

# Political communication in the age of artificial intelligence: an overview of deepfakes and their implications

SOCIETY REGISTER  
2024 / 8(2): 7–24  
ISSN: 2544–5502  
DOI: 10.14746/sr.2024.8.2.01



Daniele Battista<sup>1</sup>

<sup>1</sup> University of Salerno, Department of Business Sciences—Management & Innovation Systems, Via Giovanni Paolo II, 132, 84084 Fisciano, Italy. ORCID: 0009-0005-8418-8374, Email: dbattista@unisa.it

**ABSTRACT:** Rapid and significant advances in artificial intelligence (AI) have profoundly shaped the current technological landscape, sparking extraordinary interest in academia and beyond. Considered an unprecedented revolutionary technology, it has captured the attention of researchers, scholars, and professionals from various disciplines as it offers transformative prospects in a wide range of fields. This technological progress has sparked a broad debate on its social, ethical, and economic impacts, raising important questions that require in-depth and multidisciplinary investigations. AI thus emerges as an ever-evolving discipline with significant implications for the future of human progress. Besides being seen as an opportunity to pursue common societal goals, many observers have recognized the potential risks associated with such developments. Its integration into the political context also presents a promising opportunity to enhance the efficiency of political decisions. However, its adoption raises significant challenges that require careful evaluation. This research aims to explore in detail the relationship between AI and political communication, focusing on analysing AI's usage in this context while highlighting the phenomenon of deepfakes, which jeopardize democratic stability and security in many cases. The importance of this research contribution lies in the context where AI is assuming an increasingly prominent role in communication dynamics, making it essential to fully understand the implications and potential consequences of AI application in the political field. Furthermore, it is crucial to assess whether such initiatives can genuinely be considered democratic or if they could represent a dangerous trend towards the use of manipulative algorithms.

**KEYWORDS:** artificial intelligence, political communication, ethical impact, political innovation, deepfake



---

## INTRODUCTION

In the context of contemporary political communication, the growing intersection of digital technology and politics has sparked increasing interest among social science scholars (De Blasio, 2018; Ziccardi, 2019). This convergence phenomenon (Jenkins, 2014) has given rise to new interaction dynamics among political actors, media, and citizens, fundamentally reshaping how politics is conceived, disseminated, and consumed. Notably, the advent of Artificial Intelligence (AI) represents one of the most significant milestones in this evolution (Floridi, 2022). It is no secret that, with its ability to analyze complex data, process real-time information, and adapt to individual preferences, AI redefines how political leaders engage with the public and substantially contributes to shaping public opinion (Crawford, 2022). In this context, this research aims to comprehensively explore the crucial role played by AI in contemporary political communication, analyzing how it influences and modifies the dynamics of democratic participation and the formation of public opinion.

Throughout this study, we will examine the impacts, challenges, and opportunities that AI presents for the future of political communication and democracy itself. In response to a media and cultural discourse often characterized by a dystopian and pessimistic perspective regarding the advancement and development of artificial intelligence, this approach seeks to transcend the perception of AI as a sole existential threat, highlighting instead the potential positive implications and opportunities it offers in the political context. When analyzing the interaction between Artificial Intelligence (AI) and politics, the need for a thoughtful and informative reflection emerges, encompassing both the ethical and social challenges associated with this convergence and the potential innovation opportunities. This perspective underscores the relevance of conducting a balanced and in-depth assessment of AI's impact on the political sphere, intending to enrich the understanding of this dynamic in all its nuances. One of the crucial dimensions of this assessment concerns the potential enhancement of democratic participation through the use of AI. AI can facilitate political mobilization, streamline access to relevant information for citizens, and promote more inclusive and informed involvement in the decision-making process. However, it is essential to emphasize that such evaluation must be comprehensive and multifaceted, as AI poses significant risks to democratic functioning.

The concentration of power in the hands of opaque algorithms, the potential for manipulating public opinion through the dissemination of false or polarizing information, and the creation of online echo chambers are just some critical aspects to consider. In this context, it is crucial to carefully examine cases where AI could undermine democratic principles, aiming to mitigate adverse effects and promote responsible and ethical use of AI in the political sphere. In conclusion, the analysis of AI's impact on politics requires a balanced approach that takes into account both potential innovations and the importance of preventing harm to democratic functioning. This multidimensional approach

---

is essential for developing a comprehensive understanding of the complex interaction between AI and politics and guiding political decisions in an informed and responsible manner.

### **THEORETICAL FRAMEWORK: CONTEXTUALIZING THE INTERSECTION OF AI AND CONTEMPORARY POLITICAL COMMUNICATION**

Popular Culture, including media forms such as cinema, television, and traditional media, has presented an interpretation of artificial intelligence characterized by a focus on the competition between machines and humans (Lepri et al., 2021; Vogt, 2021). This representation has ingrained ideas, stereotypes, and expectations regarding potential threats and dangers associated with the idea of creating machines capable of exhibiting intelligent behaviours. This narrative has evolved from influences such as the Turing Test, a criterion used to assess a machine's ability to simulate human intelligence, and it has been found in films like "Terminator" (Cameron & Hurd, 1984). This perception has even led some scholars to consider AI as a potentially existential threat (Bostrom, 2014), in stark contrast to those who view the advent of a new era of innovation in AI and its capacity to reshape society (Gentili et al., 2020).

