is transforming journalism and the media, offering new opportunities for content creation, automatic translation, and recommending content. It also addresses the ethical and technical challenges associated with integrating generative AI into journalism and its implications for the quality and diversity of information.

#### 4.1.5. Media Literacy and Education

Media literacy and education refer to initiatives to improve verification skills and promote a culture of critical consumption of information. This subtopic includes the use of generative AI to develop advanced, personalized resources and strategies to educate the public on the risks of disinformation and foster consumption of more reliable information.

#### 4.1.6. Deepfakes

Deepfake technologies can create video and audio content that is almost indistinguishable from the real thing (deepfakes), with potentially negative consequences for public trust. This sub-topic includes the analysis of how deepfakes can be used to spread disinformation and manipulate public opinion, and emerging strategies to detect and combat this phenomenon.

All the articles were classified and grouped into these sub-topics. In some cases, two sub-topics were assigned to an article, when it clearly touched both. The results are shown in the following Table 2.

**Table 2.** Classification of articles by topic \*. Source: the authors.

Topic	Articles on the Topic	Number of Articles
Political disinformation and propaganda	4, 7, 13, 15, 16, 20, 24, 30, 31, 35, 38, 43, 45, 46, 48, 61	16
Scientific disinformation	3, 5, 10, 29, 34, 37, 39, 44, 49, 50	10
Fact-checking	1, 11, 21, 23, 27, 32, 36, 51, 52, 56, 57, 58, 59, 60, 64	15
Journalism and the media	2, 9, 17, 18, 41, 62	13
Media literacy and education	4, 8, 15, 25, 31, 33, 41, 42, 44	6
Deepfakes	4, 7, 13, 15, 16, 20, 24, 30, 31, 35, 38, 43, 45, 46, 48, 61	9

\* Note that the total of articles is 64 not 69, because 5 articles (4, 5, 31, 41, 44) were assigned to 2 topics.

Although we had some doubts about whether to maintain deepfakes as a standalone thematic category—since they refer primarily to a format rather than a content type—their growing relevance in public and academic debates, along with the specific challenges they pose in terms of audiovisual manipulation and narrative construction, justified treating them as an independent category rather than a transversal one. Therefore the focus is on the new problem the topic raises: the ease with which they can be made and how difficult they could become to detect, with implications for diplomatic relations ([Ikenga & Nwador, 2024](#)), election campaigns, and trust in the media ([Labuz & Nehring, 2024](#)).

These results show that, although the distribution is not symmetrical, there are publications on all the topics, particularly on political disinformation and propaganda, in fields as diverse as armed conflicts ([Cherry, 2024](#)), political strategy ([Diez-Gracia et al., 2024](#)), and propaganda ([Islas-Carmona et al., 2024](#)). Thus, there is a set of articles on the existence of disinformation and its effects on elections and on war, where there seems to be a particular interest.

While a considerable number of studies focus on the political drivers and implications of disinformation, particularly in electoral contexts and armed conflicts, other works have increasingly addressed scientific disinformation, especially during the COVID-19 pandemic, with relevant publications such as [Salaverría et al. \(2020\)](#) and [Masip et al. \(2020\)](#). These articles express a concern over the effects of generative AI in creating false content and flooding social media with such content on scientific topics ([Falyuna et al., 2024](#)).

It is possibly the combination of the three (media literacy, fact-checking, and media engagement), together with robust regulatory frameworks, that could offer the most effective strategy to counter disinformation, although this was not the subject of this study.

In this sense, it is worth stressing that AI is already being studied in relation to how it could improve verification processes ([Cuartielles et al., 2023](#)).

At the same time, depending on the view regarding how generative AI might impact the fight against disinformation, a distinction is made between positive and negative perspectives. The perspective on the impact of generative AI in the fight against disinformation is not conclusive, with both positive and negative perspectives being expressed. On the one hand, many articles provide an optimistic perspective on the ability of generative AI to combat disinformation. For example, [Moon and Kahlor \(2025\)](#) highlight how fact-checking with non-human sources could reduce bias in information, thereby improving its accuracy and efficiency. Similarly, [Gonçalves et al. \(2024\)](#) stress how AI can improve the efficiency and accuracy of fact-checking agencies, while [Lelo \(2024\)](#) examines automated fact-checking practices in Brazil, highlighting the associated advantages.

On the other hand, there are pessimistic perspectives on the risks associated with generative AI. Several authors, such as [Labuz and Nehring \(2024\)](#) and [Momeni \(2024\)](#), express this in relation to electoral campaigns and deepfakes, while [Kidd and Birhane \(2024\)](#) discuss how AI can distort human beliefs, conveying biases and false information to users.

This negative perspective underlines the need to develop advanced detection technologies and establish strong frameworks to eliminate the risks associated with generative AI.

Despite the risks, many researchers believe generative AI can be a powerful tool to combat disinformation if used adequately. Articles such as [Gutiérrez-Caneda and Vázquez-Herrero \(2024\)](#) propose strategies to improve detection and fact-checking with AI, while [Santos \(2023\)](#) analyzes automated disinformation detection using AI. Clearly, therefore, collaboration between different sectors, as proposed by [Penabad-Camacho et al. \(2024\)](#), is essential to ensuring generative AI tools are used responsibly and ethically.

