



Global Campus
Latin America-Caribbean

Awarded Theses
2024 / 2025

Agostina Dasso Boonman

Deepfakecracia

Electoral Deepfakes as Disinformation Tools
in the 2023–2024 Presidential Elections

LATMA, Master's Programme in Human Rights and
Democratisation in Latin America and the Caribbean

Agostina Dasso Boonman

Deepfakecracia

Electoral Deepfakes as Disinformation Tools
in the 2023–2024 Presidential Elections

Dasso Boonman, Agostina, *Deepfakecracia: Electoral Deepfakes as Disinformation Tools in the 2023–2024 Presidential Elections*. Supervisor: Ivan Schuliaquer, National University of San Martin (GCHR, 2026). Master's Programme in Human Rights and Democratisation in Latin America and the Caribbean (LATMA-GC Latin America and Caribbean), coordinated by National University of San Martin (Argentina).
DOI: <https://doi.org/10.25330/3187>

The Global Campus of Human Rights is a unique network of more than one hundred participating universities around the world, seeking to advance human rights and democracy through regional and global cooperation for education and research. This global network is promoted through eight Regional Programmes which are based in Venice for Europe, in Sarajevo/Bologna for South East Europe, in Yerevan for the Caucasus, in Pretoria for Africa, in Bangkok for Asia-Pacific, in Buenos Aires for Latin America and the Caribbean, in Beirut for the Arab World, and in Bishkek for Central Asia.

Every year each regional master's programmes select the best master thesis of the previous academic year that is published online as part of the GC publications. The selected GC master theses cover a range of different international human rights topics and challenges.

The Global Campus Awarded Theses of the academic year 2024/2025 are:

- Arkorful-Ewool, Peter, *Promoting Inclusivity through Legal Frameworks: An analysis of Ghana's Disability Policies and their Effectiveness in Supporting Children with Intellectual Disabilities*. Supervisor: Elvis Fokala, Center of Human Rights University of Pretoria. Master's Programme in Human Rights and Democratisation in the Caucasus (CES-GC Caucasus), coordinated by Yerevan State University.

- Baker, Isabella Eda, *The Risk of Statelessness and Humanitarian Exclusion: Legal and Structural Barriers for Non-Syrian Refugees in Jordan*. Supervisor: Ibrahim Al-Jazy, University of Jordan. Arab Master's Programme in Democracy and Human Rights (ArMA-GC Arab World), coordinated by Saint Joseph University (Lebanon).
- Dasso Boonman, Agostina, *Deepfakecracia: Electoral Deepfakes as Disinformation Tools in the 2023–2024 Presidential Elections*. Supervisor: Ivan Schuliaquer, National University of San Martin. Master's Programme in Human Rights and Democratisation in Latin America and the Caribbean (LATMA-GC Latin America and Caribbean), coordinated by National University of San Martin (Argentina).
- Lomakin, Ilya, *Status of Post-February 2022 Russian Migrants: A Comparative Study*. Supervisor: Lucia Leontiev, European University Institute (Italy). The Master's Programme in Human Rights and Sustainability in Central Asia (MAHRS-GC Central Asia), coordinated by the OSCE Academy in Bishkek (AY 2023-25).
- Mallawaarachchi, Isurinie Anuradha, *Intimate Partner Violence: Case Studies of Transgender Women in Sri Lanka as Victims and Perpetrators*. Supervisors: Kalana Senaratne, University of Peradeniya and Naparat Kranrattensuit, SJD (Judicial Science). Master's Programme in Human Rights and Democratisation in Asia Pacific (APMA-GC Asia-Pacific), coordinated by Mahidol University (Thailand).
- Ndrevataj, Entenela, *Exposing Shadows: Exploring the Intersection of Image-Based Sexual Abuse in Kosovo*. Supervisor: Zilka Spahic Siljak, University of Zenica (Bosnia and Herzegovina)-Feminism and Religion (FER) School. Master's Programme in Democracy and Human Rights in South-East Europe (GC SE Europe-ERMA), coordinated by University of Sarajevo and University of Bologna.
- Orr, Kate, *Missing from the Data, Missing from the Law: Indigenous Women and Girls and the Failures of Anti-Trafficking Frameworks in Latin America*. Supervisor: Davinia Gómez Sánchez, Univeristy of Deusto (Bilbao). European Master's Programme in Human Rights and Democratisation (EMA-GC Europe), coordinated by Global Campus of Human Rights Headquarters.

- Vasilenko, Darya, *Assessing the Extent of Implementation of the Convention on the Elimination of All Forms of Discrimination Against Women: Concluding Observations on Kazakhstan*. Supervisor: Aigerim Kussaiynkyzy, Maqsut Narikbayev University (Kazakhstan). The Master's Programme in Human Rights and Sustainability in Central Asia (MAHRS-GC Central Asia), coordinated by the OSCE Academy in Bishkek (AY 2024-26).
- Yisa, Josephine Boye, *A Human Rights-Based Approach to Overfishing in West Africa: Comparative Study of Guinea-Bissau and The Gambia*. Supervisors: Magnus Killander, University of Pretoria and Emmanuel Kasimbazi, Makerere University (Uganda). Master's Programme in Human Rights and Democratisation in Africa, coordinated by Centre for Human Rights, University of Pretoria.

Biography

Agostina Dasso studied law at University of Buenos Aires and recently graduated as Magister in Human Rights and Democratisation in Latin American and the Caribbean (University of San Martin). She is auxiliar professor at Human Rights Law, International Public Law and Criminal Law courses at UBA. Her research interest focuses on new technologies and gender from a Latin American perspective.

Deepfakes have become a tool for disinformation during election periods, using the image of candidates in fake videos, images and audios generated by artificial intelligence. This contributes to undermining citizens' access to information and reinforcing the sense of uncertainty in the information environment, eroding trust in democratic processes. The objective of this thesis is to analyse the functioning of deepfakes as a disinformation tool in election contexts based on the use of deepfakes in presidential elections in ten American countries in the period 2023–2024.

Keywords: *Artificial intelligence and human rights / deepfakes / political rights / access to information in electoral contexts.*

Acknowledgments

*To my mom for giving me wings to fly and all her love.
To my dad for his love and companionship.
To my brothers and friends for being my life companions.
To the stars Celina, Horacio, Bety I saw crossing.
Especially to my grandfather Lorenzo, for teaching me to take
risks in the simplicity of everyday life.
Superior.*

Table of Abbreviations

IA	Artificial Intelligence
UN	United Nations
UNGA	United Nations General Assembly
OECD	Organisation for Economic Cooperation and Development.
IAHR	Inter-American System of Human Rights
RELE	Office of the Special Rapporteur for Freedom of Expression
OAS	Organisation of American States
DFRL	Digital Forensic Research Lab

Table of Contents

III	Foreword
VI	Biography
VII	Abstract
VIII	Acknowledgments
IX	Table of Abbreviations
X	Table of Contents
XIII	List of Tables and Figures

1	Introduction and presentation of the problem
4	Research objectives
4	Methodology

7	1. Infocracy
8	1.1 Why continue talking about disinformation?
11	1.2 Is talking about fake news the same as talking about disinformation?
20	1.3 ABC of disinformation
20	1.3.1 Manipulative actors
21	1.3.2 Deceptive behaviour
25	1.3.3 Harmful or detrimental content
26	1.4 Electoral disinformation and social networks
28	1.4.1 How does disinformation operate in the political arena?
28	1.5 Electoral disinformation in Latin America
30	1.6 Social actors of disinformation
30	1.6.1 Citizenship

31	1.6.2 Classic or traditional media
32	1.6.3 Platforms
33	1.6.4 States and political parties
34	1.6.5 Russia in Latin America
35	1.6.6 The case of Venezuela
35	1.6.7 The Bolsonaro case in Brazil
39	1.6.8 Political consulting firms
39	1.7 What do consultancies do?
39	1.8 Strategies for disinformation
40	1.9 Research findings
41	1.10 Conclusion of Chapter 1
42	2. Deepfakecracy: The foundations
44	2.1 Artificial Intelligence
46	2.1.1 A.I. When did we start talking about artificial intelligence?
50	2.2 Artificial intelligence today
51	2.2.1 Stages of artificial intelligence
52	2.2.2 Narrow artificial intelligence
53	2.2.3 Machine learning
57	2.2.4 Generative AI
60	2.3 Deepfakes or ultra fake content
60	2.3.1 Were they always among us? Evolution of the deepfakes
62	2.3.2 How are they produced? Introduction to deepfakes production
66	2.5 Deepfakes: Political arena and disinformation
67	2.6 Conclusion of Chapter 2
69	3. American deepfakecracy
70	3.1 State of the art: Deepfakes in the 2023 election campaigns
72	3.2 American deepfakecracy
73	3.2.1 Methodology applied to comparative analysis
76	3.2.2 American electoral cycle
121	3.3 Conclusions of Chapter 3

123	4. Deepfakecracia and human rights
123	4.1 American deepfakecracy: Electoral deepfakes
126	4.1.1 Local narratives
127	4.1.2 Technopolarity
128	4.1.3 The liar's dividend
128	4.1.4 Informed consent
129	4.4.5 Deepfakes and gender
132	4.2 What is the reaction to the American deepfakecracy?
	Attempts at national and international regulations
135	4.2.1 Latin American national regulations
144	4.2.2 International regulations
147	4.3 Where do human rights stand in the face of disinformation
	by deepfakes in electoral contexts within the Inter-American
	human rights system?
157	4.4 Conclusion of Chapter 4

158 **5. Conclusion**

160 **Bibliography**

List of Tables and Figures

12	Table 1. Typologies of disinformation
18	Table 2. Biases
24	Table 3. Botnets patterns
52	Table 4. Narrow artificial intelligence categories
53	Table 5. Machine learning's systems types
53	Table 6. Machine learning techniques
55	Table 7. Deep learning layers
66	Table 8. Description of deepfakes types
73	Table 9. Variables of analysis
76	Table 10. American electoral cycle
23	Figure 1. Trolls, Bots and disinformation dynamics
49	Figure 2. Neural networks
59	Figure 3. AI types summary
65	Figure 4. Deepfakes production scheme
65	Figure 5. Types of deepfakes
79	Figure 6. Patricia Bullrich AI-altered audio
81	Figure 7. Myriam Bregman deepfake
82	Figure 8. Javier Milei deepfake
84	Figure 9. Two versions of Sergio Massa deepfake

- 84 **Figure 10. Sergio Massa deepfake from an X account**
- 86 **Figure 11. Biden's deepvoice**
- 88 **Figure 12. Kamala Harris deepfake published by Elon Musk's X account**
- 89 **Figure 13. Kamala Harris sexualised deepfake**
- 90 **Figure 14. Kamala Harris AI-manipulated audio through TikTok accounts users**
- 91 **Figure 15. Donald Trump deepfakes**
- 92 **Figure 16. Disinformation through deepfakes and photomontages of Taylor Swift in the United States**
- 92 **Figure 17. Disinformation through photomontages of Lionel Messi in Argentina**
- 96 **Figure 18. Ecuador's candidate Luisa González deepfake**
- 98 **Figure 19. Paraguay's candidate Payo Cubas video**
- 103 **Figure 20. Dominican Republic's candidate Luis Abinader deepfake**
- 106 **Figure 21. Panama's candidate Romulo Roux deepfake**
- 110 **Figure 22. Mexico's candidate Claudia Sheinbaum deepfake**
- 113 **Figure 23. Venezuela's opposition deepfake disseminated by Domingo Hernández Lárez X account**
- 116 **Figure 24. Uruguay's candidate Yamandú Orsi deepfake**
- 119 **Figure 25. Ecuador's candidate Luisa Gonzalez deepfake**

Introduction and presentation of the problem

Zelensky's surrender to Russia in the midst of armed conflict and Trump's arrest during a climate of social unrest are two cases of leaders from countries on different continents that not only share the commonality of having been generated by artificial intelligence (AI) but also of being false. This latter aspect is what is known as *deepfake*.

The term deepfake results from the combination of the words *deep learning* and *fake*, the latter alludes to the false nature of the content produced.¹ This definition gathers the coinciding elements in the definitions provided in the consulted bibliography, insofar as one of the problems surrounding the issue is that there is no univocal and precise definition of deepfakes.

Deepfakes serve various purposes, which according to Stanford University's AI Index Report 2024 are visual: face swap, lip-syncing, puppet-mastery, entire face synthesis, facial attribute manipulation and auditory types (speech recognition and transcription, and speech conversion).

Deepfakes have burst onto the global scene, appearing in multiple fields such as advertising, film production, creative arts, and the political arena is no exception. Thus, various candidates in presidential or parliamentary elections have been subjected to manipulated audiovisual materials during their campaigns, and even the images of celebrities have been used to support a specific candidacy. These examples fall under the category of 'political deepfakes'.

¹ Robert Chesney and Danielle K Citron cited in Gergely Ferenc Lendvai and Gergely Gosztonyi, 'Deepfake y Desinformacion- ¿Que puede hacer el derecho frente a las noticias falsas creadas por deepfake?' (2024) 41 IDP 1.

This work aims to establish a definition of this concept, as it has not been precisely defined. For the purposes of this thesis, only cases related to political deepfakes about candidates in presidential elections will be considered, without prejudice to the fact that the concept of ‘political’ has an extensive perspective, as it can encompass all those leaders, candidates, representatives, militants or activists who exercise a political role or influence.²

Now, these productions imply ‘one of the most significant evolutions of digital manipulation. This phenomenon has flooded media and networks, becoming integrated into the fabric of current algorithmic culture’.³

The above means that political deepfakes emerge within a context of accelerated consumption of audiovisual content, both informative and entertaining, on social media. This, combined with the development of machine learning technology, transforms how we approach reality. Specifically, it fits into a pressing context of the viral spread of false narratives and representations – in other words, as a new tool in the current disinformation landscape.

Disinformation, that is, the dissemination of false or misleading information that manipulates, confuses or influences public opinion intentionally or unintentionally⁴ is considered a constant challenge in Latin America that, in light of the adoption of technology, takes on different characteristics that could potentially violate human rights, such as citizens’ access to information.

In this dialectic between disinformation in the context of deepfakes, it means:

the deliberate dissemination of false or misleading information with the aim of strategically manipulating an audience or creating further divisions ... This phenomenon is exacerbated by the technological sophistication of deepfakes, which makes false content more convincing.⁵

² Almudueña Barrientos-Báez, María Teresa Piñero Otero and Denis Porto Renó, ‘Imágenes falsas, efectos reales. Deepfakes como manifestaciones de la violencia política de género’ (2024) 82 Revista Latina de Comunicación Social 1.

³ *ibid* 4.

⁴ Rafael Rubio Nuñez, Roeliselabell García Rossó and Vanessa Sánchez, ‘Desinformación: como entenderla, combatirla y protegerse de sus efectos’ (Friedrich Naumann Foundation for Freedom Regional Office for Latin America 2024) <https://www.freiheit.org/sites/default/files/2024-12/03dic_paper_desinformacion_versioncorta02.pdf> accessed 16 January 2026.

⁵ Lendvai and Gosztonyi, “Deepfake y Desinformación- ¿Que puede hacer el derecho frente a las noticias falsas creadas por deepfake?” (2024) N° 41 IDP 1.

To the above, it should be added that particularly in the political and electoral sphere, the digital world has become an indispensable medium in electoral campaigns or as the centre of public debate, and it is there that the devices for disinformation appear hand in hand with deepfakes.

The current context has prompted regulatory attempts by various countries. American, European and Chinese solutions have been identified⁶ which respond to different criteria, and in particular, the human rights approach is either absent or overlooked.

Therefore, this thesis starts from the idea that there is a link between deepfakes in general and politicians in particular and disinformation, especially in electoral processes, in which they are used as a tactic to generate distrust in candidates and public institutions.⁷

Thus, the research problem arises of how deepfakes manifest themselves as disinformation techniques in electoral contexts, in order to frame them as a problem that has been addressed by few states from a human rights perspective, and even then, the responses may prove insufficient. This last point refers to considering the international legal framework and reflecting on whether the current human rights paradigm is prepared to receive and regulate the development of AI in the case of political deepfakes.

The theoretical contributions that this thesis intends to make are the conceptualisation of electoral deepfakes and the characteristics they present as a disinformation tool in electoral processes in the American continent, with special emphasis on Latin America based on the previous studies consulted and the cases proposed in the methodology described below.

This presentation will also attempt to highlight electoral deepfakes as a problem that affects human rights and, at the same time, understand the regulations that have emerged to counteract them. Regarding the latter, it will examine the systems currently in place and whether or not they incorporate a human rights approach.

⁶ Lendvai and Gosztonyi, (n 5).

⁷ Rubio Nuñez, García Rossó and Sánchez (n 4).

Research objectives

The overall objective of this work is to analyse how electoral deepfakes act as a disinformation tool in the presidential elections during the period 2023–2024 and to understand what regulatory systems have been put in place in response to their emergence in the political arena.

To achieve this, the following specific objectives are set:

First, investigate what the main features of disinformation are in electoral contexts.

The second specific objective is to analyse the emergence of deepfakes and how they fit into the phenomenon of electoral disinformation in order to understand their characteristics.

The third specific objective is to try to collect, in an experimental way, cases of political deepfakes during presidential elections within the established time range and to highlight particularities of Latin American cases if they exist.

Finally, the aim is to compare the main legal systems regulating elections in order to understand the legal impact of electoral deepfakes. This process will highlight whether a human rights approach is present in these legal responses.

Methodology

To carry out the research and meet the proposed objectives, it is proposed to approach the topic from a qualitative perspective, based on the review of documents (articles, observatory studies), treaties, local regulations and reports from international and regional organisations concerning human rights in relation to the articulation between disinformation in elections, electoral deepfakes as a product of AI and citizens' access to information.

The thesis analyses a non-representative experimental sample of electoral deepfakes that use the image or refer to the person of the main candidates for the highest executive office at the national level; that is, the following will be considered: deepfakes referring to the presidential candidacies.

The study uses a sample of electoral deepfakes that have received media coverage in countries across the Americas: Ecuador, Paraguay, Guatemala, Argentina, El Salvador, Dominican Republic, Panama, Mexico, Venezuela, Uruguay and the United States.

Thus, the chosen time frame corresponds to presidential elections initiated during the selection of two calendar years with a start date of 9 February 2023 until 9 February 2025. This time frame allows for the inclusion of presidential elections in the years 2023 and 2024, as this is an important period of time in terms of democratic participation and the holding of elections in approximately 31 countries in 2023 and 64 countries in 2024 respectively. The latter is considered a super election year or year of democracy,⁸ even *TIME* magazine considered it as ‘the election year’⁹ since 49% of the global population votes. Furthermore, the selected time frame allows for a comparison of the proliferation of deepfakes in the early elections in Ecuador in 2023 and the general elections held in 2025.

For the purposes of media coverage during the election campaigns, the following will be considered: cases reported by online newspapers or fact-checking organisations, as well as those cases that have gained visibility on social media platforms such as Instagram, Facebook and Twitter, given that these are the areas where such cases proliferate and where public debate has currently migrated. The narratives presented by each of these platforms will be taken into account.

It should be clarified that the analysis will not focus on the tone of user comments or the choice of emoticons; it simply aims to collect examples of deepfake audiovisual materials (electoral deepfakes) and analyse their content to account for this small part of the disinformation phenomenon in electoral contexts. It should be added that deepfakes that are satirical or make jokes about the candidates will not be analysed.