Undeniably, with the progress of artificial intelligence (AI), the landscape of communication, especially in the realm of politics, faces a series of complex dilemmas related to information and news ecosystems. The rapid evolution of AI has opened various perspectives and intriguing issues concerning the dissemination and consumption of political information. In this context, we are confronted with challenges related to source verification, information manipulation, equitable access to news, and creating a trustworthy information environment. These developments necessitate a deep reflection on AI's ability to shape and influence the political communication process and how effective strategies can be adopted to ensure an informed and transparent political discourse in the AI era (Bareis & Katzenbach, 2022). Discourse is often condensed within a social whirlwind, where various techniques are employed to navigate and emerge (Battista, 2023a). During academic discussions, complex methodological and thematic challenges have been addressed and analyzed, including the impact on public opinion (Marmolejo-Ramos et al., 2022; Groves et al., 2023). It has been highlighted, however, that technologies related to artificial intelligence, while on the one hand, can bring improvements to human understanding and social dynamics, on the other hand, they can also introduce elements of distortion, manipulation, and misinformation (Kertysova, 2018; Bontridder & Poulet, 2021; Gallo *et al.* 2022).

This duality in the effect of AI technologies on cognition and the social sphere represents a crucial aspect to examine and understand in the context of contemporary communication challenges. However, it is undeniable to declare that due to the momentum generated by rapid technological advancements and growing public interest, Artificial

Intelligence (AI) is commonly perceived as an unprecedented revolutionary technology (Brock & Von Wangenheim, 2019). It is relevant to note that, despite originating as an academic field in the early 1950s, Artificial Intelligence (AI) remained a relatively obscure scientific domain with limited practical interest for over fifty years (Haenlein & Kaplan, 2019; Wirtz & Müller, 2019). Only recently, owing to the ongoing developments, a significant shift has brought AI out of its traditional realm within computer science, extending its influence into disciplinary areas such as the social sciences (Rosete et al., 2020).

From a broad perspective, within the specialized literature, AI is described as the capability to perform behaviours and processes typically associated with the human sphere through the use of machines, computer systems, or computational networks (Li et al., 2021). This definition emphasizes the ability of these artificial entities to emulate, replicate, or simulate a wide range of cognitive, decision-making, and adaptive activities previously considered the exclusive domain of human intelligence. This concept, extensively debated and explored in scientific literature, highlights the progressive advancement and expansion of AI as a multidisciplinary and innovative discipline that extends beyond the traditional boundaries of computer science to encompass a variety of application domains. It is not coincidental that, as stated by Russell and Norvig, Artificial Intelligence (AI) can encompass both thought and reasoning processes and a behavioural component similar to that of humans. Therefore, a comprehensive definition of AI should incorporate two key aspects: on the one hand, a more human-centered approach that embraces empirical sciences to understand human behaviour; on the other hand, a focus on the creation of mathematical and engineering models that enable the artificial realization of such behaviours. This perspective acknowledges the complexity of AI, which requires a synergy between the human-centered approach and the engineering and mathematical modeling orientation. Additionally, some scholars hypothesize that artificial intelligence has undergone three distinct stages of research development (Darlington, 2020; Sharma et al., 2020). Technological advances and increased computer processing power characterize the first phase. The second phase focuses on promoting and developing artificial neural networks, primarily aiming to emulate the functioning of the human brain. The third phase, in contrast, has been marked by the emergence of deep learning and the practical application of such approaches in real-world scenarios. This perspective suggests a historical progression in AI research, highlighting the evolution and increasing complexity of methodologies and applications over time. Understanding its evolution and application in different contexts is crucial to outlining a clear framework.

Furthermore, AI is increasingly influencing the political arena across various scenarios. For example, AI-based data analytics systems are used to collect, process, and interpret vast amounts of information from online platforms and social media, enabling politicians to understand voter opinions and preferences better (Battista & Uva, 2023). These tools can be used to guide communication strategies and election campaigns in a more targeted manner (Alvarez et al., 2023). Additionally, AI can play a crucial role in automating

---

decision-making processes within the government, contributing to improved efficiency and reducing human error (Reis et al., 2021). However, this raises important questions regarding the responsibility and transparency of decisions made by algorithms, as they directly impact citizens' lives. Political participation can also be influenced by the spread of false or manipulated news generated by AI-based text generation systems (Klinger et al., 2023). Therefore, political actors, governments, and stakeholders must fully understand the impact of AI on political participation, address the ethical and legal challenges associated with it, and develop appropriate policies to ensure that AI is used responsibly and for the public good, all while aligning these efforts with a pathway aimed at educating citizens about the use of AI to promote informed and conscious participation.

This research aims to conduct a literature review to thoroughly examine existing research and theories concerning AI in political communication. This critical analysis aims to provide a clear and detailed overview of the current state of knowledge in this interdisciplinary field, highlighting key findings, predominant theories, and gaps that require further investigation. The review contribution is fundamental for contextualizing the theoretical and methodological framework within which this research falls, identifying challenges, opportunities, and future directions in understanding the interaction between political communication and AI and its implications in the broader political landscape. A comprehensive analysis of past research and developed theories aims to lay the groundwork for understanding and guiding future research endeavours. The approach employed will allow us to identify the strengths of past research as well as the gaps that can be addressed through this investigation. We will begin by examining its integration into political communication, considering technological innovations and specific applications that have driven its growing relevance. It will be essential to explore how AI has contributed to reshaping campaign dynamics, formulating public policies, and interacting with politicians and citizens. Furthermore, the redefinition of international political dynamics, with some focus on contemporary issues, will place the current role of artificial intelligence (AI) in politics and its developments at the forefront.