In short, the perspective on the impact of generative AI in combating disinformation is complex and largely depends on how the technology is used. It is essential to maintain a balanced approach that combines optimism with pragmatism, recognizing both the opportunities and challenges associated with generative AI.

The next sections presents a number of the most significant trends and ideas that emerge from the selected articles.

#### 4.2. Deepfakes: The Cross-Cutting Challenge That Causes the Most Concern

Deepfakes represent one of the most worrying challenges in today's digital age, as the technology can create fake video and audio content that is practically indistinguishable from the genuine audiovisual material. Such technology could have devastating consequences for public trust and safety, as deepfakes can be used to spread disinformation, manipulate public opinion, and even extort individuals. [Labuz and Nehring \(2024\)](#) and [Momeni \(2024\)](#) explore how deepfakes can be used in election campaigns to create false scandals or distort reality, posing a threat to the integrity of democratic processes. [Weikmann et al. \(2024\)](#) analyze the psychological impact of deepfakes on individuals, highlighting their effect on how we see, listen to, and experience media.

To combat this problem, it is essential to develop advanced deepfake detection technologies and establish legal frameworks that target the creation and distribution of this type of content. [Gambín et al. \(2024\)](#) and [Doss et al. \(2024\)](#) propose strategies to ensure the accuracy and reliability of information, stressing the importance of collaboration between scientists, legislators, and technologists. It is also essential to educate the public on the associated risks and promote greater critical awareness of the consumption of digital content.

Despite the risks associated with deepfakes, some researchers believe that the technology could offer positive opportunities if used appropriately. For example, [Garriga et al. \(2024\)](#) advocate for media literacy as a key tool to combat deepfakes, highlighting the importance of educating the public on how to identify and avoid disinformation. Furthermore, collaboration between different sectors is essential for developing effective solutions. [Penabad-Camacho et al. \(2024\)](#) stress the need for ethical principles and transparency in the use of generative AI in scientific publication to ensure responsibility and maintain trust in the information.

What makes deepfakes particularly concerning is not only their capacity to deceive, but their corrosive effect on the broader epistemic environment. In a context where "seeing is believing" has traditionally grounded public trust, the very possibility of synthetic audiovisual content—regardless of its actual presence—introduces a kind of plausible deniability. Political figures can now discredit authentic recordings by merely invoking the specter of manipulation.

Furthermore, the potential weaponization of deepfakes goes beyond isolated instances of misinformation. State actors or powerful organizations with advanced AI capacities can manufacture doubt, confusion, or character attacks with minimal accountability, particularly in fragile democracies or conflict zones. As the sophistication of these techniques

evolves, the ethical and legal frameworks required to contain them remain fragmented and slow to adapt. In this sense, deepfakes do not merely represent a technological anomaly but a profound governance gap that challenges existing models of truth mediation and democratic oversight.

#### 4.3. Generative AI, New Tools, and Fact-Checking

Generative AI offers new and powerful tools for fact-checking, an increasingly essential task in a world flooded with disinformation. AI technologies can analyze large volumes of data in real time, identifying patterns and anomalies that might indicate the presence of fake news. This enables fact-checkers to work more efficiently and accurately, reducing the time needed to identify and debunk false information. [Moon and Kahlor \(2025\)](#) and [Gonçalves et al. \(2024\)](#) argue that generative AI improves accuracy and efficiency, automating the detection of fake news and providing detailed responses and quotes from reliable sources that debunk false claims.

One of the most promising applications of generative AI in fact-checking is its ability to generate automatic responses to false claims. Using advanced language models, generative AI can provide detailed explanations and quotes from reliable sources that debunk inaccurate claims. This not only speeds up the checking process, but also makes fact-checked information more accessible to the general public. [Lelo \(2024\)](#) and [Santos \(2023\)](#) analyze how AI can be used to improve fact-checking, highlighting the most promising technologies and their challenges.

However, the use of generative AI in fact-checking also poses a number of challenges. It is essential to ensure that AI models do not introduce their own biases or errors in the fact-checking process. Articles such as [Cuartielles et al. \(2023\)](#) discuss the importance of human supervision to ensure accuracy and impartiality in fact-checking. Furthermore, trust in AI for fact-checking requires human supervision, to ensure such accuracy and impartiality. With adequate implementation, generative AI could become a powerful tool to combat disinformation and promote more reliable information.

As generative AI becomes integrated into fact-checking ecosystems, a key shift is taking place: the transition from manual verification to AI-assisted pre-bunking and counter-narrative generation ([Linegar et al., 2024](#)). This evolution introduces new possibilities, but also new asymmetries. For example, the generation of corrective content risks being absorbed by the same attention economy that favors emotional appeal over accuracy. Moreover, the reliance on large language models for verification raises important ethical questions: How do we ensure that these systems do not reproduce biases or errors embedded in their training data? What mechanisms guarantee their transparency and accountability? As AI-generated fact-checks become more common, there is a growing risk that verification loses its human-centered deliberative dimension and becomes perceived as automatic, opaque, or contestable. Far from being a neutral enhancement, the integration of generative AI into fact-checking demands critical scrutiny of its authority and the sociotechnical imaginaries it embeds.