Finally, regarding the selection of cases, a query will be made in the Google search tools with the generic phrases: ‘deepfake in (country)’, ‘deepfake politicians in (country)’, ‘deepfake presidential candidate in (country)’, ‘deepfake and elections in (country)’. Therefore, the selection of articles, blogs and digital newspapers will be done manually, and all those cases of electoral deepfakes that do not appear in the context of specific presidential elections will be excluded.

⁸ Mateusz Labuz and Christopher Nehring, ‘On the Way to Deep Fake Democracy? Deepfakes in Election Campaigns in 2023’ (2024) 23 *European Political Science* 454.

⁹ Koh Ewe, ‘The Ultimate Election Year: All the Elections Around the World in 2024’ (*TIME*, 28 December 2023) <<https://time.com/6550920/world-elections-2024/>> accessed 16 January 2026.

Regarding the limitations of this thesis, one is that the search focuses on cases that have gained prominence in digital media and social networks. This could mean that local cases, or those ‘shared’ or ‘viralised’ within limited circles, might be omitted or go unnoticed, thus failing to receive the same visibility. It should also be mentioned that the deepfake cases analysed here are constantly being updated, so recent examples are excluded from these pages.

Furthermore, when analysing the narratives introduced, it is important to acknowledge their local element, that is, their adaptation to the idiosyncrasies and context of the country in which they are introduced. While the aim is to explore this context, the limitations that may exist in this regard cannot be ignored.

On the other hand, there is no methodological framework that establishes an adequate analysis of political deepfakes or that establishes criteria to measure the impact of deepfakes on election results or the limits of their influence.

Based on the theoretical framework and the analysis of the selected cases, this thesis is organised into four chapters.

The first chapter focuses on compiling developments on disinformation, specifically on social networks and in electoral contexts, with the purpose of analysing its main characteristics.

The second chapter then attempts to delve into the world of deepfakes, starting with an exposition on generative AI and its development leading to the emergence of deepfakes, highlighting the role of human action in their production.

The third chapter seeks to present identified cases of electoral deepfakes in American countries during the presidential elections held within the established timeframe and collected in an exploratory manner.

Finally, the fourth chapter presents the three major AI regulatory systems currently in place (American, European and Chinese) and compares the legal responses of Latin American states to electoral deepfakes. It is important to emphasise whether these responses were framed within the human rights paradigm. The chapter also examines electoral deepfakes as a problem that violates citizens’ access to information under the standards established by the Inter-American human rights system.

1. Infocracy



The phenomenon of
disinformation during
election periods

‘The mass hysteria demonstrated the power of the mass media, and this curious episode ... The program lasted almost 59 minutes: the first forty ones corresponded to the fake news report ...’

José Pablo Feinmann, *Filosofía política del poder mediático* (1st edn, Planeta 2013)

The audio to which the reader has been directed corresponds to the broadcast of 30 October 1938, in which Orson Welles and the Mercury Theatre, under the CBS banner, adapted HG Wells’s *The War of the Worlds* into a radio script. They narrated the events in the form of a news report.

The broadcast lasted 59 minutes. During that time, it was mentioned approximately twice that it was a dramatisation. This led listeners who tuned in but missed the introduction and announcements to believe it was a real news broadcast. The result was panic in the streets of New York and New Jersey, as people thought they were being invaded by aliens. Police stations and newsrooms were also overwhelmed with calls from terrified and desperate listeners trying to protect themselves from the fictional Martian gas attacks.¹⁰

The following day, protests demanded accountability and explanations, which resulted in nothing less than an apology from Welles. This episode is considered by writer José Pablo Feinmann as an example of mass manipulation through a far from innocent tool like radio, which, depending on who controls it, can produce different results for the audience.

¹⁰ José Pablo Feinmann, *Filosofía Política del poder mediático* (1st edn, Planeta 2013).

In light of the current landscape, where public life is discussed on social media, it is reasonable to question the relevance of disinformation as a factor that paradoxically undermines citizens' access to information. How many 'wars of the worlds' spreading are we aware of at the same time? Or perhaps more urgently, do we fail to notice our status as hostages to what Mark Zuckerberg calls the 'war for trust' in which 'operators are battlefields, profiles are weapons, and information is understood as ammunition in a genuine arms race to win that race'.¹¹

For this reason, this first chapter aims to present the main features studied regarding disinformation as a phenomenon in general and in particular in electoral contexts in the digital age.

1.1 Why continue talking about disinformation?

Disinformation is not a new phenomenon; it has been present throughout history and, consequently, has taken on various forms in tandem with technological evolution.

In Ancient Rome and Greece, it manifested itself through speeches and rumours that sought to influence public opinion and discredit political adversaries:¹² 'Leaders and orators manipulated information to shape citizens' perceptions and gain support. This strategy laid the foundation for the systematic use of disinformation as a tool of social and political control'.¹³

Later, the ability to print and distribute mass texts with the introduction of the printing press in the 15th century increased the speed at which misinformation spread. The same thing happened with radio and television:

¹¹ M Zuckerberg (2018) cited in Rafael Rubio Nuñez, Roelisabell Garcia Rossó and Vanessa Sánchez, 'Desinformación: como entenderla, combatirla y protegerse de sus efectos' (Friedrich Naumann Foundation for Freedom Regional Office for Latin America 2024) <https://www.freiheit.org/sites/default/files/2024-12/03dic_paper_desinformacion_versioncorta02.pdf> accessed 16 January 2026.

¹² D Lynch (1991) cited in Rubio Nuñez, Garcia Rossó and Sánchez (n 11).

¹³ *ibid* 2.

In the twentieth century, radio and television added new dimensions to disinformation. Propaganda techniques were refined and became a crucial tool during the world wars. Disinformation campaigns sought to deceive enemies while maintaining morale and controlling public opinion in times of crisis. The ability of these new media to reach a mass audience amplified the impact of disinformation, making its power in politics and society evident.¹⁴

So, if misinformation has always been among us, why are we interested in it now? Well, it has acquired new characteristics through technological advances, which has allowed a radical change in its nature and logic.

On the one hand, it has adopted these features thanks to the internet, as a decentralised network that has allowed people to express themselves in the public sphere, with access to minimal technological resources.¹⁵

To this must be added the rise of social media, as well as the development of strategies and techniques that promote disinformation. It is in this context that AI emerges through the generation of deepfakes, whose viral spread is favoured by automation products such as bots.¹⁶

While technological advancements amplified by the democratisation of information technologies strengthen public debate as well as the circulation and access to information, they also constitute a new communication ecosystem that brings with it new dilemmas:

¹⁴ D Lynch (1991) cited in Rubio Nuñez, García Rossó and Sánchez (n 11) 2.

¹⁵ Special Rapporteurship for Freedom of Expression of the Inter-American Commission on Human Rights with contributions from the Department of Electoral Cooperation and Observation and the Department of International Law of the General Secretariat of the Organization of American States, 'Guide to guarantee freedom of expression in the face of deliberate disinformation in electoral contexts' (OEA/Ser.D/XV.22 2019). <https://www.oas.org/es/cidh/expresion/publicaciones/Guia_Desinformacion_VF.pdf> accessed 16 January 2026.

¹⁶ Paulo Carlos López López, Andrea Mila Maldonado and Vasco Ribeiro, 'La desinformación en las democracias de América Latina y de la península ibérica: de las redes sociales a la inteligencia artificial (2015-2022) (2023) 8 Uru: revista de comunicación y cultura 69, 79.

such as the speed with which information is produced and distributed, the volume of information generated thanks to the exponential multiplication of sources, and the capacity of information to expand horizontally in a 'viral' way in this decentralized network, along with the ability to direct messages to very specific demographic segments.¹⁷

In other words, in the modern context of the digital age, disinformation takes on a new dimension, as Silvio Waisbord points out, due to the popularity of the internet which has eroded the vertical structure of knowledge production and dissemination.¹⁸

Similarly, digital platforms facilitate the creation and distribution of content easily on a global scale, and recommendation algorithms often amplify disinformation by prioritising sensationalist and polarising content.¹⁹

The importance of disinformation lies in its potential impact on the quality of public debate and the political sphere, due to its influence on elections and its contribution to the polarisation of society.²⁰

This last aspect constitutes the backbone of the analysis, since various studies consider that this new facet, that is, political disinformation in digital contexts, began to be studied from 2016 onwards, after Trump's presidential campaign or the Brexit referendum²¹ 'in which it was made visible how technology can be used to influence public opinion and undermine trust in democratic institutions'.²²

¹⁷ Special Rapporteurship for Freedom of Expression of the Inter-American Commission on Human Rights with contributions from the Department of Electoral Cooperation and Observation and the Department of International Law of the General Secretariat of the Organization of American States (n 15) 13–14.

¹⁸ Silvio Waisbord, 'Truth is What Happens to News' (2018) 19(13) Journalism Studies 1866 <https://is.muni.cz/el/fss/podzim2019/ZURb1608/um/readings/week_10/Truth_is_What_Happens_to_News.pdf> accessed 18 January 2026.

¹⁹ Rubio Nuñez, García Rossó and Sánchez (n 11) 2.

²⁰ Viktoriia Vovkodav, Luis Sendino Marcos and Adrián Cordero Redondo, 'Fake news y su impacto en la sociedad: una revisión de la literatura' (2022) 13(1) Pangea Revista de la Red Académica Iberoamericana de Comunicación 45 <<https://revistapangea.org/index.php/revista/article/view/220>> accessed 18 January 2026.

²¹ Francisco Sierra Caballero and Salomé Sola Morales, 'Golpes mediáticos y desinformación en la era digital. La guerra irregular en América Latina' (2020) 17 Community and Society Guadalajara e7604 <http://www.scielo.org.mx/scielo.php?script=sci_art-text&pid=S0188-252X2020000100121&lng=es&nrm=iso> accessed 17 January 2026; López López, Mila Maldonado and Ribeiro (n 16).

²² Rubio Nuñez, García Rossó and Sánchez (n 11).

1.2 Is talking about fake news the same as talking about disinformation?

In the collective imagination, the term ‘fake news’ is synonymous with ‘disinformation’. However, this term simplifies or obscures a much broader spectrum of manifestations. This point has been explored by academics who have asked themselves the same question as those reading these pages: If disinformation isn’t just ‘fake news’, then what is it? What are we talking about when we refer to this concept?

In this sense, there is agreement in the literature about those who proclaim that the phenomenon of disinformation (*information disorder*) is a broader concept than that evoked by the term *fake news*, although this word has popularly gained more traction to designate fraudulent informational content.²³ Thus, this phenomenon refers both to fraudulent informational content (*fake news*) as to deceitful (*misleading content*), hate speech (*malinformation*), deliberate false speeches (*false speech*) or unintentional misinformation by media outlets or journalists (*misinformation*).

It is worth noting that the term popularly known as fake news is an oxymoron, since it refers to a ‘news that is not real’ when what is false is not news, but a fictional narrative.²⁴

In this sense, the definition by Calvo and Aruguete stands out:

Indeed, the politics of fake news must be understood not only as an act of transmitting information but also as a performative act: an expressive event that seeks to inflict harm on an opponent. While the partisan ideologue may have difficulty interpreting an event, the troll, rather than defending their interpretation of events, seeks to cause harm. The former protects their cognitive pleasure, the latter receives a political – or economic – reward from offending another.²⁵

²³ Carlos Rodríguez Pérez, ‘No diga fake news, di desinformación: una revisión sobre el fenómeno de las noticias falsas y sus implicaciones’ (2019) 40 Comunicación (Universidad Pontificia Bolivariana) 65 <<https://revistas.upb.edu.co/index.php/comunicacion/article/view/437>> accessed 16 January 2026.

²⁴ *ibid.*

²⁵ Ernesto Calvo and Natalia Aruguete, ‘Fake News, Trolls y otros encantos: cómo funcionan (para bien y para mal) las redes sociales’ (ed Siglo Veintiuno No 2020).

In relation to the above, other authors recognise types of on-line disinformation as mentioned above, but with certain nuances, emphasising their characteristics according to the intention or origin of the various types.

On the one hand, authors such as Rubio Nuñez, Garcia Rossó and Sánchez in their work on ‘Disinformation: how to understand it, combat it and protect yourself from its effects’ recognise different types of disinformation according to their origin: misinformation as false information shared without the intention of causing harm; disinformation as false information deliberately spread to deceive; and malinformation when information based on reality is used to inflict harm on a person, organization, or country’.

For its part, the Center for Studies in Technology and Society (CETyS) of the University of San Andrés (Argentina), in a recent study, has systematised these concepts according to the character of intention and authenticity:

Table 1. *Typologies of disinformation*²⁶

Term	Definition	Intention	Authenticity
False information	Verifiable false information		False
Misinformation	False information that is shared without the intention of deceiving or causing harm.	Not to deceive	False
Disinformation	False information that is shared intentionally to deceive	To deceive	False
Malinformation	Genuine information shared with the intent to cause harm	To cause harm	Genuine

In this sense it is worth noting that the categorisation also includes simple false information where no particular intention is identified and which can be subjected to verification.

²⁶ Iván Kirschbaum and Fernanda Martinez, ‘Desinformación en la Argentina: una perspectiva multisectorial’ (Centro de Estudios en Tecnología y Sociedad (CETyS) UdeSA 2024) <<https://dspaceapi.live.udesa.edu.ar/server/api/core/bitstreams/c6b7b460-6004-4191-b9c6-7de88f8c13ca/content>> accessed 18 January 2026.

Regarding the table shown, a study on the taxonomy of information disorder published by the Council of Europe coincides in the definition of ‘misinformation’ and ‘malinformation’ But when presenting what is meant by disinformation, it adds an element absent in the definition above: the intention to cause harm. That is, it no longer only seeks to intentionally deceive but to generate harm.²⁷

Thus, although the definition of what is meant by disinformation does not seem to be settled, there is agreement in treating it as the ‘distortion of information by spreading falsified news that deceives the final recipient’,²⁸ it is important to highlight that the types of phenomena linked to the spectrum of disinformation are distinguished, which would allow a better approach in the ways of understanding them and designing measures in pursuit of the protection of the subjects exposed to them.

Notwithstanding doctrinal conceptualisations, it is important to highlight that it has also been a topic on the agenda of international entities.

In 2018 the European Commission defined disinformation as ‘information verifiably false or misleading information that is created, presented and disseminated for profit or to deliberately deceive the public and that may cause public harm’.²⁹ This definition understands ‘public harm’ as threats to democratic political and policymaking processes, as well as to public goods, such as the protection of health, the environment or the safety of EU citizens.

Informational errors, satire and parody, and news and commentary clearly identified as partisan are excluded. It is worth noting that, up to this point, the conceptualisation carried out by the European Commission does not allow for other definitional possibilities; that is, there is some agreement on the objective of the

²⁷ Claire Wardle and Hossein Derakshan, ‘Information Disorder: Toward an interdisciplinary framework for research and policy making’ (Council of Europe Report DGI(2017)09 2017) <<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>> accessed 18 January 2026.

²⁸ Carlos Rodríguez Pérez, ‘No diga fake news, di desinformación: una revisión sobre el fenómeno de las noticias falsas y sus implicaciones’ (2019) 40 Comunicación (Universidad Pontificia Bolivariana) 65, 67. <<https://revistas.upb.edu.co/index.php/comunicacion/article/view/437>> accessed 16 January 2026.

²⁹ Code of Practice on Disinformation (European Commission 2018).

policies to be promoted in this area. However, it uses broad concepts such as ‘threats to democratic political processes’, which leaves considerable room for manoeuvre for both member states and the European Union.

In 2019, the Inter-American Commission on Human Rights (IACHR or Commission) understood disinformation as ‘the mass dissemination of false information (a) with the intention of deceiving the public and (b) knowing it to be false’,³⁰ not without first clarifying in a footnote that this definition is ‘provisional’.

It clarifies that the term ‘false information’ only applies to information that can be subjected to verification or cross-checking, and excludes opinions and editorial approaches that may be shocking or misleading, or that, being opinions, are not susceptible to verification or truth assessment. An important clarification made by the Commission for the purposes of this work is that certain malicious editing practices of true content could fall within the definition of ‘false information’ if they seek to deceive the general public and falsify the original content – for example, slowing down a person’s speech to make them appear intoxicated.³¹

For her part, the United Nations (UN) Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, in her reports on ‘Disinformation and freedom of opinion and expression’, has concluded that disinformation lacks a clear definition and a consistent understanding of its meaning.³² This view was subsequently echoed by the UN Secretary-General in his 2022 report on ‘Countering disinformation to promote and protect human rights and fundamental freedoms’.³³

It includes the following conceptualisations:

³⁰ Special Rapporteurship for Freedom of Expression of the Inter-American Commission on Human Rights with contributions from the Department of Electoral Cooperation and Observation and the Department of International Law of the General Secretariat of the Organization of American States (n 15) 13.

³¹ *ibid.*

³² United Nations General Assembly Human Rights Council, ‘Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan. Disinformation and freedom of opinion and expression’ (2021) UN Doc A/HRC/47/25.

³³ United Nations General Assembly, ‘Report of the Secretary-General. Countering disinformation to promote and protect human rights and fundamental freedoms’ (2022) UN Doc A/77/287.

There is no clear definition of the term ‘disinformation,’ nor is there a common understanding or position regarding its meaning. It has been used by the International Telecommunication Union and the United Nations Educational, Scientific and Cultural Organization (UNESCO) to describe false or misleading content that can cause specific harm, regardless of motivations, awareness, or behavior. The Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression also defines disinformation as ‘false information that is intentionally disseminated to cause serious social harm’. The studies and contributions examined for this report point to the following characteristic elements of disinformation: inaccurate information, information intended to deceive, and information shared with the aim of causing serious harm.³⁴

According to this report, common elements could be drawn from the concepts presented: a) inaccurate information b) the intention to deceive c) that it is shared with the purpose of causing serious harm.

In summary, and based on the instruments presented, it is possible to highlight some aspects:

The studies consulted do not provide a universally accepted definition. However, they show agreement and concurrence regarding two elements: inaccurate information and the intent to deceive. The element on which they differ is the intent to cause harm.

International organisations also do not provide a concrete definition. In this respect, only the European Union has developed a conceptualisation that is not subject to variations, in a quasi-closed manner; however, it adopts ambiguous criteria.

There is no widespread agreement on disinformation in the Americas, which means that state action to mitigate disinformation is fragmented.

For the purposes of this work, disinformation will be understood as false information that is shared to intentionally deceive.

³⁴ United Nations General Assembly (n 33) 2 para 3.

That said, it is important to emphasise that despite these schematic classifications, there is a consensus that it should be analysed multifacetedly or treated as a ‘multicausal phenomenon’. This means that it is influenced by contextual and psychological factors, as well as by the design of the platforms.

In addition to this approach, studies highlight its use as a political tool.³⁵ In this sense, it stands out as the main instrument in political struggle, as it can be used to influence people’s perceptions, especially those who have low faith in democratic institutions or processes such as elections. At the same time, disinformation should be framed as a technological phenomenon, since technology has revolutionised public opinion.³⁶ Information is the raw material of social organisation and governance. It is capable of influencing societal behaviour. It is power.

Technology introduces tools that condition the subject-information relationship, which makes misinformation more harmful these days.

This dynamic, of disinformation as a technological phenomenon, is described by Rubio Nuñez, Garcia Rossó and Sánchez:³⁷

- The structural nature: This phenomenon, as explained above, can be negligent, malicious or innocent, but it can cause harm to people and the system by altering the perception and, therefore, individual and collective self-government. It is there that the authors consider it a threat to democracy.
- That power to change perception, to define what is reality, is achieved through disinformation. This manipulation of perception is achieved through constant, disciplined and organised actions. These actions require both human and material resources to form the ‘armies’ that exclusively carry out the task of disinformation.