## METHODOLOGY

This study, aiming to conduct an in-depth theoretical analysis of the deepfake phenomenon to identify and evaluate available countermeasures, represents a significant step in research. This field holds particular interest within the communication landscape, not only for its social implications but also for the significant political ramifications it entails. The theoretical approach adopted in this work allows for a critical examination of the various multifaceted dimensions related to artificial intelligence and deepfakes, laying the foundation for a greater understanding of how these phenomena can influence the public sphere. The social relevance of this study lies in its capacity to promote greater awareness among the public and communication practitioners regarding the challenges

posed by deepfakes, enabling more informed participation in political and social debates. Simultaneously, the political significance of this work is evident in its ability to highlight the democratic implications related to the manipulative use of media content, thus making a significant contribution to the public discourse on regulation and accountability in the digital era. Given the current landscape of scientific research, characterized by fragmentation and the relative novelty of investigations in the social sciences regarding this topic, this study serves as an innovative theoretical and prospective contribution. The originality of this research lies in its ability to combine various dimensions, such as Artificial Intelligence, the creation of manipulated audiovisual content, and the perspective of understanding and anticipating the phenomenon in the short and medium term. This research aims to fill an existing gap in the current body of knowledge, providing a comprehensive and in-depth analytical framework of the interconnections between these complex themes.

Furthermore, it seeks to lay the groundwork for a better understanding of potential future developments in this field, thus offering an explicit and informed perspective on addressing the deepfake phenomenon. Its relevance in the academic realm lies in its ability to advance the construction of a solid theoretical framework that can serve as a basis for further research and insights in this emerging field. At the same time, the short- and medium-term forecasting perspective will help prepare society and institutions for future challenges related to deepfakes, providing a basis for developing appropriate policies and strategies to address this evolving phenomenon. A diachronic research methodology has been adopted to achieve this goal, structured in two phases. Initially, a longitudinal bibliographic investigation was conducted on the deepfake phenomenon. It is a relatively unexplored and sparsely addressed field within academic publications, especially in the social sciences and, more specifically, in communication.

The novelty and complexity of this phenomenon required a descriptive, historical, and systematic approach to provide a comprehensive overview. The primary purpose of this first research phase was to share the essential characteristics of the phenomenon with the scientific community, providing a framework that synthesizes its nature, functioning, and possible applications, as previously outlined in the Introduction. A carefully planned scientific observation of emblematic cases is conducted in the second phase of the research process. This approach aims to provide a direct understanding of the deepfake phenomenon, enabling broader generalization and ensuring the results' validity. This phase delves into computational techniques and observations necessary to detect audiovisual manipulation. The primary objective of this method is to discover strategies to detect deepfakes or audiovisual falsifications, as hypothesized in the context of the conflict in Ukraine, and to raise awareness of this delicate phenomenon which, beyond military action, represents a geopolitical tsunami (Baranowski, 2023).

---

## GENERAL DISCUSSION AND THE POLITICAL ROLE OF DEEPFAKES: IMPLICATIONS AND CHALLENGES

In election campaigns, AI has introduced new perspectives and strategies beyond traditional media use (Fasan, 2019; Sottocorona, 2019). Through predictive data analysis and information processing from various sources, AI allows candidates and their campaign teams to identify target voters more precisely (Crilley, 2018). This leads to the personalization of political messages tailored to voters' specific needs and preferences, thereby improving campaign effectiveness (Nunziata, 2021; Battista, 2023b). In formulating public policies, AI, on the other hand, offers analytical and simulation tools that enable policymakers to assess the potential impacts of policies more accurately and timely (Perrucci, 2019). The ability to process large amounts of data and anticipate the consequences of political decisions represents a significant advantage in governance. AI equally influences the interaction between politicians and citizens, as the presence of chatbots and virtual assistants on politicians' websites and the use of social platforms for political engagement are examples of how AI can facilitate direct communication between representatives and citizens (Bykov & Kurushkin, 2022; Viudes, 2023). However, this also raises ethical questions regarding transparency and accessibility of information (Nida-Rumelin & Weidenfeld, 2019).