#### 4.4. Generative AI: Greater Capacity to Create Content as Well as Communicative Dysfunction in Science and Elections

Generative AI has greatly increased the capacity to create content but has also introduced new communicative dysfunctions, especially in the fields of science and elections. In science, AI can generate research articles, abstracts, and other academic documents with astonishing speed and accuracy. However, this capability can be exploited to create fake papers or manipulate data, which can mislead the scientific community and the general public. [Pecheranskyi et al. \(2024\)](#) and [Broinowski and Martin \(2024\)](#) analyze how gener-

ative AI can be used to create scientific content, highlighting both the benefits and risks associated with manipulating information.

In elections, generative AI can be used to create fake news, speeches, and other campaign materials that appear real but are designed to manipulate public opinion. This can have a profound impact on democracy, as voters may be influenced by such false information. Furthermore, consideration should also be given to AI's capacity for large-scale personalization of political messages, thereby increasing the complexity of disinformation in electoral contexts; as such, segmentation makes it more difficult to know who is reading what.

To address these dysfunctions, it is essential to develop fact-checking and detection tools that can identify AI-generated content and assess its accuracy. Along these lines, [Gonçalves et al. \(2024\)](#) and [Santos \(2023\)](#) discuss the importance of human supervision to ensure accuracy and impartiality in fact-checking. It is also important to establish regulations that enforce transparency in the use of AI in creating content, especially in sensitive contexts such as science and elections.

The disruptive potential of generative AI in science and electoral processes does not only lie in the proliferation of content, but in the altered communicative expectations it fosters. When content can be endlessly reproduced or simulated, the informational value of scientific or political discourse becomes harder to calibrate. In the case of science, this may result in a saturation of the communicative ecosystem with superficially plausible outputs, weakening the distinction between legitimate and illegitimate contributions. In electoral contexts, the credibility of candidates and institutions may be undermined not only through false claims, but by eroding the communicative norms that support public deliberation.

#### 4.5. Journalism and Media Under Pressure

Generative AI has introduced new dynamics in the field of journalism, but has also exacerbated pre-existing problems, such as labor precarity. In a context of shrinking advertising income and staff cuts, generative AI might seem an attractive solution for automating aspects of journalists' work, from writing news stories to creating multimedia content. However, the same technology could exacerbate precarity in the sector. [Forja-Pena et al. \(2024\)](#) and [Peña-Fernández et al. \(2023\)](#) discuss how reliance on AI can lead to even greater staff reductions, as media companies may see AI as a means of reducing labor costs.

This automation can have a negative impact on the quality of journalism, as AI, although efficient, cannot replace the intuition, experience, and critical judgement of human journalists. Furthermore, reliance on generative AI can lead to homogenization of content, with less diversity of viewpoints and deep analysis. [Subiela-Hernández et al. \(2023\)](#) and [Canavilhas et al. \(2024\)](#) advocate for maintaining a balance between use of the tool and human participation to ensure quality and diversity in journalism. Media companies must adopt a balanced approach, using generative AI to supplement, not replace, the work of human journalists.

To prevent generative AI from contributing to precarity in journalism, it is important to invest in life-long training for journalists, so they can work efficiently with new technologies. In this sense, [Sonni et al. \(2024\)](#) and [Moreno et al. \(2024\)](#) underline the need to promote sustainable business models and promote a culture of collaboration between journalists and generative AI technologies, ensuring these tools are used responsibly and ethically.

At the same time, there is a latent risk that AI-generated content becomes a substitute rather than a complement to journalistic work, especially in under-resourced newsrooms. As such, the impact of generative AI on journalism cannot be fully understood without considering power dynamics: who controls these tools, who sets their parameters, and who benefits from their deployment.

#### 4.6. Between Optimism and Pragmatism, Generative AI Is Already Here

Generative AI is already a reality and its presence is transforming many sectors of society. From content creation to medicine, AI is proving to be a powerful tool with enormous potential. Optimism regarding generative AI is based on its ability to improve efficiency, reduce costs, and open up new opportunities for innovation. For example, in the field of healthcare, AI can help diagnose diseases accurately and rapidly, while in education, it can personalize learning for each student. These applications could have a profound positive impact on people's quality of life (Pecheranskyi et al., 2024).

Nevertheless, it is important to maintain a pragmatic outlook and recognize the risks associated with generative AI. This includes the possibility of biases in the algorithms, loss of jobs, and the creation of false or misleading content. Kidd and Birhane (2024) and Wach et al. (2024) reflect on the risks associated with generative AI, highlighting the need to develop robust regulatory frameworks and promote transparency in the development and use of AI. In addition, it is essential to promote close collaboration between the public and private sectors to address these challenges.