³⁵ Vovkodav, Sendino Marcos and Cordero Redondo (n 20).

³⁶ Rubio Nuñez, Garcia Rossó and Sánchez (n 11).

³⁷ *ibid.*

- Social components: On the one hand, there is mention of the selective consumption of what is ‘true’ due to the excess of information, driven by the algorithms of the platforms that filter content according to preferences, which ultimately design what we will consume as our own information diet.

Information overload leads to information inflation, horizontalisation and consequently, the loss of authority of traditional media or information sources.

Among these social factors or barriers is the lack of attention to the overwhelming amount of information. This is compounded by the ‘novelty bias’, which is simply the element of time. Due to the rapid pace of information circulation as input for understanding reality, novelty is prioritised over accuracy. Publication must be rapid, even more so when there are events of high public interest, such as presidential elections or health crises like COVID-19, times when misinformation proliferates.³⁸

Information overload, coupled with the need for and immediacy in the reception and distribution of news, leads to the easy spread of misinformation. According to Rubio Nuñez, García Rossó and Sánchez, this causes ‘the loss of the temporal horizon of politics, which becomes an eternal present, where every statement, every act, seems to start from scratch, without any conditioning in the past and without any obligation for the future’.³⁹

This production and dissemination of information is decentralised, multidirectional, multichannel and participatory. It does not originate from a single source or an exclusive platform. It is within this context that the hyperconnectivity of users is leveraged, as their interconnectedness contributes to the dissemination or production of misinformation, thus hindering the determination of its origin. This dissemination is based on interpersonal trust, which is stronger than trust in organisations or institutions (including the media).⁴⁰

³⁸ Kirschbaum and Martínez (n 26).

³⁹ Rubio Nuñez, García Rossó and Sánchez (n 11) 3.

⁴⁰ *ibid.*

Another element to consider, especially for the purposes of this thesis, is the weight of the image. This element is predominant when dealing with disinformation, as it involves the capacity for abstraction and reinforces the emotional element of perceptions and decision-making processes. The image is the heart of the information, and the text is merely supporting. The content is simplified and operates within the realm of emotions.⁴¹

Within the psychological conditions, related to the mental-cognitive and affective processes in which the behavioural sciences focus on understanding why the individual is part of the disinformation process.

In this regard, Calvo and Aruguete refer to the existence of confirmation bias among users, which favours the consumption of data that coincides with and confirms their prior viewpoints, beliefs and opinions.⁴² It is worth noting that the platforms themselves, by design, contribute to these biases, as they present information to users according to what the algorithms deem relevant to their digital profiles.

In this sense, Rubio Nuñez, Garcia Rossó and Sánchez have listed a series of biases that come into play:

Table 2. Biases⁴³

Biases	What do they imply?
Confirmation	The search for and acceptance of data that confirms prior beliefs causes users to lean towards them, ignoring those that differ from or challenge their points of view.
Anchoring	The first information they receive is the first impression, although incorrect, on which they base their decisions and which can influence the consideration or evaluation of subsequent information.
Availability heuristic	It causes people to overestimate the likelihood of events based on easily remembered examples. Disinformation uses striking and memorable examples to make false information seem more frequent or probable.
Of affection	It refers to the fact that the evaluation of information and the decisions we make based on it are influenced by emotions rather than a rational evaluation.

⁴¹ Rubio Nuñez, Garcia Rossó and Sánchez (n 11).

⁴² Calvo and Aruguete (n 25).

⁴³ Rubio Nuñez, Garcia Rossó and Sánchez (n 11).

Repetition	The most repeated information is the most credible, or at least it can make it be perceived as such. The more we are exposed to an idea, the more likely we are to accept it as true.
Social pressure and conformity	Social pressure and the desire to conform to the group also play an important role in the acceptance of misinformation. People may accept and spread false information to align themselves with the opinions of the group to which they belong.

1.3 ABC of disinformation

How does Welles's new 'war of the two worlds' spread?

Disinformation spreads through platforms, and three characteristic vectors of this type of 'viral deception' are recognised. They are manipulative actors, deceptive behaviour and harmful content.⁴⁴ In this sense, it is recognised that the centrality each of these vectors has received varies. For example, the technology industry has invested in identifying the behaviour of coordinated and inauthentic accounts,⁴⁵ while governments have been more concerned with content, that is, what is published on social media. But what do these vectors consist of?

1.3.1 Manipulative actors

As Camille F Graphika argues, manipulative actors knowingly and intentionally participate in viral deception campaigns.⁴⁶ Their campaigns are 'covert', designed to conceal the identity and intentions of the actor orchestrating them. They are capable of manipulating public discourse. She adds that these are more accurately described as 'advanced manipulative actors', since they include military and AI apparatuses, or campaign equipment.

In this regard, one problem evident in the field of actors is that their detection is often not addressed by the terms of service and community guidelines, or there is little clarity regarding support for detection techniques. At this point, a central issue arises in the discussion: anonymity, yes or no? It is reasonable to question whether the validity of anonymity on social networks can be a negotiable matter in the context of disinformation.

⁴⁴ Camille Francois Graphika, 'Actors, Behaviors, Content: A Disinformation ABC. Highlighting Three Vectors of Viral Deception to Guide Industry and Regulatory Responses' (Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression 2019).

⁴⁵ *ibid.*

⁴⁶ *ibid.*

Graphika responds that the problem of manipulative actors has little to do with online anonymity for two reasons: on the one hand, anonymity allows the voices of the vulnerable to be protected and to participate critically in public debate, and on the other hand, manipulative actors respond to a structural logic; it is not a problem that can be reduced to the individuality of ‘a subject from the armchair of his house’.

Finally, it is important to consider the role of states in relation to manipulative actors, as they can covertly use platform services to carry out influence campaigns. Graphika cites as an example the promotion of content critical of the Hong Kong protests by the Chinese state. While in this case Twitter, the platform where these campaigns were disseminated, announced that it would not allow ‘state-controlled media’ to use its advertising products, this⁴⁷ opens a discussion about how platforms should distinguish between privately funded entities, independent public broadcasters and those that are financially or editorially controlled by the state.

In addition, mention should be made of the covert practice of states, in order to hide their direct participation in media that present themselves as independent, for example the Balt News network controlled by Russia.

The paradox is that just as states can act as disinformation agents, the detection and mitigation of the damage caused by manipulative actors is also in their hands.⁴⁸

1.3.2 Deceptive behaviour

This section groups together the techniques that manipulative actors can deploy to increase the reach and virality of content. These include automated tools (such as the armies of bots used to amplify the reach and effect of a message) to manual deception (such as paid participation or troll farms). The objective of this type of behaviour ‘is to allow a small number of actors to have the perceived impact that a larger number of actors would have if the campaign were organic’. The types of inputs used as a deceptive behavioural technique are:

⁴⁷ Camille Francois Graphika, ‘Actors, Behaviors, Content: A Disinformation ABC. Highlighting Three Vectors of Viral Deception to Guide Industry and Regulatory Responses’ (Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression 2019), 3.

⁴⁸ *ibid* 4.

- **Bots:** These are automated social media accounts managed by an algorithm, not a real person. The bot is designed to post content without human intervention.⁴⁹ Key engineering characteristics include anonymity, high levels of activity and the amplification of specific users, topics or hashtags. This type of automation can be grouped into a network of accounts managed by the same individual. This network is known as a ‘botnet’.
- **Trolls:** Unlike the previous example, a troll is someone who intentionally starts online conflicts or offends other users to sow division by posting inflammatory or off-topic statements in a digital community or social network. Their goal is to exploit emotional bias, as this deceptive behaviour generates a response without rational evaluation and redirects the focus of discussions.

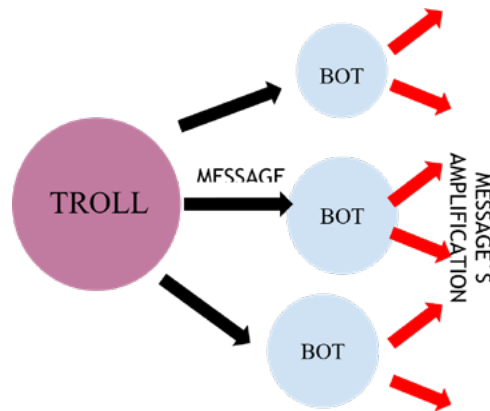
While trolls and bots are mutually exclusive because they are not the same, there is a dynamic between them, since trolls use bots to amplify their messages; in this way, automation (bots) becomes part of the trolling.

However, it should be mentioned that there are studies that point to the dynamic between bots and humans in disinformation. In this regard, González Bailon and De Domenico emphasise that while studies determine that on Twitter (now X) a large proportion of accounts are bots responsible for disinformation during election periods, while simultaneously exacerbating political conflicts, bots are vectors that amplify both true and false information, and that verified Twitter accounts, which appear to be more trustworthy, are actually more likely to receive (disinformation) retweets, and in that intelligence, it is humans who tend to spread and amplify the difference in visibility of fake news through such retweets.⁵⁰

⁴⁹ Donara Barojan, “#TrollTracker: Bots, Botnets, and Trolls” (DFRLab, 9 October 2018) <<https://medium.com/dfriab/trolltracker-bots-botnets-and-trolls-31d2bdbf4c13>> accessed 17 January 2026.

⁵⁰ S González-Bailon and M De Domenico, ‘Bots are less central than verified accounts during contentious political events’ (2018) 118(11) *Proceedings of the National Academy of Sciences* e2013443118 <<https://www.pnas.org/doi/epub/10.1073/pnas.2013443118>> accessed 18 January 2026.

Figure 1. Trolls, Bots and disinformation dynamics⁵¹



As mentioned, bot networks can be managed by a single individual or group. The administrators of such bot networks are called botnets and bot shepherds. Regarding this, Donara Baroja writes:

Bots are designed to manufacture engagement that makes it seem as if there are more 'real' users in the topics they're involved in than there actually are. On social media platforms, engagement generates more interaction, so a successful botnet manages to get the topics it's involved in to reach more real users.⁵²

The engagement is the concept used to describe the process by which a bond of commitment (usually from a brand) is generated with a community of followers or users on social media. This is achieved through certain actions performed by bots, such as interacting with these users, or by creating a sense of belonging or identification with the message recipient through a specific narrative or message.

⁵¹ Self produced scheme based on Barojan (n 49).

⁵² *ibid.*

Therefore, the bot's task is to make a hashtag, keyword or user appear more discussed, either positively or negatively depending on the objective, or more popular than they actually are. They aim to influence the 'trending' section, allowing them to expose a discussion started by a troll to other users, even those who don't follow the accounts in question.

The Digital Forensic Research Lab (DFRLab) has analysed a pattern based on the observation of large botnets on Twitter. In this regard, the aspects or indicators that stand out are:

*Table 3. Botnets patterns*⁵³

Pattern	Characteristics
Discourse pattern	Because they are run by an algorithm, bots are programmed to repeat the same discourse. This indicator can take the form of the same hashtags, the same tagged users or the same headlines on Twitter.
Identical publications	As automated products, bots repeat the same message. They do not produce authentic content.
Patterns in the alias or user	Bot creators often use the same pattern to name the alias or username of 'suspicious accounts'. The study of DFRLab cites as an example the same number of digits in the username.
Date and time of creation	A single botnet tends to share a similar creation date, the same day or week.
Identical activity on Twitter	There are certain tasks or ways to get involved in the network or platform, which increase the likelihood of becoming part of that network. In the study (DFRLab) is mentioned as an example of this indicator that different accounts follow three identical but unconnected accounts or without any link.
Location	Especially in political matters, bots or accounts categorised as 'suspicious' often share a common location. The 'shepherds' of political botnets or political bots tend to use the location where the candidate or party they are promoting is running to try to make it a trend in that particular electoral district.

To this list of patterns, we should add those identified in 'Dataset Americas #1: Honduras' by the Honduran fact-checking company Cazadores de Fake News (Fake News Hunters). According to data published by Twitter, during the period from August 2020 to November 2021, a greater number of accounts were observed

operating simultaneously and promoting common narratives, which corresponds to the first two classifications presented in the table above. However, profile avatars or photos are also analysed here as another indicator of anomalous accounts.

Thus, the faces in the photographs matched in iris position or errors or ‘glitches’ such as asymmetries in the ears, different earrings, partial deformations in faces of people outside the focused face, incongruous backgrounds, etc.⁵⁴

For her part, Graphika raises a problem regarding the disparities in the definition that technology companies give to disinformation, ergo to this type of deceptive mechanics or techniques, as well as the difficulties that arise in detecting and mitigating them, since they tend to be opaque to the public and/or can have implications for freedom of expression by reiterating content and social media accounts, especially when it comes to political content during election cycles.⁵⁵

1.3.3 Harmful or detrimental content

This vector is the most visible, because perhaps it cannot be attributed to a ‘manipulative actor’ or reveal the deceptive techniques or behaviours behind the content, but it is easily observable by the user.

This issue demands a solution for detecting and implementing strategies based on what is known as ‘content moderation’. As previously stated, governments often prioritise content and its regulation. The problem here is that legal and regulatory frameworks often struggle to define what constitutes ‘harmful content’, and the line between what is considered ‘harmful’ and discourse protected by international human rights law is blurred.

There are different types of synergy between harmful content and misinformation:

- To consider an entire category of content as harmful because it belongs to the realm of viral deception, such as ‘health misinformation’.⁵⁶

⁵⁴ Cazadores de Fake News, ‘Operaciones de Influencia en Honduras: Lo que debes saber sobre el dataset publicado por Twitter’ (Cazadores, 26 August 2022) <<https://www.cazadoresdefakenews.info/honduras-lo-que-debes-saber-operaciones-de-influencia-twitter/>> accessed 17 January 2026.

⁵⁵ Graphika (n 44).

⁵⁶ *ibid.*

- The content of a campaign (not its dissemination mechanism) itself can be manipulated to deceive users, and therefore belong to the realm of disinformation (for example, the use of manipulated media such as deepfakes).
- Harmful content is promoted by deceptive actors or campaigns that use distorting behaviours such as troll farms.

It is worth noting that within this concept of ‘harmful content’, various types of categories are linked, such as incitement to hatred, harassment and violent extremism.

1.4 Electoral disinformation and social networks

It is pertinent to draw upon Rafael Rubio Nuñez’s analysis of electoral campaigns. The author understands the electoral campaign as a specific period within the electoral process in which the aim is ‘a massive dissemination of information at no cost to the voter, in terms of both time and money’. From a traditional perspective, this moment is a symbolic dialogue between voters and candidates, between those who hold popular representation and those they represent.

Indeed, that traditional view is being redefined in light of the revolutionary introduction of technology. The author identifies a new stage in election campaigns:

- *Premodern campaigns*: This period, spanning the 19th to the early 20th century, was characterised by its personal nature. Technology was limited to a platform at the back of a train and the bare minimum needed to produce and distribute printed newspapers, primarily local ones. Election campaigns during this time relied on rallies and grassroots, personal mobilisation, rewarding party structures and diminishing the figure of a leader whom the vast majority of voters could neither see nor hear throughout the campaign.
- *Modern campaigns*: This stage began with the advent of radio and television (from the 1920s to the 1980s). Information circulated through these mass media. This increased the opportunities to learn about political proposals. Furthermore,

greater visibility allowed for a more personalised approach. Rubio Nuñez adds that this stage fostered a certain dependence on technical professionals and a greater emotional reliance on voter turnout.

- *Postmodern campaigns*: From the 1980s onward, the emergence of information technologies led to the incorporation of persuasive elements based on a better understanding of target audiences, their motivations, their logic and a greater technical capacity to deliver the message to them (segmentation). Here, technology can be used as a tool for mass communication.

In the realm of disinformation, particularly electoral disinformation, though not exclusively so, one notable feature is that hyperconnectivity, presented as one of the social elements that contribute to the phenomenon of disinformation, impacts the ‘crisis of intermediation’. This crisis stems from the questioning of traditional mechanisms of representation, which allows for the emergence of new media, generally technological.

The significance of this categorisation is that with the development of AI, a door opens to a fourth stage, named algorithmic campaigns.⁵⁷

Along these lines, García Orosa argues that in the ecology of digital democracy, four major waves can be observed:⁵⁸

- First wave: late 20th and early 21st centuries: presence of political actors on the network without major changes in the principles of participation, mobilisation and organisation.
- Second wave: until 2008 characterised by the massive emergence of social networks and digital political marketing.
- Third wave with the use of Big Data and micro-segmentation for decision-making.
- Fourth wave: since 2016 with the use of AI in election campaigns.

⁵⁷ Rubio Nuñez, García Rossó and Sánchez (n 11) 70.

⁵⁸ Cited by López López, Mila Maldonado and Ribeiro (n 16).

Although this periodisation offered by García Orosa does not strictly coincide with the periodisation, it does highlight conceptual characteristics of four eras or waves in relation to digital dynamics.

1.4.1 How does disinformation operate in the political arena?

The *modus operandi* of disinformation in general manifests itself in its political form. In this sense, bots, trolls and botnets are also used as deceptive techniques. It is worth noting that not all bots involved in the political sphere share this same nature. That is, bots can be commercial accounts hired to promote political content⁵⁹ or they can be purely political, that is, created with the sole purpose of amplifying the political content of a particular party, candidate, interest group or points of view.⁶⁰ This will be addressed in the following pages when considering political consultants as actors of disinformation.

Furthermore, there is a propaganda logic based on ‘viralisation’ in which political lies are prevalent and normalised in electoral campaigns.⁶¹

1.5 Electoral disinformation in Latin America

In the context of what is presented as ‘Post Truth’, favoured or created by disinformation, it is characterised by certain features such as the discrediting of journalists and media.

In her publication ‘Politicization of Disinformation in Contexts of Devalued Information: The Case of Latin America’, Adriana Suárez argues that these characteristics embody a long-established practice among officials, regardless of their rank, in Latin America. In other words, the authorities themselves systematically disseminate the idea that the media and journalists spread disinformation, while simultaneously establishing this as official truth through the state apparatus, which includes institutional media and pro-government think tanks.

⁵⁹ Barojan (n 49).

⁶⁰ *ibid.*

⁶¹ Dader and Campos (2017) cited by López López, Mila Maldonado and Ribeiro (n 16).

In this sense, the author argues that Latin America can contribute elements to understanding disinformation in contexts where power is concentrated in the executive branch without adequate checks and balances.⁶² Thus, she posits that in these contexts, disinformation is fostered not only by the discrediting of traditional media but also by a landscape of ‘devalued information’ comprised of severe restrictions on freedom of expression, a lack of transparency in public data and difficulty accessing reliable information for verification, low levels of trust in and consumption of news, as well as the relativisation of sources.

Added to this is the multiplicity of media and messages of varying quality as well as the instrumentalisation of emotions, common in Latin American populisms.⁶³

Rubio Nuñez, García Rossó and Sánchez argue that disinformation poses a critical threat to democratic stability in Latin America.⁶⁴ It has a pernicious impact on electoral processes, as it seeks to influence them and generate distrust in public institutions. Among the examples highlighted by the authors are the presidential campaigns in Brazil in 2018 and 2022, and in Mexico in 2024.

As already mentioned, disinformation represents a low cost and high return in elections⁶⁵ so its use in Latin America is not random. Another dynamic these authors highlight is foreign intervention in disinformation campaigns; they mention the detection of 7,000 anomalous accounts operated from abroad with servers in Russia and China.