In summary, AI has opened up new possibilities in election campaign dynamics, public policy formulation, and political interaction. Still, critical evaluation is required to understand its implications and challenges in the context of contemporary politics fully. Indeed, the advent of artificial intelligence (AI) is shaping a new phase of transformation in politics, going beyond the transformations previously introduced by the Internet, such as digital election campaigning and the widespread use of social media (Pacini, 2019; Reale & Tomasi, 2022). AI's impact on political dynamics generates deeper challenges that have significant implications for democratic processes. While the potential of artificial intelligence (AI) to improve human well-being is undeniable, its influence on democratic politics is characterized by considerable ambivalence. On the one hand, we benefit from AI-driven communication platforms that facilitate public debate, promote connections between individuals, and enhance the circulation of information. On the other hand, there are legitimate concerns regarding the use of AI applications that can undermine the foundations of democratic politics. These concerns go beyond mere election interference and touch upon the very core of democratic politics, as well as interpersonal relationships between citizens, citizens and their elected representatives, and citizens and the public institutions tasked with serving and safeguarding the common good. At this point, it is crucial to emphasize a particular concern that requires special attention: the potential ability of deepfakes, which represent manipulations of facial images, voices, content, or modes of expression generated by artificial intelligence, to influence democratic elections far beyond the data analysis techniques previously associated with cases like Cam-

bridge Analytica<sup>1</sup>. Notably, deepfakes are described as a significant source of concern for the stability of our society and the integrity of the political system (Westerlund, 2019). The recognition of counterfeit videos, known as deepfakes, constitutes one of the crucial challenges in the contemporary era of citizenship. Various methodologies exist to identify the genuineness of a video or visual image, distinguishing them from those subject to intentional manipulations. Despite the advancements made by artificial intelligence (AI) in deepfake recognition, it is worth noting that the same AI responsible for creating such fraudulent content operates in parallel. Consequently, deepfakes pose a significant issue of ethical and civic nature within the realm of modern digital languages and generative grammars, a subject that has been studied by numerous authors (Jivkova-Semova et al. 2017; Cerdán Martínez et al. 2020; Padilla & Presol, 2020; Ahmed, 2021; Appel & Prietzel, 2022). Among the issues associated with deepfakes, various pertinent questions emerge as they threaten national security by disseminating propaganda and potential interference in elections (Vaccari & Chadwick, 2020; Ursi, 2023). Indeed, the variation in the context of political marketing, election messages, and statements by leaders and politicians represents a highly problematic aspect intrinsically linked to the spread of fake news. This practice aims to discredit, attack, and delegitimize political opponents, shaping public opinion distortedly (Farkas & Schou, 2018). The dynamics underscore the importance of the timing and context in which the false news is presented, influencing public perception and the political climate in ways that vary depending on the moment and circumstances. Furthermore, these phenomena contribute to eroding citizens' trust in institutions and information from authoritative sources. In parallel, there are issues related to cybersecurity management, both for individuals and organizations. In the case of deepfakes, the threat becomes even more serious because these contents pose a greater danger than fake news as they are harder to detect and highly effective in manipulating public opinion, further complicating the task of the viewer in distinguishing between authentic and false information (Hu et al., 2022). Therefore, the main source of concern relates to the potential use of deepfakes to threaten national security through the dissemination of political propaganda or interference in electoral campaigns. This aspect is particularly critical in electoral contexts, where manipulating viral videos containing false statements by public figures can be a powerful tool for disinformation and manipulation of public opinion (Chesney & Citron, 2019). Manipulated videos have the capacity to significantly influence voters' decisions, thereby increasing the reach and effectiveness of manipulation in political choices. Consequently, it becomes evident that

---

<sup>1</sup> Cambridge Analytica was a British consulting firm that became infamous for a controversial data manipulation scandal aimed at influencing election campaigns. Its methodology integrated data mining, data management, and data analysis with a targeted political communication strategy. Using psychometrics, a field of study that analyzes human behaviors, the company could create detailed psychological profiles of users and use them to develop personalized messages tailored to their vulnerabilities and fears. The company gained notoriety, especially following the 2016 United States presidential elections, where it was accused of using data from millions of Facebook users to create highly targeted political advertisements and influence voter opinions.



the most insidious risk associated with deepfakes lies not only in the spread of disinformation but in the consequences stemming from this dissemination. The continuous exposure of citizens to manipulated content can lead them to develop increasing distrust towards most information and videos, leading to scenarios resembling an information apocalypse (Stover, 2018). Two well-known examples of deepfakes involving American presidents Obama and Trump have emerged in the political context. However, these are not the only cases in which deepfakes and video manipulation have impacted the political arena. Other instances have targeted prominent figures in American and, more broadly, international politics. A relevant example involves Nancy Pelosi, a prominent figure in American politics, who in May 2019 became the subject of ridicule through the dissemination of a manipulated video. In this video, the politician appeared to be drunk during a public speech. The clip quickly went viral and was even shared on social platforms by Donald Trump. Later, it was revealed that the video had been digitally manipulated to spread false information about the politician in question. Another significant issue is that the primary concern underlying these altered videos is not limited solely to famous individuals (Pardo, 2018). For some time now, the use of deepfakes has posed a significant danger to the conflict in Ukraine, as this technology has been exploited for purposes of propaganda, disinformation, and manipulation of public opinion, including the use of ordinary individuals

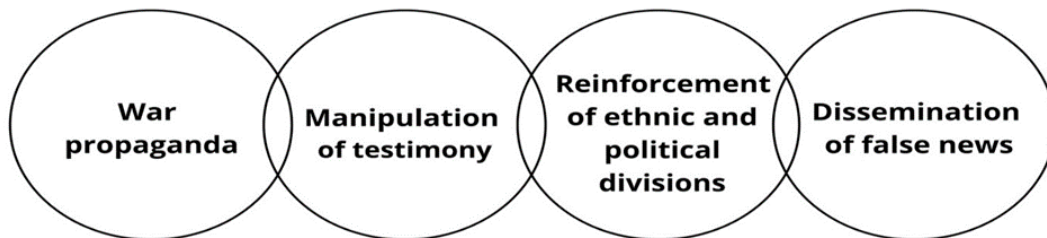


Figure 1: Conceptual framework overview

Source: The authors.