Consequently, generative AI offers significant opportunities but also poses challenges that require a balanced approach. An appropriate approach involves a combination of optimism and pragmatism, recognizing both the benefits and associated risks. Articles such as Falyuna et al. (2024) and Penabad-Camacho et al. (2024) stress the importance of collaboration between scientists, legislators, and technologists to ensure good use of the tool. Only through joint effort can the full potential be achieved while minimizing the risks.

#### 4.7. Disinformation Is a Concern, Regardless of Who Has Created It

Disinformation is a growing concern in today's society, regardless of who creates it. From fake news to conspiracy theories, disinformation can have serious consequences for democracy, public health, and social cohesion. The ability of disinformation to spread rapidly in social media and other digital platforms makes it a particularly difficult challenge to combat. Salaverría et al. (2024) and Cherry (2024) highlight how disinformation can influence political decisions, weaken confidence in democratic institutions, and foster polarization in society.

One of the main concerns is its impact on public health. The COVID-19 pandemic already showed how the spread of fake news on vaccines and treatments could put the health of millions at risk. Furthermore, disinformation can have a negative impact on the economy, as companies can be affected by false rumors that destabilize markets. In their articles, Moreno Espinosa et al. (2024) and Aïmeur et al. (2024) reflect on how disinformation can harm the reputation of individuals and organizations, creating an atmosphere of distrust and uncertainty.

To combat disinformation, it is essential to develop effective strategies that include fact-checking, media education and collaboration between different sectors. Digital platforms must take steps to identify and remove false content, while the media must promote accurate and reliable information. It is also important to educate the public on how to identify disinformation and promote a culture of critical consumption of information. Thus, Bustos Díaz and Martin-Vicario (2024) and Garriga et al. (2024) stress the importance of media literacy to combat disinformation and promote more reliable information.

#### 4.8. Beyond Generative AI Tools, There Is Still a Lack of Legislation, Collaboration, and Pressure on Platforms and Social Media

Despite recent advances in generative AI tools and important regulatory steps such as the EU Artificial Intelligence Act (AIA) (European Union, 2024), there remains a significant lack of comprehensive and harmonized legislation globally to regulate their use effectively.

The speed at which these technologies evolve often outpaces the ability of legislators to establish effective legal frameworks. This creates a regulatory vacuum that can be exploited by malicious actors to spread disinformation, manipulate public opinion, and violate people's privacy. In this sense, Kreps et al. (2020) discuss the need to develop regulatory frameworks to reduce the risks associated with generative AI. In addition to legislation, it is essential for governments, technology companies, and civil society organizations to work together in addressing the challenges posed by generative AI.

Thus, pressure must be exerted on social media platforms to assume greater responsibility in combating disinformation. For their part, Aïmeur et al. (2024) and Cherry (2024) reflect on the need for digital platforms to improve their fact-checking mechanisms and the transparency of their algorithms, and collaborate with experts in disinformation to identify and delete fake content. Without decisive action by the platforms, the efforts to combat disinformation will be insufficient. Only joint effort can ensure generative AI tools are used responsibly and ethically.

#### 4.9. Social Media: A Hotspot of Endemic Disinformation

Social media has become a hotspot for endemic disinformation, in which fake news and conspiracy theories easily spread. The viral nature of these platforms allows disinformation to reach a mass audience in minutes, amplifying their impact and impeding their correction. Aïmeur et al. (2024) highlight how social media algorithms prioritize content that generates more interaction, regardless of its accuracy, thereby contributing to the spread of disinformation.

One of the main problems is that social media facilitate the creation of echo chambers, where users are only exposed to information that confirms their pre-existing beliefs. This can reinforce cognitive bias and hinder individuals' ability to critically assess the information they consume. Cherry (2024) and Kreps et al. (2020) describe how malicious actors can use bots and fake accounts to amplify disinformation and manipulate public opinion. This dynamic means disinformation spreads more quickly and is difficult to combat.

Combating the disinformation endemic on social media requires measures that promote a diversity of viewpoints and fact-checking. This could include promoting content from reliable sources, identifying and deleting fake accounts, and improving algorithms to reduce the visibility of disinformation. Bustos Díaz and Martin-Vicario (2024) and Garriga et al. (2024) stress the importance of educating users on how to identify and prevent disinformation, promoting a culture of critical consumption of information.

#### 4.10. The Importance of Transparency and Responsibility in the Use of Generative AI

Transparency and responsibility are key elements for guaranteeing generative AI is used appropriately. The lack of transparency in AI algorithms and decision-making processes can generate distrust and increase the risk of abuse. Wach et al. (2024) highlight the importance of transparency in the development and use of generative AI, stressing that technology companies must be open about how they develop and use their AI tools, including disclosing potential biases and limitations in their models.

In addition, responsible use of generative AI means establishing supervision mechanisms and accountability. Such measures could include setting up independent ethics committees to review AI applications and implementing internal policies that guarantee responsible use of the technology.

Finally, it is important to foster a culture of shared responsibility, in which the actors involved, including developers, users, and legislators, all work together. Falyuna et al. (2024) and Penabad-Camacho et al. (2024) stress the importance of collaboration between

scientists, legislators, and technologists to develop best practices and learn from errors. This is the only way to fully use the potential of generative AI while minimizing its risks.