Like Adriana Suarez, the authors mention that in countries with authoritarian regimes, disinformation enables the justification of controlling information on social media, along with censorship and threats against journalists. Disinformation is used to divert attention and hide the truth, or to silence political dissent.

⁶² Adriana Amado Suarez, ‘Politización de la desinformación en contextos de información devaluada. El caso de Latinoamérica’ (2022) 4(17) *Revista Internacional de Comunicación y Desarrollo* <<https://doi.org/10.15304/ried.4.17.8793>>.

⁶³ Manfredi Amado and Gómez-Iniesta (2022) cited by Suarez (n 62).

⁶⁴ Rubio Nuñez, García Rossó and Sánchez (n 11).

⁶⁵ *ibid* 70.

1.6 Social actors of disinformation

The following briefly presents entities whose participation should be considered in the digital information ecosystem. It is important to emphasise that these should not be understood as isolated compartments, but rather as constantly interacting entities that disseminate (mis)information:

1.6.1 Citizenship

Users also become actors in the spread of disinformation, since it is through ‘viralisation’ that this phenomenon manages to propagate. This viralisation, among other factors, is the result of the response and the time between a stimulus (for example, a post on Twitter) and the retweet or share (response), depending on the user’s preferences or beliefs.⁶⁶

In this sense, Pedro Anguita and others⁶⁷ collects studies that establish that exposure to and credibility of false content are statistically significant predictors of whether a person will end up sharing and spreading it.

The authors present a study on the profile of users in Chile, showing that users would believe and share misinformation according to the following variables: personal and psychological factors (such as belief in conspiracy theories, trust in one’s own contacts, level of education and gender), specific uses of social networks and certain political opinions and attitudes (people who are more politically active in digital spaces and those more inclined to the right believe and share more misinformation).

Likewise, citizens participate in digital social networks without filters, verification mechanisms or controls. Furthermore, it’s worth mentioning that this group also includes users who knowingly and consciously spread misinformation, which again leads to polarisation.

However, within this group, not all the roles played by citizens are negative; they also manage to organise themselves against disinformation: by promoting regulations from political spheres, or by carrying out fact-checking, whether with public or private funding.

⁶⁶ López López, Mila Maldonado and Ribeiro (n 16).

⁶⁷ P Anguita and others, ‘El fenómeno de la desinformación: Revisión de experiencias internacionales y en Chile’ (Comité Asesor contra la Desinformación; Ministerio de Ciencia, Tecnología, Conocimiento e Innovación 2023) 3.

1.6.2 Classic or traditional media

These are central actors in political communication as well as in disinformation.⁶⁸ Although today the landscape is more ‘horizontal’ in terms of the dissemination of ‘information’ (ergo disinformation), it cannot be denied that media power still holds the capacity to determine ‘reality’.

As author José Pablo Feinmann argues in his book *Filosofía Política del Poder Mediático* [*Political Philosophy of Media Power*], someone can create everyday reality through oligopolistic media power. This reality can take the form of ‘common sense’, which is nothing more than what everyone agrees is true, and is even more than truth itself.⁶⁹

The author explains:

Millions of people speak the same way, think the same way, talk about the same things, believe that what they say, what they hear, what everyone agrees on, is the truth in its simplest, most everyday sense, the truth that everyone agrees on without questioning it. That is called common sense.⁷⁰

According to the writer, media power manages to shape common sense as a steamroller of the freedom of subjects, since it imposes a ‘system of values’ shared by all; whoever is outside of it is not within the spectrum of ‘normal’.

Anyone who lacks common sense falls under a thousand suspicions. They’re either crazy. Or they’re high. Or they’re on hard drugs. Or they’re a fool. A blockhead ... A stubborn person. A weirdo. Being ‘within’ common sense reassures people. It calms them. They’re not crazy. They belong to a community of healthy, sensible, grounded, and trustworthy people.⁷¹

⁶⁸ López López, Mila Maldonado and Ribeiro (n 16).

⁶⁹ Feinmann (n 10) 83.

⁷⁰ *ibid.*

⁷¹ *ibid.*

It is worth noting that cases of disinformation by traditional media ‘mostly stem from poor journalistic practices, a lack of controls, or giving excessive prominence to sources and actors’.⁷² This context, compounded by the denigration of media outlets by leaders in Latin American countries – especially⁷³ – influences the perception of the media as disseminators of disinformation.

Like citizens, the media have also introduced mechanisms to ‘fight’ disinformation, such as verification in the production and circulation of news (*news-making*) and the introduction of specific personnel or departments intended for *fact-checking*.⁷⁴

Although their role and power as builders of reality remain relevant, authors like Silvio Waisbord argue that today, the creation of meaning is not limited to journalism or specific news companies.⁷⁵ Rather, reality, truth, is based on the way multiple actors actively engage with a vast amount of news and information of diverse nature. In this sense, it is worth considering the actors mentioned below, key players in the digital world: the platforms.

1.6.3 Platforms

It is undeniable that the advent of platforms has changed the information landscape and the role that information and news play within it. Thus, in the book *The Power of Platforms: Shaping Media and Society*, Kleis Nielsen and Anne Ganter argue that large companies such as Facebook, Google, Amazon, Apple and Microsoft define the way the internet works and therefore influence the structure of the entire digital communication ecosystem.⁷⁶

Specifically, Facebook, Google and Twitter are all examples of companies relevant to the production and distribution of news, and therefore information. Furthermore, these types of platforms establish standards, both social and technical, for those who wish to participate in the networks and markets enabled by these platforms today.

⁷² López López, Mila Maldonado and Ribeiro (n 16) 72.

⁷³ Suarez (n 62).

⁷⁴ López López, Mila Maldonado and Ribeiro (n 16) 72.

⁷⁵ Waisbord (n 18).

⁷⁶ Rasmus Kleis Nielsen and Sarah Anne Ganter, *The Power of Platforms: Shaping Media and Society* (OUP 2022) 8 <<https://doi.org/10.1093/oso/9780190908850.001.0001>>.

The authors emphasise the idea that platforms do not control the means of production but rather the means of connection.⁷⁷ In the information ecosystem, companies not only gain economic benefits such as advertising, attention and popularity by acting as intermediaries between users – whose numbers are ever-increasing – and third parties. For this reason, the authors understand that platforms occupy a dual position: they are the market and they are in the market. This position enables conflicts of interest that warrant governance rules to balance these conflicts of interest among the platforms, users and other parties.

In light of the above, where it is determined that the platforms have the ability to determine how the internet works and how the aforementioned means of connection are executed, it is important to question what responsibility they bear in the process of disinformation, which entails engagement, virality and monetisation.

In this sense, tensions between the (de) regulations from platforms like X or Meta, such as the latter's decision to change its content policies by eliminating the fact-checking programme and modifying its content moderation system, while countries like Brazil are emphasising the digital responsibilities of these platforms regarding content generated by their users.

In this regard, the Brazilian Supreme Federal Court established that companies such as X, Facebook, Instagram and TikTok must immediately remove publications that promote anti-democratic acts or hate speech without a court order, understanding that the companies' self-regulation has failed since their business model is designed to promote content that generates more interaction, even if it is detrimental to access to truthful information and democracy.⁷⁸

1.6.4 States and political parties

'States and intelligence agencies flooded social media with disinformation intended to sow confusion and influence public opinion and election results.'⁷⁹ These lines are relevant to illustrating the role that governments and political parties play in the disinformation ecosystem.

⁷⁷ Nielsen and Ganter (n 76) 160.

⁷⁸ According to Decision 06/26/2025 about Topic 533 and Topic 987 [Supreme Federal Court of Brazil].

⁷⁹ Bennet and Livingston (2018) cited by Waisbord (n 18).

This is how the actions of political parties are viewed, insofar as they constitute political organisations that may or may not hold political power (as a government) and equally contribute to, deploy or respond to the same disinformation operation.

That said, regarding the deployment of disinformation by state governments, Rubio Núñez mentions two cases: Russia's actions in Latin America (as a case of disinformation interference) and the case of Venezuela (as an intra-Latin American example, within its territory).⁸⁰ The case of Bolsonaro, as presented by Calvo and Aruguete, should also be added to the list.⁸¹

1.6.5 Russia in Latin America

Thus, the author states that since the first Information Security Doctrine, Russia has waged an information war or deployed its sharp power in the West through the use of the internet and global media, particularly exploiting narratives that favour its political interests.⁸² In Latin America, these interests are fuelled by resentment toward the United States and the presence of radical ideas that span the ideological spectrum. At the same time, Russia collaborates with autocratic allies to promote these very interests.

In this way, Russia establishes relationships with intellectuals and academic institutions to disseminate its vision, aided by regional media outlets such as Telesur and El Correo del Orinoco.⁸³ These outlets, more than their audience, offer another asset: their 'legitimacy' derived from their status as media outlets. This is what is known as 'information laundering' prior to its mass distribution through other alternative channels.

It is also mentioned that it uses other vectors of disinformation such as agents of influence, alternative journalism and social networks.⁸⁴

⁸⁰ Rubio Núñez, Garcia Rossó and Sánchez (n 11).

⁸¹ Calvo and Aruguete (n 25).

⁸² Rubio Núñez, Garcia Rossó and Sánchez (n 11) 6.

⁸³ *ibid* 7.

⁸⁴ *ibid*.

1.6.6 The case of Venezuela

The peculiarity mentioned by Rubio Nuñez, García Rossó and Sánchez is that the disinformation deployment is coordinated (*astroturfing*) and sophisticated. Furthermore, state entities play a crucial role in this dissemination. This stems from the closure of independent media outlets and the expansion of state media that replicated the official discourse.⁸⁵

Disinformation has served as a tool to silence dissenting voices and maintain control in contexts of polarisation. Even in the face of the migration to social networks, disinformation tactics were ‘renewed’; thus, it is understood that ‘Venezuela was a pioneer in the use of bots to create opinion trends on Twitter in 2010 and to harass opponents with verbal attacks since 2009’.⁸⁶

The study is based on a report from Oxford University, which states that Facebook categorises Venezuela as part of a group of countries with ‘high-capacity’ cyber-terrorist forces that have a ‘daily content target’ to influence public opinion. This group includes official, unofficial and some bot accounts.

Finally, it’s worth mentioning that during the last elections in 2024, disinformation *narratives* circulated, such as claims that the opposition candidate, Edmundo González, had serious health problems or terminal illnesses, as well as a lack of support for the opposition movement against Nicolás Maduro. However, despite this context, the study in question reveals that the ruling party leader (Nicolás Maduro) also suffered disinformation campaigns, including claims that he had lost the election.

1.6.7 The Bolsonaro case in Brazil

This example also demonstrates ‘*astroturfing*’. This concept refers to ‘a coordinated activity with the intention of creating the false impression that there is a widespread popular movement of support for or opposition to a candidate or party’.⁸⁷

⁸⁵ Rubio Nuñez, García Rossó and Sánchez (n 11) 9.

⁸⁶ *ibid.*

⁸⁷ Keller and others (2017); Llewellyn and others (2018) cited in Calvo and Aruguete (n 25).

Thus, Natalia Aruguete and Ernesto Calvo argue that in Brazil, during Jair Bolsonaro's 2018 campaign, artificial users and digital media were created to populate social networks and establish apparent support for security demands, among other proposals, on which the candidate based his electoral platform. These non-traditional digital media that spread the content, consisted of recently created media, and many of them were artificial.

On the other hand, regarding political parties and the introduction of disinformation as part of their political strategy, this is often coordinated by companies or individuals acting as 'political consultants'. In short, the politicians spread misinformation as part of their political strategy. The ways of doing so can be direct and indirect. The former refers to the use of fake news or spreading hoaxes through their own media⁸⁸ while the latter are related to inaccurate or biased information.

What types of parties promote disinformation?

This question must be addressed because several authors mention that, in the Latin American context, the ideology to which political parties respond constitutes a relevant factor, since 'the extreme right gains prominence through narratives based on hatred and the exclusion of others'.⁸⁹ Thus, the use of this 'land' becomes a political weapon both in the American continent and in the Iberian European area.

In this regard, the actions of Jair Bolsonaro are mentioned, who used a digital strategy based on the circulation of fake news, in addition to the use of hate speech⁹⁰ in the 2018 elections, at the end of his term during the Covid crisis and in the 2022 elections. This same deployment was observed in Peru linked to far-right movements, where fake news about vaccines and their effects has also been disseminated during said pandemic.⁹¹

⁸⁸ López López, Mila Maldonado and Ribeiro (n 16) 72.

⁸⁹ Rodríguez Serrano, García and Martín (2019) cited in López López, Mila Maldonado and Ribeiro (n 16).

⁹⁰ Moura and Verissimo (2019) cited in López López, Mila Maldonado and Ribeiro (n 16).

⁹¹ Gonzales (2022) cited in López López, Mila Maldonado and Ribeiro (n 16).

Authors -such as Paulo Lopez Lopez, Andrea Mila Maldonado and Vasco Ribeiro- mention that the global far right has a coordinated global strategy, which is evident in issues they oppose, such as feminism, migration and social movements, where they disseminate misleading or manipulated content. In this sense, Bolsonaro's performance is compared to that of Trump⁹² and to the Spanish far right, specifically Vox, in relation to their communication strategies, which are primarily linked to emotional appeals. While Bolsonaro's publications appealed to positivity and hope, the Vox leader's rhetoric was more closely tied to anger and disgust.

This, which would seem to be a characteristic feature of only America and the Iberian Peninsula, has its counterpart in the study carried out by Peter Tönberg and Juliana Chueri in their work 'When do parties lie? Misinformation and Radical-Right Populism across 26 Countries'.⁹³

In it, the authors analyse a database of 32 million tweets from 8,198 parliamentarians in 26 countries, during the years 2017–2022. These parliamentarians correspond to the member states and candidate states of the European Free Trade Association, where at least 45% of the parliamentarians are on Twitter. The group analysed includes European countries, the United States and Canada.

The conclusion they reach is that neither left-right ideological positions nor populist ideology alone explain the party members' propensity to spread misinformation on Twitter. In other words, the link between populism – the latter being associated with distrust of institutions such as mainstream media – and disinformation is not supported by the research.

Rather, the main problem of disinformation according to this research is not populism, but radical right-wing populism, which emerges in a crisis of legitimacy of democratic institutions⁹⁴ favoured by a period of neoliberal consensus, increased levels of inequality and growing power of business elites.⁹⁵

⁹² Arbuet and Cáceres (2019) cited in López López, Mila Maldonado and Ribeiro (n 16).

⁹³ Petter Tönberg and Juliana Chueri, 'When Do Parties Lie? Misinformation and Radical-Right Populism Across 26 Countries' (2025) 31(2) *The International Journal of Press/Politics* 367 <<https://doi.org/10.1177/19401612241311886>>.

⁹⁴ *ibid.*

⁹⁵ Crouch (2019) cited in *ibid.*

Thus, there is a link between radical right-wing populism and disinformation in the face of the loss of trust and legitimacy in conventional institutions, strategically exploiting vulnerability through disinformation to attack the media, governments and elections.

Something that should be highlighted is that the authors mention a kind of symbiotic relationship between these populist parties and the model of ‘*clickbait media*’.⁹⁶ This refers to a digital strategy that seeks to capture users’ attention through sensationalist headlines, generate curiosity or intrigue, and encourage them to click on a link to access specific content. This ‘attention economy’ model promotes the retention of user interest.

Radical right-wing populisms have created alternative media ecosystems that amplify their views, such as blogs or news websites, while also fostering a sense of community among followers and providing a counter-narrative to mainstream media reports, something that left-wing populists have failed to do effectively.⁹⁷

In sum, this research would allow us to show disinformation as an expression of the politics of radical right-wing parties. However, it should be noted that this study is based on metrics that Twitter no longer provides and focuses primarily on European countries. It is worth considering, while keeping in mind that, as Paulo Carlos López López and investigators have pointed out, there is a lack of research on disinformation in Latin America.⁹⁸

Finally, to this contextualisation it is important to add a journalistic investigation that, while not having the rigor of a scientific investigation, allows us to consider cases in Latin America in which disinformation deployments have been detected in electoral contexts as part of campaigns of parties associated with the ‘left’ based on the actions of political consultants (Buendía, Eduardo and others under the coordination of the Latin American Center for Journalistic Investigation (CLIP)).⁹⁹

This last point allows us to introduce another actor that must be considered in the scheme of dissemination of (mis) information:

⁹⁶ Munger (2020) cited in *ibid*.

⁹⁷ *ibid*.

⁹⁸ López López, Mila Maldonado and Ribeiro (n 16).

⁹⁹ Centro Latinoamericano de Investigación Periodística [CLIP], ‘*Mercenarios Digitales*’ (2023) <<https://www.elclip.org/mercenarios-digitales/>> accessed on 27 May 2025

1.6.8 Political consulting firms

In the disinformation ecosystem whose arena of dispute is currently social networks such as Facebook, Instagram or Twitter (now X), it is worth considering the role of political consultancies.

In this sense, a relevant example that invites us to think about the place these actors occupy in the political arena and the search for influence on the electorate, in conjunction with social networks, is given by the case of the then British consultancy Cambridge Analytica.

These actors are not frequently mentioned, as there is little consistent research on their actions, particularly in Latin America. However, there are journalistic compilations, such as those produced by CLIP, which, while not academic texts, provide a general overview of their possible operational logic.

1.7 What do consultancies do?

This group comprises both companies and individuals who advise candidates on how to present their ideas in a way that resonates with the electorate and wins their vote through catchy slogans, charisma and emotional appeals. However, the information and communication strategies they have implemented have loosened the ethical boundaries of their profession, a conclusion supported by the findings of the aforementioned research.

1.8 Strategies for disinformation

According to the CLIP journalistic study, the methods used by the aforementioned consulting firms include: purchasing or acquiring private data from any citizen without their knowledge in order to design messages that appeal to their emotions. They also resort to ‘warnings of danger’ to encourage the circulation of messages.

The report also mentions the creation of ‘fake opinion-makers on social media’, or as they have been called in previous pages: bots, to give the impression of a ‘majority’ in their favour and to set the agenda for traditional news outlets. It also describes the creation of numerous digital media outlets to act in unison to attack an ‘enemy’ and aggrandise the figure of the politician who uses their services.

These types of tactics act as a kind of match to ‘light the fuse’ in people on social media, in order to generate engagement and gain greater visibility on social media. As this study explains, they operate the ‘influence machine’ and contribute to exporting deceptive political agendas.

1.9 Research findings

The aforementioned CLIP investigation¹⁰⁰ deals with disinformation and propaganda operations in 15 Latin American countries where ‘political manipulators’ capable of influencing elections were found.

In this sense, to respect the length limitations of this work, interested readers are invited to delve into the lines of these articles, but not before mentioning that there are recorded cases of interference by political consultancies in the information ecosystem (mainly social networks) for political purposes through strategies such as image positioning, disinformation actions through bots or fake accounts that disappear after the electoral contests, creation of digital news programmes on social networks, profiling of citizens, media and ‘opinion leaders’ in Colombia, Guatemala, Argentina, Venezuela, Panama, Brazil and Chile.

Thus, this point attempts to raise as a topic for discussion and reflection: the responsibility and economic benefit that this type of activity represents for consulting companies or individuals, who end up influencing perhaps even more than politicians themselves in the dissemination of disinformation.

1.10 Conclusion of Chapter 1

Based on the foregoing, this first chapter outlines the main characteristics of the information ecosystem. It begins by defining the phenomenon of disinformation, the biases at play and the strategies used to spread it, before finally identifying the actors involved in disinformation. This analysis will allow us to understand how AI fits into this paradigm as a tool for disinformation.