As highlighted in the graphical representation (Fig. 1), a conceptual framework has been outlined that encompasses the underlying reasons for creating such videos within the context of an ongoing conflict. It is important to note that, frequently, these products result from a complex and interconnected combination of various components rather than being the result of a single isolated factor. Initially, the factions involved in the conflict context may exploit deepfake technology to generate falsified video content representing fabricated or carefully manipulated events. These videos could be disseminated through online channels and traditional media to shape public opinion at the national and international levels.

Such a strategy could be adopted to justify their own actions and discredit the adversary, deceptively harnessing the persuasive power of moving images. Another application of deepfakes lies in the manipulation of testimonies provided by victims or witnesses of alleged war crimes. This practice would aim to question the authenticity of evidence presented in judicial settings, both at the national and international levels, casting doubt on the credibility of the statements provided. In this context, using deepfakes could compromise the integrity of legal proceedings and challenge the validity of testimonial evidence, posing a significant challenge to courts and legal institutions. Such a strategy could be adopted to alter the perception of the reality of events and influence the outcome of legal proceedings.

The widespread use of deepfake videos as vehicles for promoting divisive messages or exploiting ethnic or political tensions represents another significant concern. Such a practice could act as a catalyst in accentuating existing divisions within a particular community or nation. The ability to manipulate speeches and images to spread polarizing narratives could fuel internal conflicts and further complicate delicate peace and reconciliation processes. In this context, the strategic use of deepfakes could destabilise a country's social and political fabric, undermining cohesion and trust between ethnic groups or political factions. This could lead to increased instability and make it more difficult to achieve diplomatic solutions and peace agreements. Therefore, the irresponsible use of such technologies could have serious consequences for political stability and peaceful coexistence within a society. The use of deepfakes for spreading false news and manipulating reports on conflict developments represents another problematic dimension. More generally, using deepfakes can severely compromise the political reputation of the targeted individual. Creating fake or manipulated videos that appear to show a politician or leader in compromising situations or expressing ideas and positions contrary to reality can irreparably damage their public image and negatively influence public opinion. This type of manipulation can undermine the trust of voters and the public in the politician in question, thereby weakening their ability to exercise leadership and effectively represent the interests of the public. In India, for example, deepfake technology has been used as a tool against female politicians and journalists. Many of them have reported instances where their images were inserted into pornographic photos and videos disseminated on the internet. This abuse of artificial intelligence-generated images and deepfakes in political advertisements could easily be exploited to sexualize female politicians, significantly compromising their credibility during electoral campaigns (Tomić et al., 2023).

This issue is not limited to India but is observed in various parts of the world. A glaring example is the heavily manipulated fake video depicting Ukrainian President Volodymyr Zelenskyy purportedly ordering his soldiers to lay down their arms and abandon the fight against Russia. The consequences of these deepfakes are not only misinformation but also the broader erosion of truth, polluting the information ecosystem and casting

---

doubt on all content. This practice could significantly influence public perception and the interpretation of events related to the ongoing conflict. Creating false news through the application of deepfakes could create confusion among the public, undermine trust in reliable sources of information, and weaken the public's ability to access accurate and verifiable information. Misinformation fuelled by deepfakes could undermine the stability of political decisions and conflict management, making it difficult for leaders and international organizations to take effective measures. The spread of false information could lead to unforeseen reactions and further polarize public opinions, thus compromising the possibility of reaching a peaceful resolution to the conflict. Furthermore, the strategic use of deepfakes in this context could complicate the processes of verification and attribution of responsibilities, contributing to maintaining a state of uncertainty and discord. To mitigate these risks, it is essential to adopt strategies for detecting and verifying deepfakes, raise public awareness of the use of this technology, and develop mechanisms for rapid response to expose and refute false content. The international community and organizations involved in resolving the conflict in Ukraine should pay attention to these emerging challenges related to the use of deepfakes. During the conflict, the widespread use of deepfake videos has led to a general loss of trust among the public, significantly affecting perceptions of the authenticity and veracity of visual information. This phenomenon has fostered widespread doubts about the credibility of videos, sometimes leading to the erroneous rejection of genuine footage or the development of conspiracy theories suggesting that national leaders may have been replaced by mannequins (Twomey et al., 2023). This underscores the negative impact of advanced digital manipulation on public trust and the integrity of visual information, potentially destabilizing perceptions of political leadership and the management of international conflicts. The uncontrolled spread of deepfakes also threatens the stability of public information. It necessitates effective strategies to detect and counteract such manipulations, safeguarding trust in visual media and promoting an accurate understanding of global events.

## CONCLUSION

This article examined the interaction between artificial intelligence and the political context, particularly emphasising the need for a comprehensive and balanced assessment of AI's impact on politics. Through a multidimensional perspective, both the positive aspects, such as the potential enhancement of democratic participation and efficiency in addressing complex political issues, and the associated risks, including the concentration of algorithmic power, manipulation of public opinion, and the formation of information bubbles, were analyzed.