## 5. Conclusions

This scoping review set out to analyze the impact of generative AI (particularly ChatGPT) on disinformation, with special attention to how it is being used to create, disseminate, and mitigate misleading content. The study also aimed to examine how the emerging academic literature conceptualizes this phenomenon, in order to outline potential directions for future research and public policy.

The analysis of 64 academic articles published between 2021 and 2024 reveals that generative AI is already profoundly transforming the disinformation ecosystem. On the one hand, it enables an unprecedented capacity to generate synthetic text, images, audio, and video that are increasingly difficult to distinguish from authentic content. This raises significant risks for democratic processes, scientific credibility, and public trust, particularly when these tools are used to manipulate, fabricate, or strategically distort information.

Generative AI also facilitates the dissemination of disinformation by making it more targeted, personalized, and scalable. The combination of synthetic content and online platform recommendation algorithms amplifies the reach of false narratives—often beyond the control of traditional oversight mechanisms. This dynamic challenges existing models of media regulation and raises urgent questions around accountability, traceability, and platform governance.

At the same time, several studies highlight the potential of generative AI as a resource for fact-checking and education. Some fact-checking organizations are already experimenting with AI-assisted tools that help detect false claims, generate counter-narratives, or automate verification processes. Likewise, some educational initiatives have begun to incorporate generative AI as a means to promote critical thinking and to improve understanding of information manipulation dynamics.

One of the most relevant findings of this review is the ambivalent nature of generative AI. It is neither inherently beneficial nor harmful; its impact depends on the context in which it is used, the intentions behind its deployment, and the regulatory frameworks that govern it. This dual potential is clearly reflected in the literature analyzed: while some authors emphasize its opportunities for enhancing informational quality and institutional resilience, others warn of the risks associated with deepfakes, epistemic instability, and the erosion of public trust.

The review also emphasizes the importance of collaboration across different sectors to combat disinformation. Governments, technology companies, and civil society organizations must work together to develop common standards and regulations, as well as share resources and knowledge. Another important aspect identified is the need to educate the public about the risks associated with disinformation and to promote a culture of critical information consumption. This includes fostering media literacy and training in verification skills. Digital platforms must also take steps to identify and remove false content, as well as improve algorithms to reduce the visibility of disinformation.

As a result, generative AI offers new opportunities for fact-checking—an existing trend—but it also poses challenges that must be addressed. Key steps in this process include the development of more advanced detection and verification technologies, the establishment of solid regulatory frameworks, and the promotion of cross-sector collaboration to ensure the appropriate use of generative AI.

Moreover, it is confirmed that generative AI enables greater content creation capacity but also introduces significant informational and communicative dysfunctions in scientific and electoral contexts. In this sense, a balanced approach is proposed—one that combines

technological innovation with regulation. Only through joint efforts and a comprehensive approach can the negative impact of this technology be prevented while harnessing its potential benefits.

Beyond the thematic categories examined, several cross-cutting concerns have emerged. These include the vulnerability of journalism in an increasingly automated media environment, the role of social media as a central hub for disinformation, and the psychological and cognitive effects of synthetic content. These dynamics reinforce the idea that generative AI is not merely a technological development but a sociopolitical challenge that requires a holistic response.

The review also highlights the need for cross-sector collaboration in addressing the challenges of disinformation. Governments, technology companies, academic institutions, and civil society organizations must work together to develop shared standards, coordinate regulatory efforts, and foster innovation in detection and response tools. Isolated or fragmented efforts are unlikely to succeed in the face of such a dynamic and globalized information environment.

Equally essential is the task of strengthening public resilience through education and training. Promoting a culture of critical information consumption requires investing in media literacy, verification skills, and the integration of these competences into both formal and informal education systems. Digital platforms must also take responsibility for identifying and removing false content, as well as adjusting their algorithms to reduce the visibility and virality of disinformation.

Finally, the literature suggests the need for a balanced and comprehensive approach. While generative AI presents clear opportunities for innovation in communication, research, and education, it also introduces major communicative dysfunctions, particularly in scientific and electoral domains. Addressing these tensions requires regulatory frameworks that safeguard informational integrity while supporting innovation. Only through coordinated and sustained action will it be possible to harness the benefits of generative AI while mitigating its risks in the disinformation landscape.

Social networks without AI-generated content identifiers, without automated and manual monitoring systems for fake content, and traditional media without empowerment and control capabilities will allow harmful effects to emerge. Therefore, it will be necessary to find a balance to avoid ungovernable situations, as happened in the past with intellectual property and the advent of the internet.

This study presents a series of limitations that should be considered when interpreting the results. First, the scoping review focused primarily on academic literature published in English and Spanish, which may have excluded important studies in other languages. Furthermore, the search was limited to the Scopus database, which may have reduced the scope of the results. For this reason, we believe it may be important in the future to develop new research that includes more languages.