2. Deepfakecracy: The foundations

Which one is real and which one is AI-generated?



*Only 3 are real.*¹⁰¹



The answers regarding what seems real and what appears to be the product of AI can vary. This depends on several factors, such as the source and how the information is presented. Therefore, let's consider the scenario outlined in the first two pages of this chapter as a constant reality in which citizens are immersed every day.

However, if this scenario, favoured by the incursion of AI, is evaluated in light of the disinformation ecosystem described in Chapter 1, the outlook becomes even more alarming.

Therefore, it is reasonable to question what role AI plays in the 'war for trust' proposed by Mark Zuckerberg, and even more so if that war unfolds in the American political-electoral arena.

For this reason, this chapter aims to establish a connection with the preceding pages in order to investigate what is meant by AI, its characteristics and how this relates to deepfakes.

All of this is aimed at fulfilling the second specific objective proposed in this thesis and building a bridge to Chapter 3 of this work, in which an attempt will be made to account for the role of the deepfakes in electoral contexts as a disinformation tool based on the analysis of cases in Latin America and the United States.

¹⁰¹ Some of the images shown belong to the test on the next page, Detect Fakes <<https://detectfakes.kellogg.northwestern.edu>> accessed 17 January 2026.

2.1 Artificial Intelligence

When discussing AI, different definitions can be found depending on the branch of knowledge addressing it, as well as the state of the art at the time of its description. In other words, there is no single, universally accepted definition. In fact, Stuart Russell and Peter Norvig, in their book *Artificial Intelligence: A Modern Approach*,¹⁰² considered a classic in the field, refer to eight different definitions from four approaches to AI, all of them accurate but distinct.

The truth is that there is some agreement in defining it as an area of computer science that ‘focuses on the development of systems and programs capable of performing tasks and solving problems in a way that differs from conventional programming methods and with results similar to those obtained by people’.¹⁰³ In other words, they are machines that attempt to simulate human intelligence.

While Chapter 4 of this work will be dedicated to studying the normative regulation from a human rights approach referring exclusively to the deepfakes, at this point, it is important to refer to the definition that has been attempted to be given to AI in general from the perspective of human rights.

In 2018, the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, in his 2018 report,¹⁰⁴ addressed the issue from a human rights perspective, defining AI as a ‘constellation of processes and technologies that enable computers to complement or replace specific tasks that would otherwise be performed by human beings, such as decision-making and problem-solving’.¹⁰⁵ Importantly, he emphasised the crucial role of human intervention in applying, selecting and labelling the data used by AI.

For their part, UNESCO, in its Recommendation on the Ethics of Artificial Intelligence, has defined that:

¹⁰² SJ Russell and P Norvig, *Inteligencia Artificial. Un Enfoque Moderno* (2nd edn, Pearson Educación SA 2004).

¹⁰³ Marcos J Gómez and others, ‘Diez preguntas frecuentes y urgentes sobre Inteligencia Artificial’ (Fundación Sadosky- Program AR 2024).

¹⁰⁴ United Nations General Assembly (UNGA), ‘Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression’ (2018) UN Doc A/73/348 <<https://docs.un.org/en/a/73/348>> accessed 17 January 2026.

¹⁰⁵ *ibid* para 3.

AI systems are information processing technologies that integrate models and algorithms, producing the ability to learn and perform cognitive tasks, resulting in outcomes such as prediction and decision-making in physical and virtual environments. AI systems are designed to operate with varying degrees of autonomy, through knowledge modeling and representation, data exploitation, and correlation calculations.¹⁰⁶

Another definition is that adopted by the European Union Regulation, since a) in accordance with the literature on the subject, it focuses on the protection of fundamental rights;¹⁰⁷ b) it has been adopted by the representatives of the states that make up this integration bloc, thus representing a unified definition at least for a group of states, and c) its definition is flexible enough to 'adapt to the continuous progress of this technology while maintaining an adequate level of precision, providing the legal certainty that any definition of a concept adopted in the legal order should have'.¹⁰⁸ Thus, in article 3(1), it defines what is meant by an 'AI system':

a machine-based system that is designed to operate with varying levels of autonomy and that can demonstrate adaptability after deployment, and that, for explicit or implicit objectives, infers from the input information it receives how to generate output results, such as predictions, content, recommendations, or decisions, that can influence physical or virtual environments.

In short, while there are basic and transversal characteristics of AI that will be discussed in these pages, the truth is that AI is not a single entity, but refers to a broad term that encompasses different technologies and fields of study.

¹⁰⁶ United Nations Educational, Scientific and Cultural Organization (UNESCO), 'Recommendation on the Ethics of Artificial Intelligence' (UNESCO 2021) <<https://unesdoc.unesco.org/ark:/48223/pf0000381137>> accessed 15 January 2026.

¹⁰⁷ Gergely Ferenc Lendvai and Gergely Gosztönyi, 'Deepfake y Desinformación- ¿Que puede hacer el derecho frente a las noticias falsas creadas por deepfake? (2024) 41 IDP 1.

¹⁰⁸ Mariana Sánchez Caparrós, 'Prevenir y controlar la discriminación algorítmica' (Rubinzal-Culzoni 2021) <<https://www.justierradelfuego.gov.ar/wp-content/uploads/2021/07/Sánchez-Caparrós.-Prevenir-y-controlar-la-discriminación-algorítmica.pdf>> accessed 17 January 2026.

2.1.1 A.I. When did we start talking about artificial intelligence?

Background

During World War II, with the creation of the first computers, computer science emerged. This encouraged scientists to devise algorithms, although it should be mentioned that mathematical methods that were precursors to AI had already been developed before their emergence.¹⁰⁹ In this sense, several articles highlight the role of Ada Byron, also known as Lady Lovelace, as the ‘mother’ of programming, since she is credited with creating the first algorithm intended to be processed by a machine, that is, the first computer programme.

This last fact is not recognised by academic texts, which instead portray the beginnings of AI as a product of the work of Alan Turing, nicknamed as one of the ‘fathers’ of AI or computer science.

In 1936, he published an article called ‘The Turing Machine’ where he defined an abstract computing machine, capable of performing any calculation through a programme (algorithm), which would allow the development of modern computing later.¹¹⁰

Genesis

As Russell and Norvig point out,¹¹¹ the genesis of AI (1943–1955) lies in the work of Warren McCulloch and Walter Pitts, considered the first work on AI. Inspired by the structure and physiological behaviour of neurons in the brain and Turing’s theory of computation, they proposed a basic model made up of artificial neurons, which, with the subsequent contributions of F Rosenblatt, was consolidated as a basic neural network that performs calculations to detect trends, patterns, classifications and information storage. Here, the neuron is ‘trained’ to discriminate between two linearly separable classes. This is the basis for what is now known as deep learning, which will be discussed later.

¹⁰⁹ César Federico Caiafa and Sergio Eduardo Lew, ‘¿Qué es la Inteligencia Artificial?’ (CONICET 2020).

¹¹⁰ *ibid.*

¹¹¹ Russell and Norvig (n 102).

The birth of AI

In 1950 he first spoke of the intelligence of an artificial system in his work ‘Computing Machinery and Intelligence’, in which he proposes the well-known Turing test.

Nevertheless, it is generally accepted that the true birth of AI corresponds to 1956, when the term was first coined at the Dartmouth College Summer Conference.¹¹² This conference consisted of a two-month workshop held that summer, aimed at increasing the interest of American researchers in automata theory, neural networks and the study of intelligence. John McCarthy proposed adopting the name AI for this new field, a choice that was widely accepted. This initiative also brought together the American figures who would dominate the field for the next 20 years.¹¹³

Russell and Norvig place the years of greatest hope and success in the early days of AI between 1952 and 1969. Likewise, they place the emergence of the first problems in the face of that promising future between 1966 and 1973, as the first AI systems had been used to solve simple problems but in most cases had failed when faced with more varied or difficult problems.¹¹⁴

Among the most significant questions are those raised by mathematicians such as Marvin Minsky regarding the *perceptrons* alluding to the fact that this simple neural network had a very limited representational capacity. These criticisms in the late 1960s halted the development of artificial neural networks for more than ten years.¹¹⁵ Furthermore, this period was accompanied by the withdrawal of funding for AI research.

The 1980s: Rise of AI

The cited literature agrees that AI grew rapidly from the 1980s onward, becoming an industry. The countries leading this exponential development are the United States, Japan and the United Kingdom.

¹¹² Moor (2006) cited in Caiafa and Lew (n 109).

¹¹³ Russell and Norvig (n 102).

¹¹⁴ *ibid* 24.

¹¹⁵ Caiafa and Lew (n 109).

On the other hand, in the field of neural networks, which had been abandoned by computer science, it gained a great boost with the work of Rumelhart, Hinton and Williams that advanced the idea of multilayer neural networks through a clear and practical methodology to be able to train them,¹¹⁶ thus overcoming the classification problems that were presented to perceptrons.

From 1987 to the present, the revolutionary field of AI has germinated not only in content but also in methodology, becoming established as a science.¹¹⁷

21st century: Deep learning

It was only in the 2000s that the construction of large multilayer networks could effectively be implemented to solve more complex problems such as image recognition, natural language processing, etc. Caiafa and Lew¹¹⁸ explain that this was possible due to two factors:

- A — Availability of large volumes of training data as a result of the proliferation of digital devices such as cell phones with cameras and
- B — The availability of inexpensive hardware with high computing power, such as the first graphics processing units (GPUs) (commonly known as video cards or graphics cards).

This technological evolution ‘allowed the construction of neural networks with tens and even hundreds of layers of neurons, known as deep neural networks’ which gave rise to the term deep learning.¹¹⁹ The aforementioned authors describe how this technology works:

¹¹⁶ Caiafa and Lew (n 109).

¹¹⁷ Russell and Norvig (n 102).

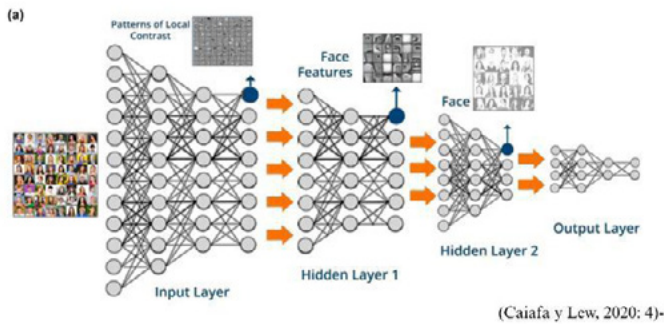
¹¹⁸ Caiafa and Lew (n 109).

¹¹⁹ Goodfellow (2014) cited in *ibid*.

In the case of natural image analysis, the architectures learn from the training data by adjusting the synaptic weights of the neurons. Thus, the first layers of neurons learn to detect local elementary patterns at the pixel level of the image, while the final layers learn to identify complex groupings of elementary patterns ... In this way, a deep neural network manages to capture the structural characteristics of the dataset with which it is trained. Unlike the detection methods of the 1980s, where experts defined which characteristics best defined the data, in deep convolutional networks these characteristics are learned automatically during training.¹²⁰

The photograph accompanying the explanation is as follows:

*Figure 2. Neural networks*¹²¹



It was this development of neural networks on GPUs that enabled other advances in the area, such as generative models for artificially synthesising new data with the essential characteristics of the training dataset. For the purposes of this work, it is worth mentioning that these types of networks are called Generative Adversarial Nets (GAN), were introduced by Ian Goodfellow and constitute the best known technology for producing deepfakes.¹²²

¹²⁰ Caiafa and Lew (n 109) 4.

¹²¹ *ibid.*

¹²² Pantserev (2020) cited in Lendvai and Gosztanyi (n 107).

2.2 Artificial intelligence today

The landscape we face today is very different from the early days of AI. This is due to the exponential increase in data that fuels AI, the computation of which began in the 2000s with the introduction of GPUs, as mentioned earlier. Added to this are more complex algorithms, more efficient processors, increased computing and storage capacity, and simultaneously, growing public interest in AI and investments in its development.

In the information society, the volume of data that surrounds us today is constantly growing, and in turn, technologies like the internet generate it at an accelerated pace, leading to the era of 'Big Data'.

In the field of computer science, this term is not new, having emerged in the mid-1990s; however, it is more prevalent today. Simply put, the term Big Data refers to the processing of large masses of data. The literature on the subject agrees in highlighting some characteristics such as: volume, velocity (created in real time) and variety of sources, natures and formats.¹²³ Likewise, authors such as Boyd and Crawford expand on the definition, considering that Big Data not only refers to large datasets and the tools and procedures used to manipulate and analyse them, but also to a shift in computational thinking and research.¹²⁴

To perform this data processing, algorithms are applied, which are at the heart of AI. The Special Rapporteur on Freedom of Opinion and Expression, in his report, reiterates the definition made by the Council of Europe, which defines algorithms as:

¹²³ R Kitchin and G McArdle, 'What makes Big Data, Big Data? Exploring the Ontological Characteristics of 26 Datasets' (2016) 3(1) *Big Data & Society* 1 <<https://doi.org/10.1177/2053951716631130>> accessed 17 January 2026.

¹²⁴ Lorenzo Cotino Hueso, 'Big Data e Inteligencia Artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales' *Ética de datos, Sociedad y Ciudadanía* (2017) 131–50.

At the heart of AI are algorithms, computer code designed and written by humans that executes instructions to translate data into conclusions, information, or products. Algorithms have long been essential to the functioning of everyday communication and infrastructure systems. The sheer volume of data in modern life and the ability to analyze it drive AI. The private sector certainly sees data this way: the more data available to feed algorithms, and the better the quality of that data, the more powerful and accurate the algorithms can become. Algorithmic systems can rapidly analyze large volumes of data, enabling AI programs to perform decision-making functions that were previously the domain of humans acting without computer tools.¹²⁵

It is worth adding that these codes ‘can involve different modes of epistemic or practical reasoning, and later, they can self-improve through the development of new heuristics, modify their internal information and even generate new algorithms’.¹²⁶

Considering what has been presented so far, it is important to highlight the idea that currently AI systems ‘can detect and recognize patterns of information in data, which is achieved by combining computers, the internet, algorithms and programming languages to solve problems or make decisions that previously could only be made by our cognitive abilities’.¹²⁷

2.2.1 Stages of artificial intelligence

The literature agrees that there are three stages of development of AI according to the capacity of the machines, that is, according to what they can do:

- *Narrow, weak or restricted AI (ANI)*: This type of AI is characterised by being trained and focused on specific tasks. It involves the execution of programmed tasks to achieve a specific result. Most scholars consider this to be the current state of affairs.

¹²⁵ UNGA, ‘Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression’ (n 104).

¹²⁶ Sánchez Caparrós (n 108).

¹²⁷ *ibid* 3.

- *Artificial general intelligence (AGI)*: For its part, this stage refers to systems that do not yet exist according to which AI can reproduce or simulate human behaviour, with the ability to manage several fields of knowledge at once (‘multitasking’).
- *Artificial superintelligence (BUT)*: The third stage corresponds to a stage in which AI surpasses human capacity.

2.2.2 Narrow artificial intelligence

As explained, humans are currently in the first stage of AI evolution, as it is designed to perform specific tasks. Thus, AI is subdivided into categories with their own approaches and techniques:

Table 4. Narrow artificial intelligence categories

Robotics	It is the most visible expression of AI, associated with the creation of robots that can perform tasks in the physical world. Examples of this include drones, industrial robots and autonomous vehicles.
Natural language processing	The area of AI where it is trained to understand human natural language. Examples include text generation and chatbots such as ChatGPT.
Computer vision	This subcategory enables computers to ‘see’ and understand their surroundings through image and video processing. This technology is used in medical image analysis.
Machine learning	In simple terms, this narrow branch of AI seeks to enable machines to learn. Thus, it is based on a set of techniques and methods that allow algorithms to receive a large amount of data about a ‘problem to be solved’ and, from a computational process, to wield correlations based on probabilities.

Now, from what has been said, it is important to continue with the subcategories that comprise machine learning to begin charting the course towards the object of this thesis: deepfakes, not without first warning that these subcategories can be white box or black box.

Table 5. Machine learning's systems types¹²⁸

Machine learning	
White box	Black box
These systems possess explainability as a property, which allows them to provide traceability or explanation of actions for the purpose of being understood by human experts.	These systems, on the other hand, do not allow for tracing the steps through which AI made certain decisions; they are as opaque to human experts as they are to those who designed the system. A prime example of this is AI systems based on complex neural networks known as deep learning.

2.2.3 Machine learning

It is important to note that the modern relevance of AI is due to the success of this type of AI. This type of learning involves inputting a set of data called ‘training data’, which is then subjected to an analysis process (‘training’) through computational processing (a programme is executed on the dataset). This process results in a solution to a problem with a certain probability of success, which is called a ‘machine learning model’. Since the responses are based on probabilities, they are often called ‘predictions’, which accounts for the possibility of a margin of error.¹²⁹

Table 6. Machine learning techniques¹³⁰

Machine learning techniques can be classified into:

Supervised learning	AI training is done with labelled (structured) data with the expected results.
Unsupervised learning	This method does not require labelled data for training. It is used when labelled data is unavailable or difficult to label.
Semi-supervised learning	The methods are combined. There is a small amount of labelled data along with a large amount of unlabelled data.
Reinforcement learning	The system learns to make decisions as it interacts with its environment: the automated agent is trained using input from the environment (whether it's a real-world scenario or a simulated one), the decision made and its effect, evaluated as success or failure. For each decision made by the agent, the model's behaviour is adjusted to improve its overall performance.

¹²⁸ Sánchez Caparrós (n 108).

¹²⁹ Gómez and others (n 103).

¹³⁰ Table created by the author using information from *ibid*.

It's worth mentioning that neural networks can be applied to any of these types of machine learning models. This system, as mentioned in previous pages, is inspired by the functioning of neurons in the human brain and represents a field of great advances in AI in recent years. The computational system consists of nodes organised in layers.

Each node performs a mathematical calculation in which it receives information from other nodes (generally the previous layer), combines it using coefficients and produces an output result (by applying a function to this result) and shares it with the following nodes. The layers are divided into three types: input layers, hidden layers and output layers. Information propagates through the neural network from the input layer, where the data is encoded, through the hidden layers, and finally to the output layer, where the final result is produced. When these networks have many hidden layers, they are called *deep neural networks*. Training these networks requires 'enormous volumes of data to estimate the value of the internal coefficients that produce the best results'.¹³¹

Thus, the highlighting refers to what is known as deep learning. It constitutes a subfield within machine learning. However, its distinguishing feature is that it is composed of multilayered neural networks trained on massive datasets, simulating the learning and decision-making processes of the human brain. It differs from 'general' machine learning in so far as the latter is usually made up of simple neural networks with one or two computational layers.

Deep neural networks consist of several interconnected layers of nodes, each building upon the previous layer to refine and optimise prediction or categorisation. This progression of calculations is called 'forward propagation'. For system training, 'back-propagation' is also used to calculate errors in predictions, allowing the neural network to make predictions and corrections, thus making the algorithm more accurate over time.

In simple terms, the training is structured as follows:

- 01 —The network makes a prediction,
- 02 —Compare that prediction or result with the actual one and calculate an ‘error’,
- 03 —This error propagates backward through the network, and the connection weights are adjusted slightly to reduce the error in future predictions, and
- 04 —The cycle is repeated thousands or millions of times until the network reaches a desired level of accuracy.

Deep learning, as mentioned, is made up of many layers, which can be visible as the input layer and the output layer, while in between are the hidden layers.