The study revealed the crucial importance of adopting a balanced and informed approach in the political segment to maximize benefits and mitigate disadvantages. AI can play a significant role in enriching democratic participation and facilitating data-driven political decision-making. However, taking preventive measures to avoid abuses and

distortions of the political process is equally essential. This is the case with deepfakes, representing a concrete manifestation of two interconnected processes. On one hand, the rapid technological progress, based on the use of artificial intelligence and big data management, has generated applications whose impact on the social sphere largely escapes control. On the other hand, a complex media ecosystem has emerged in which it is increasingly difficult to discern what is authentic from what is counterfeit, giving rise to a society shaped by simulacra and simulations, in line with the insights of the French philosopher Jean Baudrillard from the 1970s. This phenomenon represents one of the fundamental challenges of the modern era, where representations and realities have become intricately intertwined and overlapped. In this context, the notion of authenticity has been distorted and blurred as reality itself has become increasingly susceptible to manipulation and interpretation. New technologies, particularly artificial intelligence, have enabled the creation of content that can be difficult to distinguish from reality, if not impossible to recognize, without the use of specialized tools. This environment of simulacra and simulations has led to a situation where what is “true” has become a complex and contested issue. Representations and narratives can be created, altered, and disseminated with ease, often to promote certain perspectives or manipulate public opinion. Baudrillard’s theory emphasizes how this confusion between reality and representation profoundly affects our understanding of the world and our ability to make informed decisions. In an age where information is abundant but often counterfeit, it becomes essential to develop critical skills to navigate this environment of simulacra and simulations to preserve trust in institutions, information sources, and the very concept of reality. The COVID-19 pandemic has intensified the debate on the need to strengthen control mechanisms to prevent misinformation and develop fact-checking tools (Krause et al., 2020). Fact-checking has thus become a fundamental task of communication and represents a cornerstone of social responsibility, as it automates news verification by identifying sources, analyzing content, and managing various streams of information.

Many scholars (Graves, 2018) highlight the positive impact of fact-checking on political communication, promoting greater accountability among institutions and enhancing internal control mechanisms within journalistic enterprises. This phenomenon represents a necessary response to social pressure aimed at inducing changes in behaviors and attitudes, with three main objectives: ensuring the accuracy of the information, scrutinizing power, and transforming information into concrete knowledge (Rodríguez Pérez, 2020). Artificial Intelligence, while sometimes used for illicit purposes and often unchecked, can equally be harnessed to counteract such activities. Its applications in detecting cyber threats, preventing fraud, and combating misinformation are concrete examples of its positive potential. Through advanced data analysis and machine learning, AI can identify suspicious behavioral patterns and contribute to developing effective counterstrategies. Moreover, integrating AI-based solutions into institutional decision-making processes enhances the capacity to respond to contemporary challenges, underscoring AI’s crucial role in safeguarding information integrity and societal security. Tools like TruthBuzz,

---

designed to create compelling narratives, and Facterbot, a chatbot integrated into Facebook Messenger and available via email, offer innovative approaches to fact-checking. The Fake News Detector categorizes news based on credibility, employing sophisticated algorithms to assess source reliability. Additionally, Les Décodeurs, a section of Le Monde's website, employs a multidimensional analysis model, examining historical data of authors and sources to evaluate information reliability (Sitaula et al., 2019). These tools represent crucial technological solutions for tackling misinformation, enhancing the quality of public discourse, and promoting a culture of verified information while improving communication in political contexts. In this intricate scenario, where every representation takes on an artificial character and every artificial element translates into a representation, reality risks dissolving, subject to a series of technologies predisposed to reinforce preconceived beliefs or ideologies. According to a well-known sociological law called the "Thomas theorem," if individuals perceive certain situations as real, the consequences will be considered real in every respect. In other words, when individuals strongly believe in something or perceive a situation in a certain way, they will act based on those convictions, influencing the surrounding reality. This principle underscores the power of individual perceptions in shaping human behaviour and actions. For instance, if an individual perceives a threat where there may not be one, their response to this perception can have tangible effects on social dynamics and interpersonal relationships. This theorem highlights human experiences' subjectivity and capacity to impact the surrounding world through the reactions and decisions stemming from these perceptions. In the context of the article, the "Thomas theorem" is relevant because it emphasizes how the creation and dissemination of manipulated content, such as deepfakes, can lead people to perceive false situations as real. These perceptions will then influence their behaviour and could significantly impact political and social dynamics. Therefore, it is crucial to understand how individual perceptions, even when based on counterfeit information, can shape the social and political reality. For these and many other reasons that have emerged, future research should continue to investigate these complex dynamics, thereby contributing to a more comprehensive and nuanced understanding of the interaction between AI and politics and informing political decisions with solid and democratic principles-respecting foundations.