Another limitation is that the review focused on studies published between 2021 and 2024, which may have excluded earlier works that could offer a broader perspective. Additionally, the nature of the review means that the results are preliminary, and further research is needed to confirm the findings and examine in more depth the aspects identified—both chronologically and in terms of the use of search equations that include other AI- and disinformation-related terms and concepts.

As for future research directions, an interesting avenue would be to study the impact of generative AI on disinformation in other fields, such as the third sector, health, or education. It would also be relevant to investigate how to develop more accurate detection and verification technologies to identify AI-generated content. Additionally, it is necessary

to study how to establish more effective regulatory frameworks that ensure transparency and accountability in the use of generative AI.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/publications13030033/s1>, Table Scoping review dataset.xlsx.

**Author Contributions:** Conceptualization, methodology, investigation, data curation, A.L.-B. and C.L.; writing—original draft preparation, A.L.-B.; writing—review and editing, C.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** No new data were created or analyzed in this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Aimeur, E., Amri, S., & Brassard, G. (2024). Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining*, 13(1), 30. [\[CrossRef\]](#)
- Almasri, F. (2024). Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research. *Research in Science Education*, 54(5), 977–997. [\[CrossRef\]](#)
- Bianchini, S., Müller, M., & Pelletier, P. (2022). Artificial intelligence in science: An emerging general method of invention. *Research Policy*, 51(10), 104604. [\[CrossRef\]](#)
- Bircan, T., & Salah, A. A. A. (2022). A bibliometric analysis of the use of artificial intelligence technologies for social sciences. *Mathematics*, 10(23), 4398. [\[CrossRef\]](#)
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Sin Wang Chong, S. M., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21(1), 4. [\[CrossRef\]](#)
- Broinowski, A., & Martin, F. R. (2024). Beyond the deepfake problem: Benefits, risks and regulation of generative AI screen technologies. *Media International Australia*. [\[CrossRef\]](#)
- Bustos Díaz, J., & Martin-Vicario, L. (2024). Media literacy in a hyperconnected world: From social networks to Artificial Intelligence. *European Public and Social Innovation Review*, 9, 1–17. [\[CrossRef\]](#)
- Canavilhas, J., Ioscote, F., & Gonçalves, A. (2024). Artificial Intelligence as an opportunity for journalism: Insights from the Brazilian and Portuguese media. *Social Sciences*, 13(11), 590. [\[CrossRef\]](#)
- Chan, C. K., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, 43. [\[CrossRef\]](#)
- Cherry, S. (2024). Modern armed conflicts: Disinformation campaigns shaping the digital information landscape. *Serials Librarian*, 85(1–4), 19–31. [\[CrossRef\]](#)
- CIPR (Chartered Institute of Public Relations). (2023). *Artificial Intelligence (AI) tools and the impact on public relations (PR) practice*. Available online: [https://cipr.co.uk/CIPR/Our\\_work/Policy/AI\\_in\\_PR/\\_AI\\_in\\_PR\\_guides.aspx](https://cipr.co.uk/CIPR/Our_work/Policy/AI_in_PR/_AI_in_PR_guides.aspx) (accessed on 15 November 2024).
- Codina, L. (2021). *Scoping reviews: Características, frameworks principales y uso en trabajos académicos*. Available online: <https://www.lluiscodina.com/scoping-reviews-guia> (accessed on 15 November 2024).
- Correas, M., & Kenneally, E. (2019). *Fake news. Bulos que cambiaron la historia*. Titilante Ediciones. ISBN 9788494967030.
- Cuartielles, R., Ramon-Vegas, X., & Pont-Sorribes, C. (2023). Retraining fact-checkers: The emergence of ChatGPT in information verification. *Profesional de la Información*, 32(5). [\[CrossRef\]](#)
- Diez-Gracia, A., Sánchez-García, P., & Martín-Román, J. (2024). Disintermediation and disinformation as a political strategy: Use of AI to analyze fake news as Trump's rhetorical resource on Twitter. *Profesional de la Informacion*, 32(5), e320523. [\[CrossRef\]](#)
- Doss, C., Mondschein, J., Shu, D., Wolfson, T., Kopecky, D., Fitton-Kane, V. A., Bush, L., & Tucker, C. (2024). Deepfakes and scientific knowledge dissemination. *Scientific Reports*, 13, 13429. [\[CrossRef\]](#) [\[PubMed\]](#)
- European Commission. (2018). *Tackling online disinformation: A European approach. Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions*. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0236> (accessed on 15 November 2024).
- European Union. (2024). Regulation (EU) 2024/1689 of the European parliament and of the council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence act) and amending regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013 and (EU) 2018/858. *Official Journal of the European Union*, 1689, 1–157. Available online: [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689) (accessed on 15 November 2024).