Table 7. Deep learning layers

Input layer	This layer receives the initial data. Each neuron in this layer processes a feature of the input data.
Hidden layers	In these layers, neurons perform complex calculations on the data received in the input layers. Each neuron has an ‘activation function’ that decides whether the signal is transmitted to the next layer and with what intensity.
Output layer	This layer produces the result of the prediction or categorisation.

Deep learning identifies and encodes the patterns and relationships that exist within that large mass of data in order to later understand user prompts. The success of this deep learning network will depend on the shape or architecture of the network, that is, what type of connections exist between the nodes.

Some possible configurations are:

- *Convolutional neural networks (CNNs)*: These are a popular neural network model for image and video processing, as they can identify visual features and patterns. They are commonly used for facial recognition and are also beneficial for AI-based image recognition applications.¹³²

¹³² Martin Ortiz Dominguez, ‘Artificial neural networks’ (2025) 12(23) Ingenio y Conciencia Boletín Científico de la Escuela Superior Ciudad Sahagún 38, 40.

- *Neural networks deconvolutional*: These function in the opposite way to the previous ones. Their objective is to detect missing signals or features that were discarded while the convolutional scanners performed their assigned task. They are used for image analysis and synthesis.¹³³
- *Recurrent neural networks (RNNs)*: This neural network model works by storing the output generated by its processing nodes and feeding it back into the algorithm, thus improving its predictive capacity. They have the ability to be self-taught from their errors. They are commonly used for audio or text-to-speech processing and sales and stock market predictions.¹³⁴
- *Transformative neural networks (transformers)*: This architecture transforms an input sequence into an output sequence. It was crucial for machines to understand natural language processing. They better understand contexts and meanings.
- *Feedforward neural networks*: They circulate information in only one direction. They are designed to process a large volume of ‘noisy’ input data and create ‘clean’ outputs. They are the basis of facial recognition and computer vision.¹³⁵
- *Modular neural networks*: These are a set of independent neural networks, where each independent group constitutes a module without interfering with the processing of the other modules and are supervised by an intermediary. They are used for the efficient and rapid execution of extensive processes.¹³⁶

¹³³ Ortiz Dominguez (n 132) 41.

¹³⁴ *ibid.*

¹³⁵ *ibid.*

¹³⁶ *ibid.*

- *Generative adversarial or antagonistic networks*: These networks are a generative modelling solution that leverages convolutional neural networks along with other deep learning configurations to automate pattern discovery in data.¹³⁷ They create new data that resembles the original training data. These networks are self-training, and their output is highly realistic.

These models can be used to develop one of the types of deep learning: generative AI.

2.2.4 Generative AI

This category of AI is based on models of deep learning. It is necessary to discuss it not only because it is the type of AI that is currently prominent, but also because one of the types of productions it favours are deepfakes.

The unique feature of this type of AI is that it can create original and unique content in seconds, according to the request made to it (prompt) and the existing information (massive amount of data with which this system was trained). This information usually comes from the Internet: social networks, research articles, while its processing is possible thanks to Big Data.

Generative AI can create everything from images, videos, texts and conversations to songs. This creation process is carried out in three phases: training, adjustment and generation, and evaluation and readjustment.

How do they work?

Generative AI begins with a deep learning model that serves as a foundational model, that is, the basis for other generative AI applications. Foundational models can be language learning models (LLM) or models for generating videos, images and on-demand video and music. There are also multimodal models that support the production of a variety of content.

Thus, an algorithm is trained to deep learning using unstructured data – that is, unlabelled, raw data – training involves evaluating this amount of received data by attempting to ‘predict’ the next element in a sequence, for example, the next word in a sentence or the next element in an image. It continues to refine itself

¹³⁷ Ortiz Dominguez (n 132).

until it minimises the difference between its predictions and the actual data. The result is a ‘parameter neural network’ (coded representations of the entities, patterns, and relationships of the data) that can autonomously generate content in response to inputs or instructions.¹³⁸

This process is costly, requiring not only calculations but also money, primarily thousands of GPUs and weeks of processing time. However, technologies now exist that allow us to eliminate this step and its associated costs.

The adjustment stage is the part of the process that requires significant human intervention. This is because, in order to refine the results or outputs, the model needs to be fed with a large amount of now structured or labelled data, instructions or questions (this technique is called tuning). Another technique involves human feedback to update and/or improve the accuracy of the system; this generally occurs when people rate or correct the system’s responses.¹³⁹

Finally, the last stage of generation, evaluation and readjustment consists of obtaining the result of the generative AI that is constantly evaluated in order to improve the performance of said generative model.¹⁴⁰

Just as a side note, the most relevant architectures within generative AI models are variational decoders (VAEs), generative adversarial networks (GANs) and transformers. For the purposes of this paper, we will only focus on the second, due to its role in the production of deepfakes.

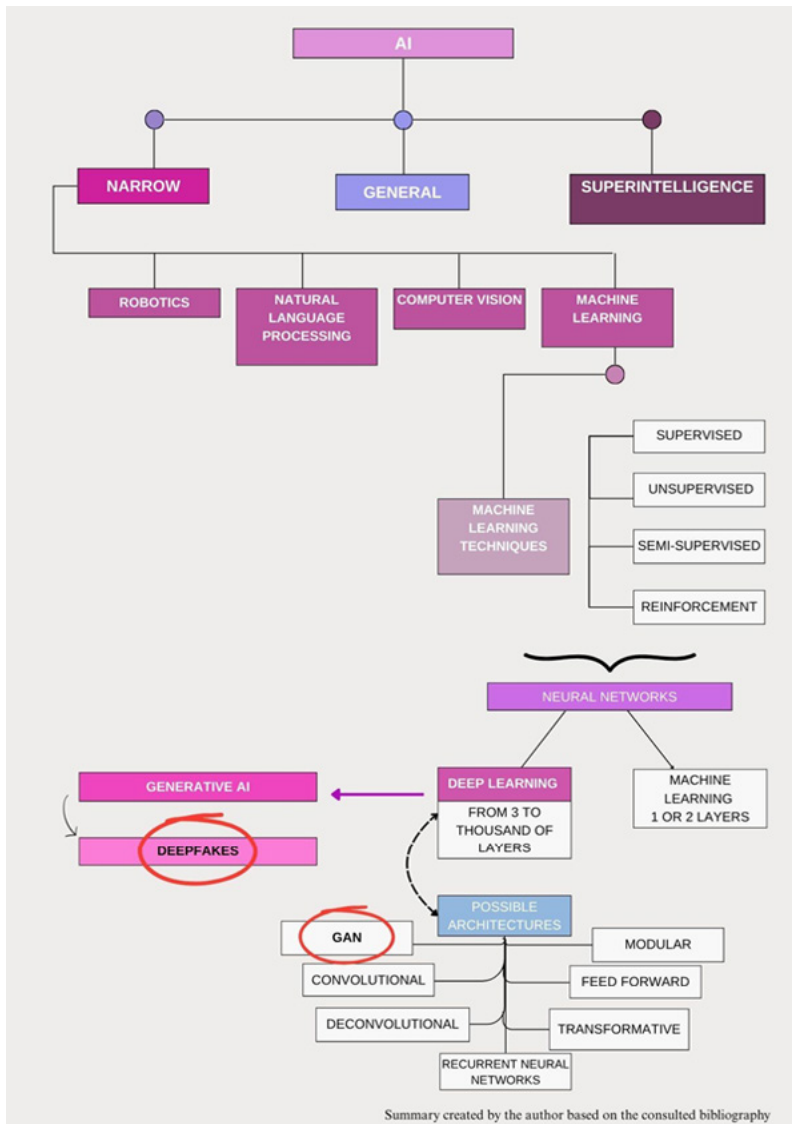
From the above, it is clear that this is a rapidly evolving technology and, as such, presents risks. In this regard, we propose to address one of the challenges in this area, especially in crisis and electoral contexts: deepfakes.

¹³⁸ Cole Stryker and Mark Scapicchio, ‘¿Qué es la IA generativa?’ (IBM, 2024) <<https://www.ibm.com/es-es/think/topics/generative-ai>> accessed 17 January 2026.

¹³⁹ *ibid.*

¹⁴⁰ *ibid.*

Figure 3. AI types summary



2.3 Deepfakes or ultra fake content

Generative AI impacts software that creates content, text, images, videos and audio. One type of production is popularly known as deepfakes.

The term deepfake is a neologism that arises from the union of deep and fake. These are understood as image, video and audio content that appears authentic but is actually created using AI algorithms.¹⁴¹ They depict individuals saying or doing things they have genuinely and never actually said or done; that is, what is shown in a deepfake never happened. Such content appears real to the human eye, which can make it challenging and difficult to identify as technically manipulated.

2.3.1 Were they always among us? Evolution of the deepfakes

A — Foundational bases (1990s-2000s)

While the term deepfake is relatively recent, its foundational bases can be found between the 1990s and 2000s. In this period, video manipulation techniques and editing software were developed, which made this technology more accessible,¹⁴² for example, Adobe Premiere is one of the types of programmes that laid the foundation for what would later become attempts at facial replacement in videos.

In the early 2000s, techniques were developed to video manipulation which allowed people to simulate saying words they hadn't actually said. This was a precursor to lip-syncing used in deepfakes. The same thing happened with the face-swapping technique.¹⁴³

Furthermore, advances were made in the field of machine learning, particularly on neural networks, as well as in facial recognition technology, the latter being a key component of modern deepfakes.¹⁴⁴

¹⁴¹ S Alanazi and S Asif, 'Explorando la tecnología deepfake: creación, consecuencias y contramedias' (2024) 6 Human-Intelligent Systems Integration 49 <<https://doi.org/10.1007/s42454-024-00054-8>>.

¹⁴² Robin Chataut and Aadesh Upadhyay, 'Introduction to Deepfake Technology and its Early Foundations' in *Deepfakes and Their Impact on Business* (IGI Global Scientific Publishing 2025).

¹⁴³ *ibid.*

¹⁴⁴ *ibid.*

B — The 2010s

This period is characterised by developments in face-swapping technology. Face2Face, a method for capturing and recreating faces in real time, is launched constituting a big step towards realistic face swapping.¹⁴⁵

While deepfake technology was initially driven in mid-2010 by the availability of large datasets and the possibility of processing them, it experienced two irreversible advances in this decade:

- In 2014, Ian Goodfellow and his team incorporated GANs or adversarial neural networks into the AI spectrum which constitute a fundamental piece in the production of high-quality ultra-fake images, video and audio.
- On the other hand, in 2017 the term deepfake gained popularity when a Reddit account of the same name began posting manipulated videos that superimposed the faces of female celebrities, such as Taylor Swift, onto pornographic videos, creating content that was convincing to the public (high-impact content). It's worth noting that these videos were created using deep learning neural networks at a time when the necessary tools were becoming more accessible and available to the general public with basic or moderate technical skills.

The following year, FakeApp, an application that allows the creation of this type of ultra-fake content, was launched and has made this technology more accessible to a wider audience. Furthermore, Chataut and Upadhyay point out that in the same year concerns about the implications of deepfakes increased when they became known on the internet deepfake Obama, portrayed by Jordan Peele.¹⁴⁶

C — From 2019–present

From 2019 onwards, as the aforementioned authors point out, legislative measures began to be adopted in response to the deepfakes. Between 2020 and 2021, technologies such as NVIDIA's StyleGAN emerged simultaneously, enabling cloning a perfect

¹⁴⁵ Chataut and Upadhyay (n 142).

¹⁴⁶ *ibid.*

video and voice. OpenAI also launched DALL-E, which represented a breakthrough in the field of generative AI. Since 2022, efforts are being made to establish regulatory policies to reduce its potential misuse and misinformation through deepfakes.¹⁴⁷

2.3.2 How are they produced? Introduction to deepfakes production

The deepfakes are produced using neural networks through deep learning algorithms, particularly GANs. The latter revolutionised the creation of deepfakes because they allowed the synthetic production of such highly convincing content.

According to Alanazi and Asif, the task is specifically performed through decoders. This procedure implies that neural networks are trained to encode and decode images or videos.

The role of the encoder is to take the initial training information (input) which can be an image or video and condenses it into a latent code, while preserving vital characteristics and filtering out unnecessary details.¹⁴⁸

Next, that latent code is transmitted to the decoder, which reconstitutes the original content based on this new code produced by AI.

As the aforementioned authors state, adversarial neural networks known as GANs represent a powerful type of neural network for creating deepfakes.

In this particular case, the typical architecture of GANs requires two basic components: a generator and a discriminator:

- The generator (G): creates new data samples such as synthetic images or videos, while
- The discriminator (D): evaluates them by comparing them with real data to distinguish authenticity. That is, it must distinguish between those created by the G network and the real ones.

These functions are carried out during the training process in which the generator strives to outsmart the discriminator, who in turn adapts by improving in identifying false data.

¹⁴⁷ Chataut and Upadhyay (n 142).

¹⁴⁸ Alanazi and Asif (n 141).

As mentioned earlier, the advancement of GANs was introduced by Ian Goodfellow and team in 2014 in their work '*Generative Adversarial Nets*'. They propose a very simple way to understand the architecture of GANs:

The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistinguishable from the genuine articles.¹⁴⁹

In other words, this popular architecture for generating images and videos consists of two competing neural networks: the generator network and the discriminator network. According to Goodfellow, this competition can be analogous to a game between counterfeiters, who produce coins trying to avoid detection (generator), and the police, who try to identify the counterfeit coins (discriminator). Each team or network is forced to improve its methods until the counterfeit coins are indistinguishable from the real ones. All of this occurs during the processing of data by the neural networks.

¹⁴⁹ Ian Goodfellow and others, 'Generative Adversarial Nets' (University of Montreal 2014).

Basically, the procedure for creating deepfakes according to Chataut and Upadhyay involves the following steps:¹⁵⁰

01 — Data collection: A database is established with all images or videos collected from various sources about the individual or person who is the subject of the deepfakes.

02 — Training: This database is used to train neural networks, which learn to recreate the facial features, expressions and movements of the target individual.

03 — Face swap or original content: The model trained with the data is used to generate new content, either by swapping faces or by synthesising completely new or original content.

04 — Refinement: Post-processing techniques are applied to improve the realism of the generated content. Issues such as lighting, skin texture and their appearance are addressed.

It is crucial to highlight that AI has currently improved the quality and efficiency in generating deepfakes. Some models can create this type of content using few images or a few seconds of audio. In this way, while increasing its accessibility and speed of production, it also increases its risks.¹⁵¹

At the same time, it is important to be clear that human intervention is involved in this production process. This is clearly evident when selecting who will be the target of the deepfake and for what purpose will it be done? In other words, there is a first stage of decision.

On the other hand, this interference is observed at the initial stage, that is, to collect data and establish the dataset of the person one wants to imitate to train the AI. In other words, someone must search for, select and organise this data. Human involvement is also involved in configuring the learning model to achieve better results.

There is also human intervention at the time of obtaining the output or result, to control its 'quality' during refinement, in which edits are applied to it (for example, before the desynchronisation (lip and audio) with a view to making it more convincing or realistic.

¹⁵⁰ Chataut and Upadhyay (n 142) 1.

¹⁵¹ Mitra and others cited in *ibid*.

Based on the above, the following is proposed: a production scheme of deepfakes, following the guidelines of the Report of the Special Rapporteur on Freedom of Expression which emphasises human action as an integral part of AI:

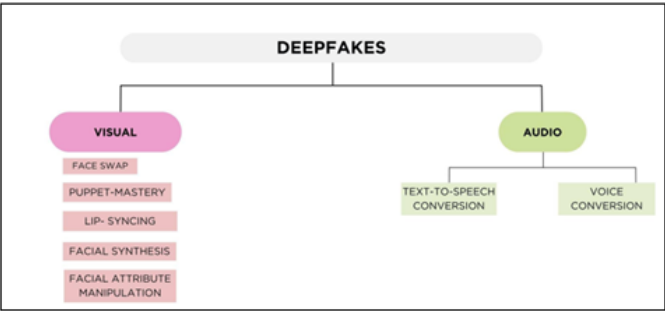
Figure 4. Deepfakes production scheme¹⁵²



2.4 Types of deepfakes

According to the 2024 Artificial Intelligence Report, different types of deepfakes can be created: For the purposes of its presentation, the following figure is presented:

Figure 5. Types of deepfakes¹⁵³



¹⁵² Based on a scheme presented in UNGA, 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression' (n 104) 5.

¹⁵³ Table presented in Reuel Anka, 'Chapter 3: Responsible AI' in Maslej N and others, *The AI Index 2024 Annual Report* (AI Steering Committee, Institute for Human-Centered AI, Stanford University 2024) 48.

So, the potential forms that the deepfakes correspond to:

Table 8. Description of deepfakes types

Face swap	The facial image of the original subject of the image or video is replaced. They are the most accessible and basic form of deepfake.
Puppet-mastery	This type of deepfake involves creating a completely new video, in which the movements of an actor or individual are recorded and that image is superimposed onto another subject, as if it were a puppet, hence its name. Image and voice samples of the target person is sufficient for the creation of this type of content.
Lip-syncing	An existing or real video is altered so that the objective subject of the deepfake seems to be talking about new content.
Entire face synthesis	This involves creating a highly realistic image or video of a face, real or imagined. This is where GAN networks, explained on previous pages, become crucial. Data provided as input (eg, photographs of real people) passes through the layers of these networks.
Facial attribute manipulation	In this case, facial expressions are altered in a real image or video. It also uses GAN networks.
Text-to-speech conversion	Text synthesis involves transforming it into natural speech in real time.
Voice conversion	It allows you to transfer a person's speech style, tone and accent to another person in a video or audio recording. You can even manipulate real audio or create a fake recording using a real person's voice, making them say something they never said.

2.5 Deepfakes: Political arena and disinformation

Deepfakes have a wide range of applications across various fields. For example, in the film industry, they have allowed for the de-aging of actors or the recreation of deceased artists. In the realm of video games, deepfakes allow for a more immersive experience for users by enabling the creation of more realistic facial animations. In the educational field, they allow historical figures to be brought to life. Also, in medical training, they allow for the simulation of interactions with patients. However, just as there are valuable uses of deepfakes, they also pose ethical challenges and there are indeed worrying points about their (mis)use.

In this sense, they highlight its use in the political arena due to its ability to influence public opinion and election results: ‘One of the most concerning areas of deepfake misuse is in the political arena. The ability to create convincing fake videos of political figures has the potential to significantly influence public opinion and election outcomes’.¹⁵⁴

Similarly, Stanford University’s 2024 AI Index Report identifies the use of AI in political processes as a tool for disinformation as a major concern, since the development of the technology discussed in these pages makes it easier to generate disinformation. Furthermore, as the report points out, since the 2020 US elections, AI tools have become increasingly prevalent and deepfakes have improved significantly. This can undermine trust in democratic institutions, manipulate public opinion and polarise public discourse.¹⁵⁵

2.6 Conclusion of Chapter 2

In this chapter, a necessary journey is traced on the conceptualisation of AI in general until reaching the object of study of this thesis: deepfakes.

These pages address one of the main topics on the global agenda, namely the rise of AI, prioritising one of its most popular products today.

This document outlines the historical evolution of the technology behind ultra-fake content, while also highlighting a series of characteristics and details of its operation. This allows us to begin addressing part of the second specific objective of the research.

It is worth considering as relevant aspects of this chapter that, on the one hand, show that it is evident that the technology we are discussing has made great leaps in terms of development in a drastically reduced period of time. At the same time, this chapter has also highlighted the potential, in negative terms, that it could have as a tool for disinformation, especially in the political arena (as indicated in point V).