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

**CONFLICT OF INTEREST:** The author declares no conflict of interest.

## REFERENCES

Alvarez, R. M., Eberhardt, F., & Linegar, M. (2023). Generative AI and the Future of Elections. California Institute of Technology Center for Science, Society, and Public Policy. Retrieved from <https://lindeinstitute.caltech.edu/research/csspp>

- Ahmed, S. (2021). Who inadvertently shares deepfakes? Analyzing the role of political interest, cognitive ability, and social network size. *Telematics and Informatics*, 57, 101508.
- Appel, M., & Prietzel, F. (2022). The detection of political deepfakes. *Journal of Computer-Mediated Communication*, 27(4), zmac008.
- Baranowski, M. (2023). The Paradox of Two Wars in Ukraine: Between Independence and the Social Welfare of Others. *World Affairs*, 186(1), 226-234.
- Battista, D. (2023a). For better or for worse: politics marries pop culture (TikTok and the 2022 Italian elections). *Society Register*, 7(1), 117-142.
- Battista, D. (2023b). Knock, Knock! The Next Wave of Populism Has Arrived! An Analysis of Confirmations, Denials, and New Developments in a Phenomenon That Is Taking Center Stage. *Social Sciences*, 12(2), 100.
- Battista, D. & Uva, G. (2023). Exploring the Legal Regulation of Social Media in Europe: A Review of Dynamics and Challenges—Current Trends and Future Developments. *Sustainability*, 15(5), 4144.
- Bareis, J., & Katzenbach, C. (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, & Human Values*, 47(5), 855-881.
- Baudrillard, J. (1970). Fetichisme et ideologie: la reduction semiologique. *Nouvelle revue de psychanalyse*, 2, 213-224.
- Bykov, I. A., & Kurushkin, S. V. (2022). Digital political communication in Russia: Values of humanism vs. technocratic approach. *RUDN Journal of Political Science*, 24(3), 419-432.
- Bontridder, N., & Pouillet, Y. (2021). The role of artificial intelligence in disinformation. *Data & Policy*, 3, e32.
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford: Oxford University Press.
- Brock, J. K. U., & Von Wangenheim, F. (2019). Demystifying AI: What digital transformation leaders can teach you about realistic artificial intelligence. *California Management Review*, 61(4), 110-134.
- Cameron, J., & Hurd, G. A. (1984). *Terminator*. USA: Orion.
- Cerdán-Martínez, V., García-Guardia, M.-L., & Padilla-Castillo, G. (2020). Alfabetización moral digital para la detección de deepfakes y fakes audiovisuales. *Cuadernos de Información y Comunicación*, 25, 165-181. <https://doi.org/10.5209/ciyc.68762>
- Chesney, B., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *Calif. L. Rev.*, 107, 1753.

- 
- Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.
- Crilly, R. (2018). International relations in the age of 'post-truth' politics. *International Affairs*, 94(2), 417-425.
- Darlington, K. (2020). How Artificial Intelligence is helping prevent the spread of the COVID-19 pandemic. *IQ: The RIMPA Quarterly Magazine*, 36(3), 19-21.
- De Blasio, E. (2018). *Il governo online: le nuove frontiere della politica* (No. 98, pp. 1-260). Roma: Carocci editore.
- Fasan, M. (2019). Intelligenza artificiale e pluralismo: uso delle tecniche di profilazione nello spazio pubblico democratico. *BioLaw Journal-Rivista di BioDiritto*, 1, 101-113.
- Floridi, L. (2022). *Etica dell'intelligenza artificiale: Sviluppi, opportunità, sfide*. Milano: Raffaello Cortina Editore.
- Gallo, M., Fenza, G., & Battista, D. (2022). Information Disorder: What about global security implications? *Rivista di Digital Politics*, 2(3), 523-538.
- Graves, L. (2018). Boundaries not drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies*, 19(5), 613-631.
- Groves, L., Peppin, A., Strait, A., & Brennan, J. (2023, June). Going public: the role of public participation approaches in commercial AI labs. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1162-1173).
- Haenlein, M. & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5-14.
- Hu, L., Wei, S., Zhao, Z., & Wu, B. (2022). Deep learning for fake news detection: A comprehensive survey. *AI Open*, 3, 133-155, <https://doi.org/10.1016/j.aiopen.2022.09.001>
- Lepri, B., Oliver, N., & Pentland, A. (2021). Ethical machines: The human-centric use of artificial intelligence. *IScience*, 24, 102249.
- Li, B., Chen, P., Liu, H., Guo, W., Cao, X., Du, J., ... & Zhang, J. (2021). Random sketch learning for deep neural networks in edge computing. *Nature Computational Science*, 1(3), 221-228.
- Kertysova, K. (2018). Artificial intelligence and disinformation: How AI changes the way disinformation is produced, disseminated, and can be countered. *Security and Human Rights*, 29(1-4), 55-81.
- Klinger, U., Kreiss, D., & Mutsvauro, B. (2023). Platforms, Power, and Politics: A Model for an Ever-changing Field. *Political Communication Report*, 27, 1-6. <http://dx.doi.org/10.17169/refubium-39045>