- Falyuna, N., Feher, K., Demeter, M., Szüdi, G., & Uscinski, J. E. (2024). Against disinformation: Bridging science and public discourse. *Információ Társadalom*, 24(2), 68–84. [CrossRef]
- Forja-Pena, T., García-Orosa, B., & López-García, X. (2024). The ethical revolution: Challenges and reflections in the face of the integration of artificial intelligence in digital journalism. *Communication and Society*, 37, 237–254. [CrossRef]
- Gambín, A., Yazidi, A., Vasilakos, A., Haugerud, H., & Djenouri, Y. (2024). Deepfakes: Current and future trends. *Artificial Intelligence Review*, 57, 64. [CrossRef]
- Garriga, M., Ruiz-Incertis, R., & Magallón-Rosa, R. (2024). Artificial intelligence, disinformation and media literacy proposals around deepfakes. *Observatorio*, 18(5). [CrossRef]
- Gonçalves, A., Torre, L., Oliveira, F., & Jerónimo, P. (2024). AI and automation's role in Iberian fact-checking agencies. *Profesional de la Información*, 33(2). [CrossRef]
- González, M. S., Sánchez Gonzales, H. M., & Gonzalo, S. M. (2022). Artificial intelligence in Hispanic verifiers of the IFCN network: Innovative projects and perception of experts and professionals. *Estudios Sobre el Mensaje Periodístico*, 28(4), 867–879. [CrossRef]
- Gutiérrez-Caneda, B., & Vázquez-Herrero, J. (2024). Redrawing the lines against disinformation: How AI is shaping the present and future of fact-checking. *Tripodos*, 55, 55–74. [CrossRef]
- Hosseini, M., Horbach, S. P. J. M., Holmes, K., & Ross-Hellauer, T. (2024). Open science at the generative AI turn: An exploratory analysis of challenges and opportunities. *Quantitative Science Studies*, 6, 22–45. [CrossRef] [PubMed]
- Ikenga, F. A., & Nwador, A. F. (2024). The intersection of artificial intelligence, deepfake, and the politics of international diplomacy. *Ianna Journal of Interdisciplinary Studies*. [CrossRef]
- Islas-Carmona, J. O., Gutiérrez-Cortés, F. I., & Arribas-Urrutia, A. (2024). Disinformation and political propaganda: An exploration of the risks of artificial intelligence. *Explorations in Media Ecology*, 23(2), 105–120. [CrossRef]
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys (CSUR)*, 55(12), 1–38. [CrossRef]
- Kharroub, T. (2023). *Disinformation and hate speech on social media contribute to inciting war crimes against Gaza*. Arab Center Washington DC. Available online: <https://arabcenterdc.org/resource/disinformation-and-hate-speech-on-social-media-contribute-to-inciting-war-crimes-against-gaza> (accessed on 15 June 2025).
- Kidd, C., & Birhane, A. (2024). How AI can distort human beliefs: Models can convey biases and false information to users. *Science*, 380(6651), 1222–1223. [CrossRef] [PubMed]
- Kreps, S., McCain, R. M., & Brundage, M. (2020). All the news that's fit to fabricate: AI-generated text as a tool of media misinformation. *Journal of Experimental Political Science*, 9(1), 104–117. [CrossRef]
- Labuz, M., & Nehring, C. (2024). On the way to deep fake democracy? Deep fakes in election campaigns in 2023. *European Political Science*, 23, 454–473. [CrossRef]
- Lelo, T. (2024). Fostering Artificial Intelligence to face misinformation: Discourses and practices of automated fact-Checking in Brazil. *Journalism and Mass Communication Quarterly*, 101(2), 320–345. [CrossRef]
- Linegar, M., Sinclair, B., van der Linden, S., & Alvarez, R. M. (2024). Towards generalizable AI-assisted misinformation inoculation: Protecting confidence against false election narratives. *arXiv*, arXiv:2410.19202. [CrossRef]
- Lopezosa, C., Codina, L., & Rovira, C. (2023a). Investigant la SEO des de les ciències socials: "scoping review". *BiD: Textos Universitaris de Biblioteconomia i Documentació*, 51. [CrossRef]
- Lopezosa, C., Guallar, J., Codina, L., & Pérez-Montoro, M. (2023b). Curación de contenido y periodismo: Revisión sistematizada exploratoria y visión experta. *Revista Mediterránea de Comunicación/Mediterranean Journal of Communication*, 14(1), 205–223. [CrossRef]
- Masip, P., Aran-Ramspott, S., Ruiz-Caballero, C., Suau, J., Almenar, E., & Puertas-Graell, D. (2020). Consumo informativo y cobertura mediática durante el confinamiento por el COVID-19: Sobreinformación, sesgo ideológico y sensacionalismo. *Profesional de la Información*, 29(3), e290312. [CrossRef]
- McNair, B. (2018). *Fake news: Falsehood, fabrication and fantasy in journalism*. Routledge. ISBN 9781138306790.
- Momeni, M. (2024). Artificial Intelligence and political deepfakes: Shaping citizen perceptions through misinformation. *Journal of Creative Communications*, 20(1), 41–56. [CrossRef]
- Moon, W., & Kahlor, L. A. (2025). Fact-checking in the age of AI: Reducing biases with non-human information sources. *Technology in Society*, 80, 102760. [CrossRef]
- Moreno, Á. P., Manche, W. V. S., & Artigas, W. (2024). Artificial Intelligence in journalism. Bibliometric review in scopus (1989–2022). *Anuario Electrónico de Estudios en Comunicación Social Disertaciones*, 17(2). [CrossRef]
- Moreno Espinosa, P., Abdulsalam Alsarayreh, R. A., & Figueiro-Benítez, J. C. (2024). Big data and artificial intelligence as solutions to disinformation. *Doxa Comunicacion*, 38. [CrossRef]
- OpenAI. (2023). *GPT-4 technical report*. Available online: <https://openai.com/research/gpt-4> (accessed on 30 June 2025).