¹⁵⁴ Chataut and Upadhyay (n 142).

¹⁵⁵ Anka (n 153).

Regarding the body of literature consulted, it should be noted that it consists mostly of non-Latin American works. American and European research predominates, and Indian literature has also been used as a reference, all of which are available in English. Furthermore, the cases of deepfakes have taken place in those regions.

This aspect, far from being trivial, motivated a question to be posed to Google's Gemini AI: Why are there no Latin American authors that speak about deepfakes? The answer provided highlights that there are no Latin American authors who are 'pioneers' or 'key figures' in the fundamental development of deepfakes for 'interconnected reasons that reflect the global landscape of research and development in Artificial Intelligence'.

These reasons stem from: 1) the gap in investment and resources, as they tend to come from large technology corporations such as Google, Meta and OpenAI; 2) in the United States, Europe and parts of Asia, research institutions and centres have a long history of funding that fosters the study of AI; 3) the connectivity of Latin American authors with the global scientific community on AI is proportionally lower, hindering the visibility of their work; 4) research into deepfakes is a very specific and advanced branch of deep learning, ergo, the countries that lead its development in this field are the same ones that have invested massively in fundamental AI research for decades.

At the same time, Gemini believes that Latin America faces a context of significant limitations in public and private investment in research, especially AI, and less sustained long-term funding, as well as lacking the computing infrastructure for the massive data processing required by neural networks.

In this regard, it is worth considering that although this response does not follow a scientific method, but rather the filtering of large amounts of data, and may contain biases or partialities, it is nonetheless a relevant aspect that invites reflection and from which the significance of this research arises.

On the other hand, in order to fulfil the third specific objective of the thesis, you are invited to continue reading Chapter 3 of this work in order to continue the journey but now with a greater focus on the political arena of the American continent.

3. American deepfakecracy



It's an attempt at digital fraud ... it's sad because what we're trying to rebuild is trust among Argentinians, and the world's trust in us. And if we attack the quality of democracy, the rules of democracy, we'll be undermining it. We put at risk all the effort that Argentinians are making to get out of the economic crisis ... We are in a country that wants to have more transparent rules of the game, which is what has not happened today, it has been a very violent campaign.

Statements by the former President of Argentina and leader of PRO, Mauricio Macri¹⁵⁶

What happens if, the day before the election, the leader of the political party you were planning to support announces that he won't be participating? You see him in a video; he looks like him, he talks like him, but he isn't.

This chapter begins with a clear and concrete example of deepfakecracia in the Americas. It is presented as a new way of conducting state elections, in which the lines between reality and falsehood are blurred.

Deepfakecracy accompanies the emergence and constant refinement of generative AI, and in its wake challenges the traditional rules of the game in politics and public debate.

¹⁵⁶ [A24.com](https://www.a24.com), 'Votó Mauricio Macri: Habló del temporal y sobre la difusión de videos falsos' (YouTube, 18 May 2025) <https://youtu.be/Al06Kd5ud4Q?si=2ilj7F_INKTLqnEB> accessed 18 January 2026.

The deepfake to which the reader is referred circulated during the electoral silence period of the legislative elections in the city of Buenos Aires, Argentina. The AI-manipulated video was published on X (formerly Twitter) in a coordinated effort by anonymous accounts and others that were not so anonymous. The fact that the Buenos Aires electoral district was under an electoral reflection period did not constitute an obstacle, nor did the order issued by the Electoral Court to remove the blatantly fake content.

Thus, Chapter 3 begins with just one example of what happens during election periods following the emergence of generative AI in a world that is beginning to be characterised by the technopolarity. In this sense, the chapter intends to address, in an experimental way, cases of political deepfakes during the last presidential elections that took place from 9 February 2023 to 9 February 2025.

This last point tends to fulfil the third objective proposed by this thesis, in which an attempt will be made to understand the articulation of the deepfakecracia, understood as a possible government or power of the ultra-false in electoral contexts, in the American continent, while at the same time attempting to highlight aspects of its way of operating in said continent, if any.

3.1 State of the art: Deepfakes in the 2023 election campaigns

The application of generative AI has been the focus of concerns due to the understanding that it undermines democratic processes in elections. In this sense, the main media tool is deepfakes.¹⁵⁷

In this sense, the authors argue that they contribute to reinforcing the feeling of uncertainty among citizens and negatively affect the information environment.

Labuz and Nehring, in their study, analyse the use of deepfakes in 11 countries during 2023 in the context of elections.¹⁵⁸

¹⁵⁷ Mateusz Labuz and Christopher Nehring, 'On the Way to Deep Fake Democracy? Deepfakes in Election Campaigns in 2023' (2024) 23 *European Political Science* 454.

¹⁵⁸ *ibid.*

The importance of this empirical precedent is that it records the use of electoral deepfake processes in European, Middle Eastern and some Asian countries and highlights the lack of analysis of this issue in Latin America. Furthermore, the analysis uses 2023 as its timeframe, while the study proposed in this work focuses on two calendar years (9 February 2023 to 9 February 2025), which allows for the inclusion of two presidential election periods in the same country, as in the case of Ecuador, and enables an evaluation of the greater or lesser proliferation of deepfakes between the periods.

Thus, the aforementioned authors verify that deepfakes are used to undermine confidence in elections, their results and potential candidates. In Poland, the United Kingdom, India, Taiwan and Slovakia, political leaders were targeted by deepfakes, deepvoices were circulated on social media and in some cases political opponents were registered as the perpetrators.

In other cases, such as Bulgaria or Indonesia, deepfakes were recorded for video that alluded to opaque or corrupt relationships to discredit the image of leaders or candidates.

Some of the most prominent examples in this analysis are Turkey and Slovakia: In Turkey, a pornographic deepfake of a candidate went viral, leading him to withdraw his candidacy. Although his vote share was only 5%, this withdrawal represents an unforeseen redistribution of votes, impacting the final election results. In Slovakia, two days before the election and during the campaign blackout period, an AI-generated audio recording was released in which the liberal party candidate, Michael Simecka, and journalist Monika Todová discussed a plan to rig the election by buying votes from the country's marginalised Roma minority.¹⁵⁹

The peculiarity of this case lies in the viral spread of a single audio file, and because it was released during the electoral ban period, it was not quickly denied.

The researchers understand that while pre-election polls showed a close race between Simecka and the SMER party, where SMER appeared to be winning by a narrow margin, deepfakes have the potential to make a significant difference in these kinds of polarised elections. In this case, it is understood that the audio's objective was to demobilise voters for Simecka and to encourage the far right. The SMER ultimately won by 5%.

¹⁵⁹ Meaker (2023) cited in *ibid*.

This case was also featured in Stanford University's AI Index Report as an example illustrating how AI-based disinformation can be used in a political context.¹⁶⁰

It should be added that the study carried out by Labuz and Nehring deals with only two cases in the Americas: Argentina and the United States, whose two examples will be taken up again in conjunction with other cases identified from the fieldwork presented in this research.¹⁶¹

Furthermore, to this sum of countries we must add another empirical work such as that prepared by Barrientos-Báez, Piñero Otero and Porto Renós in which cases of deepfakes in Spain, especially aimed at women in the political context.¹⁶²

To this list should be added cases compiled in Antigua and Barbuda, highlighted by the Special Rapporteur for Freedom of Expression (OAS) in its recent 2024 Annual Report. The report indicates that there were reports of deepfakes in which the figure of leaders like Donald Trump is used to speak disparagingly of Antigua and Barbuda. One of the news items on which this point of the report is based also mentions the same operation for Jamaica.¹⁶³

3.2 American deepfakecracy

The aforementioned empirical studies account for the political use of the deepfakes in electoral contexts, in countries in Europe and Asia. Therefore, it is worth asking what happens with its implementation in the Americas. Is there an American deepfakecracy? Does it share techniques or patterns with Europe and Asia? In this sense, prior to explaining how it is articulated the American deepfakecracy in America, it is possible to clarify the methodological guidelines that guide the collection of case studies.

¹⁶⁰ Reuel Anka, 'Chapter 3: Responsible AI' in Maslej N and others, *The AI Index 2024 Annual Report* (AI Steering Committee, Institute for Human-Centered AI, Stanford University 2024) 49.

¹⁶¹ Labuz and Nehring (n 157).

¹⁶² Almudena Barrientos-Báez, María Teresa Piñero Otero and Denis Porto Renó, 'Imágenes falsas, efectos reales. Deepfakes como manifestaciones de la violencia política de género' (2024) 82 *Revista Latina de Comunicación Social* 1.

¹⁶³ Inter-American Commission on Human Rights Office of the Special Rapporteur for Freedom of Expression, 'Annual Report of the Office of the Special Rapporteur for Freedom of Expression' (OEA/Ser L/V/II, Doc 39 2024) 32 <<https://www.oas.org/en/iachr/expression/reports/IA2024RELEEN.pdf>> accessed 18 January 2026.

3.2.1 Methodology applied to comparative analysis

As Donatella della Porta points out, comparative analysis, such as that which will be carried out in this work, has a central place in social science research. In this sense, such comparative analysis can be based on the study of variables or case-oriented research.¹⁶⁴ Both approaches are legitimate, and each allows for a data analysis as rich as the other. For this reason, a case-oriented methodology will be applied to the study of Argentina and the United States, as it allows for a more detailed examination of two of the countries with the most examples of deepfakes put into circulation during presidential elections in order to understand this phenomenon of disinformation through deepfakes in those contexts.

On the other hand, in the remaining countries, a single case that has gained media attention, as explained above, will be selected. The analysis in these cases will be based on the study of variables with the aim of establishing general relationships between them.

This establishes the following variables:

Table 9. Variables of analysis

Country	Panama
Deepfake content	
Date	
Account/or media outlet that exposed it	
Interactions	
Figure of the deepfake	
Context in which it was introduced	
Effect/s	
Typology	
Link to the campaign	

¹⁶⁴ Donatella della Porta, 'Comparative Analysis: Case-Oriented Versus Variable-Oriented Research' in *Approaches and Methodologies in the Social Sciences: A Pluralist Perspective* (CUP 2008) ch 11, 198–222.

The table outlines the variables that will be considered when analysing cases by country. These include:

- Deepfake (content): Accordingly, the deepfake is described taking into account the narrative used and what platform or medium was used.
- Date: This is relevant for the purpose of considering whether it was put into circulation during the pre-campaign period, campaign, electoral ban or after it.
- Profile/account or media outlet that exposed it.
- Interactions: This indicates the potential reach of the deepfake if it is displayed on social media, such as views, likes or shares.
- Deepfake figure: Refers to the person whose identity is used (in this case, presidential candidates) to spread disinformation in an electoral context.
- Context in which it was introduced: This variable allows us to analyse at what point in the electoral campaign the deepfake was put into circulation, whether it occurred as a result of a presidential debate or arose after the celebration of the first or second round of elections (ballotage).
- Effects: Refers to the immediate consequences following the disclosure of the deepfake in question: whether it is denied or not by the candidate or party of the candidate whose identity is used, whether it generates the withdrawal of the candidacy of the person affected by the deepfake in the elections, or whether it puts AI and the use of deepfakes on the agenda in electoral contexts.
- Typology: This indicates the type of deepfake: audio, video or image. This helps to determine if there is a pattern of selection in the Americas or a trend toward a specific type.

- **Connection to the campaign:** This analysis considers the date of dissemination and the content. Specifically, whether it was displayed during the electoral blackout period or a few days before the election, and whether the content of what is said or done in the highly fabricated video, image or audio constitutes a sensitive issue for the society in question during the campaign period or presents a proposal from the electoral platform of the party or parties.
- **Source:** Whenever possible, the analysis focuses on the primary source, that is, where the deepfake originates. However, it also considers when it gains media attention in online newspapers or reports from election observation missions or fact-checking organisations.

Limitations

One of the potential limitations of this thesis is that the search focuses on cases that have gained prominence in digital media and social networks. This could mean that local cases, ‘shared’ or ‘viralised’ within limited circles, might be overlooked or go unnoticed – that is, they might not receive the same visibility. Along these lines, it should be mentioned that cases of deepfakes. The examples analysed here are constantly being updated, so there may be recent examples that are excluded from these pages.

In turn, it is possible that cases that have gained media relevance may no longer be active in the digital ecosystem, so it is not possible to identify the primary source from which they originate in order to subject them to the analysis proposed here.

Furthermore, when analysing the narratives introduced, it is important to acknowledge their local element, that is, their adaptation to the idiosyncrasies and context of the country in which they are introduced. While the aim is to explore this context, the limitations that may exist in this regard cannot be ignored.

On the other hand, there is no methodological framework that establishes an adequate analysis of political deepfakes or that establishes criteria to measure the limits of influence of deepfakes or their impact on electoral results.

Finally, it should be noted that although the work will mention the role of deepfakes in the candidacies of women politicians due to its significance in the link between AI and human rights, it will be the only point that refers to previous studies on women candidates for other positions of power, since women candidates for the presidency are scarce during the 2023–2024 period.

Furthermore, while the campaign periods established by the electoral calendars published by the electoral authorities of each country will be considered, it cannot be ignored that these periods may begin much earlier than established by said authorities.

3.2.2 American electoral cycle

Table 10. American electoral cycle

2023 (From 9 February onwards)	Ecuador (extraordinary presidential elections to complete Guillermo Lasso's term)
	Argentina
	Paraguay
	Guatemala
2024	El Salvador
	Dominican Republic
	Panama
	Mexico
	Venezuela
	Uruguay
	USA
2025 (Until 9 February 2025)	Ecuador

3.2.1.1 Case-based study

Argentina

The Argentine election campaign in 2023 included three elections: the primary, open and simultaneous and mandatory (PASO) elections, which took place on 13 August 2023, to define the internal party elections while also setting the voting intentions for the general elections.

Subsequently, the general presidential elections were held on 22 October 2023 (first round). The candidates were: Sergio Massa for the Frente de Todos coalition, Javier Milei for the libertarian party La Libertad Avanza, Patricia Bullrich for Juntos por el

Cambio, Juan Schiaretti representing the Hacemos por Nuestro País coalition (Partido Justicialista Nacional) and Miriam Bregman for the Frente de Izquierda y Trabajadores (Workers' Left Front). In this first round, none of the presidential candidates obtained the required majority of votes, necessitating a second round to elect the president and vice president between the two tickets that received the most votes in the first round.

Thus, the candidates Javier Milei and Sergio Massa competed in the runoff election on 19 November, with the former being elected President of the Argentine Republic.

The campaign period was characterised by a marked and widespread use of deepfake technology to discredit political opponents and in political advertising, so much so that *The New York Times* described the elections as 'The First Elections with AI'.¹⁶⁵

Based on the fieldwork conducted and the data obtained, the use of parodic deepfakes and memes stands out, as well as their role in political propaganda, alluding to figures from the arts and film. All of this is aimed at gaining sympathy and popularity among users of the social networks where they are disseminated.

However, these types of tools have also been used to implement various strategies such as discrediting the political opponent, dismissing criticism, modifying narratives and dividing communities and groups to increase divisions.

In this regard, it is worth noting that in most of the identified cases their parodic nature or rudimentary AI manipulation made them easily recognisable. However, some cases are mentioned below that stand out for blurring the lines between reality and artifice.

In the lead-up to the 2023 primary elections (PASO), a video of presidential candidate Patricia Bullrich went viral, which was later manipulated. Specifically, two videos circulated online: the original and the fake one. The original video was posted by Ivo Foguet Poser on X (formerly Twitter) and showed the candidate with young PRO party activists chanting.

¹⁶⁵ Jack Nicas and Lucia Cholakian Herrera, 'Is Argentina the First AI Election?' (*The New York Times*, 15 November 2023) <<https://www.nytimes.com/2023/11/15/world/america/argentina-election-ai-milei-massa.html>> accessed 16 January 2026.

Meanwhile, the fake video began circulating on the same platform and contained AI-altered audio that showed the people from the initial video making discriminatory remarks: ‘we are the pituquitos from Recoleta, we are Patricia’s and Larreta’s people, the provinces are very different, we are upper class and beautiful people’.¹⁶⁶ This post was shared multiple times. The plausibility of this deepfake was based on the fact that the lyrics referenced an expression used by Martín Llaraya (governor-elect of Córdoba) criticising the interference of the ‘Recoleta rich kids’,¹⁶⁷ that is, the political class of the Autonomous City of Buenos Aires, in the interior of the country, particularly his province.¹⁶⁸

Ana Silmovich examined this case and analysed the responses of political internet users to both the original video and the deepfake. She argued that the reactions to both the original and the fake video were made by ‘militant, fanatical internet users’ and ‘logical and passionate Kirchnerist followers’.¹⁶⁹ This latter point stems from the fact that the source video mentioned the desire for a country without Cristina Fernández de Kirchner.

It’s reasonable to consider that even if one paid close attention to the deepfake to determine its authenticity, it was identifiable as a deepfake because the manipulated audio didn’t match the facial or lip movements of the people appearing in it. However, it’s important to remember that not all users take the time to verify what they see or hear given the vast amount of information available online.

Finally, only one example of a deepfake post from an X account is shown, which received an average of 1,000 likes, 661 shares and 631,600 views.

¹⁶⁶ Ana Silmovich, ‘Desinformación política e Inteligencia artificial en la campaña presidencial argentina de 2023’ in Oskar Almazá-Lopez, Begoña Cabanes-Cacho and Eva Blumbury (coords), *IA, educación y medios de comunicación: modelo TRIC* (2024) 275.

¹⁶⁷ *ibid* 275.

¹⁶⁸ La Nación, ‘Martín Llaraya clarified why he used the controversial phrase “Recoleta posh kids”’ (*La Nación*, 28 July 2023) <<https://www.lanacion.com.ar/politica/martin-llaraya-aclaro-por-que-dijo-la-polemica-frase-pituquitos-de-recoleta-nid28072023>> accessed 16 January 2026.

¹⁶⁹ Silmovich (n 166).

Figure 6. Patricia Bullrich AI-altered audio¹⁷⁰



For its part, the Feminist Economics Association (ECOFEMINITA), in its report on ‘Gender Disinformation during the elections in Argentina, October-December 2023’, compiled cases of women who were targeted by smear campaigns for expressing support for feminism during the electoral campaign.¹⁷¹ Specifically, as stated in the report, these campaigns primarily originated from the Libertarian Party, given its agenda’s opposition to that of

¹⁷⁰ Screenshot of a viral tweet spreading the analysed deepfake. @diegowainstein (X (formerly Twitter), 27 July 2023, accessed 29 June 2025.

¹⁷¹ ECOFEMINITA, ‘Gender Disinformation during the elections in Argentina, October-December 2023’ <<https://ecofeminita.com/wp-content/uploads/2024/03/InformeECOFEMIBOT.pdf>> accessed 17 June 2026.

feminist movements. It is within this context that ECOFEMINITA presents a case that fits the general criteria for analysis: a case involving a presidential candidate during the campaign within the timeframe established in this thesis. Specifically, it explains that after the presidential debates held on 1 and 8 October 2023, Myriam Bregman, after speaking out against Javier Milei, faced significant repercussions on social media, where AI was used to create distorted images where gender stereotypes are reproduced and she appears sexualised and objectified.

Based on the research and fieldwork undertaken, this case was chosen to verify its continued existence in the digital space. Specifically, the case of the deepfake circulating on X (Twitter) was considered, where a screenshot of it is displayed and the user can be seen. In the report of ECOFEMINITA the visible user is @OOC_Milei and the deepfake is accompanied by the tweet 'I always trusted you, blonde'.