- Krause, N. M., Freiling, I., Beets, B., & Brossard, D. (2020). Fact-checking as risk communication: the multi-layered risk of misinformation in times of COVID-19. *Journal of Risk Research*, 23(7-8), 1052-1059.
- Jenkins, H. (2014). *Cultura convergente*. Rimini: Maggioli Editore.
- Jivkova-Semova, D., Requeijo-Rey, P., & Padilla-Castillo, G. (2017). Usi e tendenze di Twitter nella campagna per le elezioni generali spagnole del 20 dicembre 2015: hashtag che erano argomenti di tendenza. *Professionista dell'informazione*, 26(5), 824-837.
- Nida-Rumelin, J., & Weidenfeld, N. (2019). *Umanesimo digitale: un'etica per l'epoca dell'Intelligenza artificiale*. Milano: FrancoAngeli.
- Nunziata, F. (2021). Il platform leader. *Rivista di Digital Politics*, 1(1), 127-146.
- Pacini, F. (2019). Una modesta proposta: il politico virtuale. Prime note sull'intelligenza artificiale come provocazione in tema di rappresentanza politica. *BioLaw Journal-Rivista di BioDiritto*, 1, 115-133.
- Padilla Castillo, G. & Presol Herrero, Á. (2020). Ética y deontología en publicidad. Nike'Dream Crazier'2019 como campaña feminista en Instagram. *Comunicación y Género*, 3(1), 3-15.
- Pardo, L. (2018). *Statistical inference based on divergence measures*. Florida: CRC press.
- Perrucci, A. (2019). Dai Big Data all'ecosistema digitale. Dinamiche tecnologiche e di mercato e ruolo delle politiche pubbliche. *Analisi Giuridica dell'Economia*, 18(1), 61-88.
- Reale, C. M. & Tomasi, M. (2022). Libertà d'espressione, nuovi media e intelligenza artificiale: la ricerca di un nuovo equilibrio nell'ecosistema costituzionale. *DPCE online*, 51(1), 325-336.
- Reis, J., Santo, P. E., & Melão, N. (2021). Influence of artificial intelligence on public employment and its impact on politics: a systematic literature review. *Brazilian Journal of Operations & Production Management*, 18(3), 1-22.
- Rodríguez Pérez, C. (2020). Una reflexión sobre la epistemología del fact-checking journalism: retos y dilemas. *Revista de comunicación*, 19(1), 243-258.
- Russell, S. & Norvig, P. (1995). A modern, agent-oriented approach to introductory artificial intelligence. *Acm Sigart Bulletin*, 6(2), 24-26.
- Sottocorona, C. (2019). *# AI Challenge: Amica o nemica? Come l'Intelligenza Artificiale cambia la nostra vita*. Milano: HOEPLI EDITORE.
- Sitaula, N., Mohan, C. K., Grygiel, J., Zhou, X., & Zafarani, R. (2020). Credibility-based fake news detection. *Disinformation, misinformation, and fake news in social media: Emerging research challenges and Opportunities*, 163-182.



- Sharma, A., Rani, S., & Gupta, D. (2020). Artificial intelligence-based classification of chest X-ray images into COVID-19 and other infectious diseases. *International Journal of Biomedical Imaging*, 1-10.
- Stover, D. (2018). Garlin Gilchrist: Fighting fake news and the information apocalypse. *Bulletin of the Atomic Scientists*, 74(4), 283-288.
- Thomas, W. I. (2005). La definición de la situación. *CIC. Cuadernos de Información y Comunicación*, 10, 27-32.
- Tomić, Z., Damnjanović, T., & Tomić, I. (2023). Artificial intelligence in political campaigns. *South Eastern European Journal of Communication*, 5(2), 17-28.
- Twomey, J., Ching, D., Aylett, M. P., Quayle, M., Linehan, C., & Murphy, G. (2023). Do deepfake videos undermine our epistemic trust? A thematic analysis of tweets that discuss deepfakes in the Russian invasion of Ukraine. *Plos one*, 18(10), e0291668.
- Ursi, R. (Ed.). (2023). *La sicurezza nel Cyberspazio*. Milano: Franco Angeli.
- Vogt, J. (2021). Where is the human got to go? Artificial intelligence, machine learning, big data, digitalisation, and human-robot interaction in Industry 4.0 and 5.0: Review Comment on: Bauer, M.(2020). Preise kalkulieren mit KI-gestützter Onlineplattform BAM GmbH, Weiden, Bavaria, Germany. *AI & Society*, 36(3), 1083-1087.
- Viudes, F. J. (2023). Revolucionando la política: El papel omnipresente de la IA en la segmentación y el targeting de campañas modernas. *Más poder local*, 53, 146-151.
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 40-53.
- Wirtz, B. W., & Müller, W. M. (2019). An integrated artificial intelligence framework for public management. *Public Management Review*, 21(7), 1076-1100.
- Ziccardi, G. (2019). *Tecnologie per il potere: come usare i social network in politica*. Milano: Raffaello Cortina Editore.

#### BIOGRAPHICAL NOTE

Daniele Battista, PhD, is a researcher at the Department of Business Sciences - Management & Innovation Systems, University of Salerno, Italy. His main research interests revolve around media theory, with a specific focus on the relationship between social networks and political communication.

**OPEN ACCESS:** This article is distributed under the terms of the Creative Commons Attribution Non-commercial License (CC BY-NC 4.0) which permits any non-commercial use, and reproduction in any medium, provided the original author(s) and source are credited.

**JOURNAL'S NOTE:** *Society Register* stands neutral with regard to jurisdictional claims in published figures, maps, pictures and institutional affiliations.

**ARTICLE HISTORY:** Received 2024-02-12 / Accepted 2024-06-30