- Pecheranskyi, I., Oliinyk, O., Medvedieva, A., Danyliuk, V., & Hubernator, O. (2024). Perspectives of generative AI in the context of digital transformation of society, audio-visual media and mass communication: Instrumentalism, ethics and freedom. *Indian Journal of Information Sources and Services*, 14(4), 48–53. [CrossRef]
- Penabad-Camacho, L., Penabad-Camacho, M. A., Mora-Campos, A., Cerdas-Vega, G., Morales-López, Y., Ulate-Segura, M., Méndez-Solano, A., Nova-Bustos, N., Vega-Solano, M. F., & Castro-Solano, M. M. (2024). Heredia declaration: Principles on the use of Artificial Intelligence in scientific publishing. *Revista Electronica Educare*, 28(S), 1–10. [CrossRef]
- Peña-Fernández, S., Peña-Alonso, U., & Eizmendi-Iraola, M. (2023). The discourse of journalists on the impact of generative artificial intelligence on disinformation. *Estudios Sobre el Mensaje Periodístico*, 29(4), 833–841. [CrossRef]
- Pierri, F., Luceri, L., Jindal, N., & Ferrara, E. (2023, April 30–May 1). Propaganda and misinformation on facebook and twitter during the Russian invasion of Ukraine. 15th ACM Web Science Conference 2023 (WebSci '23) (pp. 65–74), Austin, TX, USA. [CrossRef]
- Rochlin, N. (2017). Fake news: Belief in post-truth. *Library Hi Tech*, 35(3), 386–392. [CrossRef]
- Salas Abad, C. (2019). La primera fake news de la historia. *Historia y Comunicación Social*, 24(2), 411–431. [CrossRef]
- Salaverría, R., Bachmann, I., & Magallón-Rosa, R. (2024). Desinformación y confianza en los medios: Propuestas de actuación. *Comunicación*, 14(2), 13–32. [CrossRef]
- 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. [CrossRef]
- Santos, F. (2023). Artificial Intelligence in automated detection of disinformation: A thematic analysis. *Journalism and Media*, 4, 679–687. [CrossRef]
- Sonni, A. F., Putri, V. C. C., & Irwanto, I. (2024). Bibliometric and content analysis of the scientific work on Artificial Intelligence in journalism. *Journalism and Media*, 5(2), 787–798. [CrossRef]
- Subiela-Hernández, B. J., Gómez-Company, A., & Vizcaíno-Laorga, R. (2023). Challenges and opportunities in the fight against misinformation and copyright in Journalism: MediaVerse (AI, blockchain and smart contracts). *Estudios Sobre el Mensaje Periodístico*, 29(4), 869–880. [CrossRef]
- Tomassi, A., Falegnami, A., & Romano, E. (2025). Disinformation in the digital age: Climate change, media dynamics, and strategies for resilience. *Publications*, 13(2), 24. [CrossRef]
- Uribe, A., & Machin, J. D. (2024). Past, present and future of information literacy in Latin America. *Journal of Information Literacy*, 18(2), 6–36. [CrossRef]
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. [CrossRef] [PubMed]
- Wach, K., Duong, C. D., Ejdys, J., Kazlauskaité, R., Korzynski, P., Mazurek, G., Palisziewicz, J., & Ziembra, E. (2024). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, 11(2), 7–30. [CrossRef]
- Wardle, C. (2020). Understanding information disorder. Available online: <https://firstdraftnews.org/long-form-article/understanding-information-disorder> (accessed on 15 November 2024).
- Weikmann, T., Greber, H., & Nikolaou, A. (2024). After deception: How falling for a deepfake affects the way we see, hear, and experience media. *International Journal of Press/Politics*, 30(1), 187–210. [CrossRef]
- World Health Organization. (2020). Novel coronavirus (2019-nCoV). Situation report—13. Available online: <https://www.who.int/docs/default-source/coronavirus/situation-reports/20200202-sitrep-13-ncov-v3.pdf> (accessed on 15 November 2024).
- Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C., Qiu, J., Hua, K., Su, W., Wu, J., Xu, H., Han, Y., Fu, C., Yin, Z., Liu, M., ... Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4), 100179. [CrossRef] [PubMed]
- Zhou, H. (2024, November 5). The top ten challenges, needs, and goals of publishers—And how AI can help in digital transformation and the open science movement. The Scholarly Kitchen. Available online: <https://scholarlykitchen.sspnet.org/2024/11/05/the-top-ten-challenges-needs-and-goals-of-publishers-and-how-ai-can-help-in-digital-transformation-and-the-open-science-movement> (accessed on 15 November 2024).

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.