In this regard, this research contributes that as of the date of writing this thesis, the deepfake image remains on the aforementioned platform, accompanied by the same inscription, 'I always trusted you, blonde', but now under the username @Kuke_cj. In other words, this case not only demonstrates the particular treatment given to deepfakes in light of gender stereotypes, but also the way accounts associated with a user operate. From the time the deepfake circulated (23 October 2023) to the present, the alleged user has not remained unchanged.

Figure 7. Myriam Bregman deepfake¹⁷²



Among the other discrediting materials that have circulated, a deepfake was detected in which Javier Milei explained how the organ trafficking business could work.¹⁷³ This case was highlighted by *The New York Times* as a deepfake that crossed the line into disinformation. Its veracity is established by the consistency between the theme or narrative addressed by the deepfake and previous statements by Javier Milei in which he suggested that organ donation could think of itself as a market and that philosophical fits with its libertarian views.¹⁷⁴

The deepfake video is featured in the online article of the newspaper *PERFIL*. The highly fake video includes a watermark with the inscription '@iaxlapatria'. This information comes from the newspaper *La Nación*:¹⁷⁵ 'The account @iaxlapatria was published by the campaign as the one used by Sergio Massa's

¹⁷² Deepfake is active in X as of 27 June 2025. Source *ibid*.

¹⁷³ Labuz and Nehring (n 157).

¹⁷⁴ *PERFIL*, 'From Soviet to Indiana Jones: The use of Artificial Intelligence in Massa's campaign according to The New York Times' (*PERFIL*, 16 November 2023) <<https://www.perfil.com/noticias/politica/uso-inteligencia-artificial-campana-sergio-massa-segun-the-new-york-times.phtml>> accessed 29 June 2025.

¹⁷⁵ Jack Nicas and Cholakian Herrera, 'Argentina, a testing ground for the use of Artificial Intelligence in political campaigns' (*La Nación*, 15 November 2023) <<https://www.lanacion.com.ar/politica/la-argentina-un-campo-de-pruebas-para-el-uso-de-la-inteligencia-artificial-en-las-campanas-politicas-nid15112023/>> accessed: 29 June 2025.

campaign on his Instagram account for AI called IAxlaPatria. However, after a verification process, said deepfake is not visibly available to the public on the Instagram, X or TikTok accounts registered under that user’.

It is worth noting that for this reason, the exact date of publication and the level of virality it may or may not have achieved cannot be identified; however, the news reports date from November 2023 and agree that it was circulated during the election campaign.

*Figure 8. Javier Milei deepfake*¹⁷⁶



On the other hand, in the same campaign, a particular deepfake was addressed by bibliographic studies¹⁷⁷ as well as journalistic entities and companies of fact-checking such as Chequeado (Argentina) in conjunction with AFP Factual (France) and Verificador (Peru).¹⁷⁸ The deepfake video in question showed Sergio Massa, then Minister of Economy and presidential candidate, consuming an apparent narcotic substance.

¹⁷⁶ Deepfake published in the digital article of PERFIL (n 174).

¹⁷⁷ Silmovich (n 166) 277; Labuz and Nehring (n 157).

¹⁷⁸ Ezzio Ramos, 'Video does not show Sergio Massa inhaling cocaine: it is a "deepfake"' (Verificador, 10 November 2023) <<https://larepublica.pe/verificador/2023/11/10/vid-eo-no-expone-a-sergio-massa-inhalando-cocaina-es-un-deepfake-474240>> accessed 29 June 2025.

The deepfake circulated in two formats: a short version in which he can be heard saying ‘this meme is great’, published by the account @sindicatodedom4 (currently suspended), and a longer version published on 8 November 2023, by the account @LiberalDeMilei accompanied by the hashtag ‘#SergioTomásDeLaBuena’. It should be noted that the validity of the deepfake and the aforementioned account could not be verified on platform X.

This means that the deepfake was published by anonymous accounts days before the run off election and under a narrative that attempted to link Sergio Massa to drugs.

It’s also worth noting that deepfakes have been disseminated on platforms like X, Instagram and TikTok amidst a plethora of decontextualised or digitally manipulated videos that share two common characteristics: the same narrative and dissemination through anonymous accounts, often using pseudonyms or aliases. According to a fact-check led by Chequeado, across all the platforms where it circulated, it accumulated a total of 3 million views.

In this regard, another controversial video went viral on 7 November 2023, under the same hashtag. Its origin cannot be determined, but it was posted by the account @GordoDAN¹⁷⁹ where it garnered 2.7 million views and 5,000 shares, while also being replicated by other accounts, both anonymous and not, on the same date, alleging that Massa was in an altered state. Regarding its authenticity, contradictory opinions circulated; some users claimed it was an AI-generated product, while others defended its veracity, but neither position has been supported by evidence to verify its authenticity.

The latter is mentioned as an example where the epistemological feature of the videos that are viewed and circulated on social networks is put in crisis, which, favoured by political polarisation, transforms the truth into a disputed terrain, where ultimately only doubt is sown and no doubt is cleared about the veracity of what the eyes of the voters perceive.

¹⁷⁹ See the tweet from @GordoDa_ (X (formerly Twitter), 7 November 2023) <https://x.com/GordoDan_/status/1722010056160403556> accessed 29 June 2025.

Figure 9. Two versions of Sergio Massa deepfake¹⁸⁰



Figure 10. Sergio Massa deepfake from an X account¹⁸¹



¹⁸⁰ Screenshots corresponding to the two viral versions of the deepfake on 8 November 2023. Chequeado, 'Home page' <<https://chequeado.com/>> accessed 29 June 2025.

¹⁸¹ Screenshot of the deepfake posted by @LIBERALDEMILEI. Source: Fact Checker.

The fieldwork reveals a large number of deepfakes circulating on social media platforms like X, Instagram and TikTok, as well as in traditional media. All of these are used for electoral purposes, with the particularity that most of them responded to a parodic or satirical nature regarding presidential candidates, which allows us to recognise their AI origin. In this sense, some online newspapers have called Argentina the 'AI laboratory in elections'. However, among this widespread use of deepfakes, there are cases where the electoral goal is to discredit the political opponent, either personally or through their campaign proposals, as has been discussed. Furthermore, it should be noted that, also in pursuit of political discrediting and undermining, deepfakes even act as a means of reproducing gender stereotypes, a topic that will be addressed later.

United States of America

The presidential elections in the United States were held on 5 November 2024. The main candidates vying for the presidency were Donald Trump for the Republican Party and Kamala Harris for the Democratic Party, after Joe Biden withdrew from the race seeking re-election.

Donald Trump emerged as the winner of the election with 49.80% of the votes, becoming President of the United States.

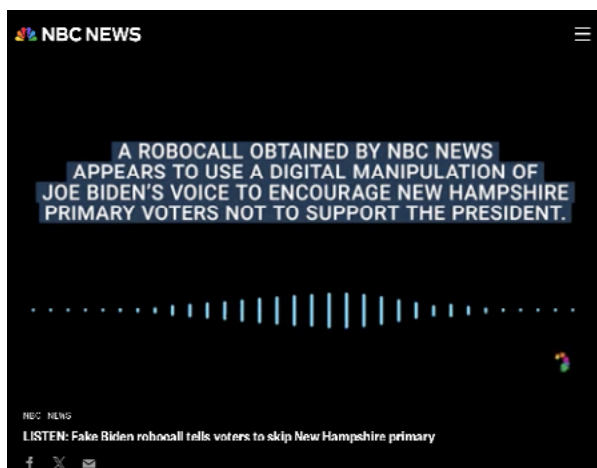
This time, as in Argentina, the election campaign was marked by the massive use of deepfake technology. The aim of these productions was to harm opposing parties during the election campaign, as well as satirise, mock and delegitimise them.

After Joe Biden announced his intention to run for president, he was the target of deepfake videos and satirical videos intended to cast doubt on his ability to hold the presidency due to his age and/or health. A relevant incident in this regard occurred during the presidential primary in New Hampshire, in which a large number of people received a 'robocalling', that is, calls made automatically by a machine, which played an extremely fake message with a deep voice, that is, manipulated by AI that reproduced Joe Biden's voice.

There he was trying to persuade voters not to go to the polls in the primary elections; journalistic sources even point out that the audio used a phrase frequently used by Biden: ‘What a bunch of malarkey!’. This last part can be analysed as a way to lend credibility to the recording, deceiving those who heard the deep voice into believing it was actually Biden. The audio lasts approximately 30 seconds.

Regarding this incident, and as reported by the United States public broadcasting service (NPR), a Democratic political consultant was held responsible and had to face charges.¹⁸² He also proposed a fine for the communications company Lingo Telecom for the automated transmission of deep voice. Another consequence to consider is that the Federal Communications Commission ruled a month later that robocalls using AI-generated voices are illegal because they are used to deceive voters and defraud people.¹⁸³

Figure 11. Biden's deepvoice¹⁸⁴



¹⁸² Shannon Bond, ‘A political consultant faces charges and fines for Biden deepfake robocalls’ (NPR, 23 May 2024) <<https://www.npr.org/2024/05/23/nx-s1-4977582/fcc-ai-deepfake-robocall-biden-new-hampshire-political-operative>> accessed 29 June 2025.

¹⁸³ Shannon Bond, ‘The FCC says AI voices in robocalls are illegal’ (NPR, 8 February 2024) <<https://www.npr.org/2024/02/08/1230052884/the-fcc-says-ai-voices-in-robocalls-are-illegal>> accessed 29 June 2025.

¹⁸⁴ NBC News, ‘Listen: Fake Biden Robocall Tells Voters To Skip New Hampshire Primary’ (NBC News, 22 January 2024) <<https://www.nbcnews.com/video/listen-fake-biden-robocall-tells-new-hampshire-not-to-vote-in-primary-202609733664>> accessed 29 June 2025.

Another relevant case, due to its particular characteristics, has been the deepfake of Kamala Harris, in which the audio of a video used to promote her campaign was manipulated. This AI-powered manipulation was circulated on the X platform by the social network's owner himself: Elon Musk, who had publicly demonstrated his financial and symbolic support for Donald Trump's campaign.

The deepfake was created and first uploaded by the anonymous account @MrReaganUSA, which claimed it was a 'mockery' of the 'We Choose Freedom' clip released by Harris's campaign team. However, when Elon Musk uploaded the deepfake directly (ie, without attribution or sharing), he failed to clarify that it was an AI-generated image, a requirement stipulated by the platform's own rules.

In other words, in this case, the platform owner himself published manipulated content intended to deceive and/or confuse people without any disclaimer informing users that it is a deepfake on his account with 221.4 million followers, even violating the rules established by his own social network.

The deepfake of Kamala Harris released on 26 July claimed that President Joe Biden appeared senile during the debate. It also stated that she was chosen because she was the right person to represent diversity as a woman of colour, and that anyone who criticised her would be sexist and racist. Furthermore, the manipulated voice of the candidate asserted that she knew nothing about running a country.

The post on Elon Musk's profile received 137.1 million views, 55,000 replies and 234,000 likes. At the same time, the viral spread of this alleged incident had media repercussions and brought the regulation of deepfakes to the forefront of the public agenda, as political figures like California Governor Gavin Newsom used this case to speak out against the possibility of manipulating voices in political advertisements, deeming it illegal. All of this discussion took place on the X platform.

This case demonstrates how platform owners become key political actors, not only in presidential campaigns but also in discrediting and spreading political disinformation against opposition candidates. Specifically, it is the owner of X himself who, using these resources, leverages deepfakes as weapons of disinformation, with the potential to erode public trust among voters.

Figure 12. Kamala Harris deepfake published by Elon Musk's X account¹⁸⁵



Kamala Harris was also the subject of deepfakes through which she was sexualised and objectified. For example, on 7 August, she circulated a deepfake on the social network Facebook, in which she was presented dressed in tight clothing and an attempt was made to portray her as a sex worker, all in pursuit of delegitimising her political career. The publication was accompanied by the phrase 'Just so everyone knows how she got her start... DNC must be proud'. The post was shared 24,000 times.

¹⁸⁵ Screenshot of the tweet from the account @elonmusk in which he spreads the deepfake of Kamala Harris.

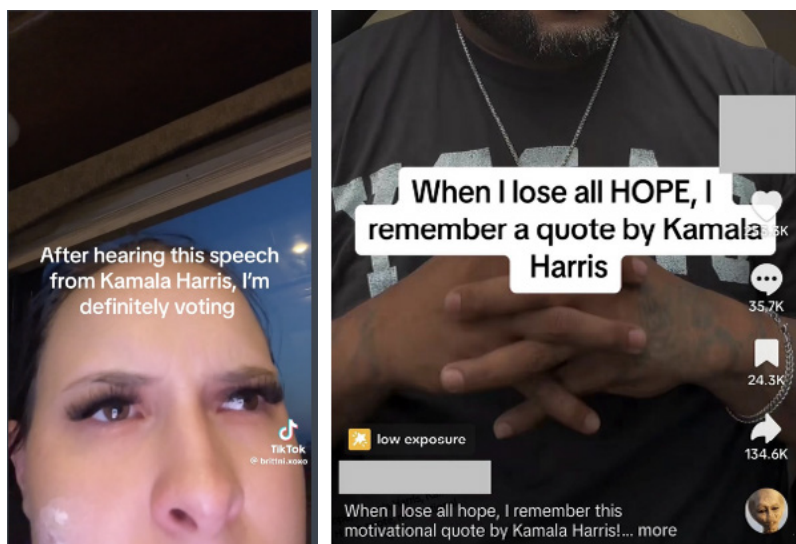
Figure 13. Kamala Harris sexualised deepfake¹⁸⁶



The presidential hopeful was also the victim of a fake audio (deep voice) after Joe Biden announced his withdrawal from the presidential race. This audio was generated by AI and made Vice President Kamala Harris say incoherent things: ‘Today is today and yesterday was today yesterday. Tomorrow will be today tomorrow. So, live today so the future today will be as the past today, as it is tomorrow’. The audio was from an AI-manipulated video of Kamala giving a speech in 2023. While the deep voice was removed from TikTok, it appeared in 2,300 TikTok videos from individual accounts as ‘background sound’. Many of these videos also included descriptions such as, ‘After listening to this speech, I definitely ... “vote” in an ironic tone and continued with the phrase “Don’t vote for her just because she’s a woman”’.

¹⁸⁶ Screenshot of the screenshot uploaded by Gabrielle Settles (*Internet Archive*, 23 August 2024) <<https://archive.org/details/screenshot-2024-08-23-at-5.41.29-pm>> accessed 29 June 2025.

Figure 14. Kamala Harris AI-manipulated audio through TikTok accounts users¹⁸⁷



Donald Trump, on the other hand, was more often the subject of satirical and caricatured deepfakes. Among the cases that gained prominence was the use of deepfakes showing Trump alongside members of the African American community in the United States, in an effort to create a strategic narrative portraying the Republican candidate as a popular figure within that community. These deepfakes were created by Trump supporters. Even a radio host and his team created and shared these kinds of highly fake images on Facebook.

In this regard, Cliff Albright, co-founder of the campaign group Black Votes Matter, in an interview with BBC¹⁸⁸ said that it was an attempt to misinform African American communities, particularly a portion of this demographic population such as young people, in order to obtain their votes.

¹⁸⁷ Screenshots from videos using AI-manipulated audio. Source: (*Media Matters for America*, 22 July 2024).

¹⁸⁸ Marianna Spring, 'Las imágenes falsas creadas con IA para intentar atraer el apoyo de los votantes negros hacia Trump' (*BBC*, 4 March 2024) <<https://www.bbc.com/mundo/articulos/c3g415xgvryo>> accessed 29 June 2025.

It should be clarified that although these deepfakes tend to imitate reality, their manipulated nature can be noticed relatively easily if one pays attention to admiring the characteristics of the photo.

*Figure 15. Donald Trump deepfakes*¹⁸⁹



Finally, it is worth highlighting some common features between the deepfake campaigns undertaken in both Argentina and the United States:

On the one hand, the use of popular figures to allege supposed support for a particular candidate stands out: in the United States this strategy was used by Donald Trump who published fake images and AI-manipulated images of Taylor Swift and her fandom to show support for her campaign. This prompted a reaction from the singer on social media through a statement in which she expressed her concern about AI and the dangers of spreading misinformation while also demonstrating her support for Kamala Harris.

This type of strategy has also been used in Argentina, employing Lionel Messi's image to simulate support for electoral campaigns. It applies to both Javier Milei and Sergio Massa. However, these are simple photomontages or edited television graphics with captions; in other words, they are not manipulated by AI.

¹⁸⁹ Capture of the manipulated images that have circulated on social media.

Figure 16. Disinformation through deepfakes and photomontages of Taylor Swift in the United States¹⁹⁰

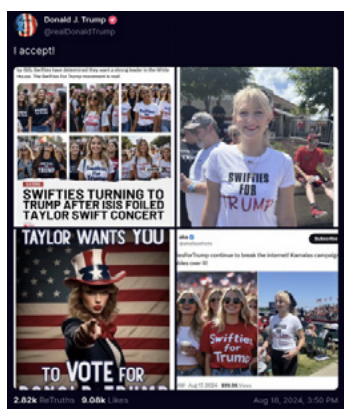


Figure 17. Disinformation through photomontages of Lionel Messi in Argentina¹⁹¹



¹⁹⁰ Screenshot of a tweet from Donald Trump in which he posts deepfakes of Taylor Swift and her fans supporting his campaign. It circulated in traditional media outlets such as La Nación, 'Qué hay detrás de la imagen falsa de Donald Trump con Taylor Swift como protagonista' (*La Nación*, 20 August 2024) <<https://www.lanacion.com.ar/estados-unidos/que-hay-detras-de-la-imagen-falsa-de-donald-trump-con-taylor-swift-como-protagonista-nid20082024/>> accessed 29 June 2025.

¹⁹¹ Both screenshots of the verification performed by <[Chequeado.com](https://chequeado.com)> regarding Messi's alleged support for the campaigns of the two candidates with the highest voting intentions: Corral Ignacio, 'No, Lionel Messi no publicó en sus redes sociales una imagen de un león en apoyo a Javier Milei' (*Chequeado*, 24 October 2023) <<https://chequeado.com/ultimas-noticias/no-lionel-messi-no-publico-en-sus-redes-sociales-una-imagen-de-un-leon-en-apoyo-a-javier-milei/>> accessed 29 June 2025; Corral Ignacio, 'No, Lionel Messi no apoyó públicamente a Sergio Massa, como afirman estas placas falsas' (*Chequeado*, 4 November 2023) <<https://chequeado.com/ultimas-noticias/son-falsas-estas-supuestas-placas-de-tyc-sports-sobre-lionel-messi-y-un-apoyo-a-sergio-massa/>> accessed 29 June 2025.

At the same time, the use of deepfakes has proliferated in the election campaigns of both countries, depicting catastrophic scenarios should the political opponent win the election, as well as associating candidates like Sergio Massa or Kamala Harris with the Communist Party. These instances were easily detectable as fake or AI-generated, but this does not diminish their potential as vehicles for narratives associated with certain political candidates.

As stated, while most deepfakes often responded to satirical nature, deepfakes have also been widely disseminated and circulated to discredit political opponents and consequently misinform the electorate. Although this doesn't establish a direct link between the proliferation of deepfakes and the final vote count, it does represent an objective erosion of trust in information, which is crucial for transparency in presidential elections.

3.2.1.2 Variable-based study

From a first stage of case collection under the established general parameters, that the cases of deep forgeries analysed in this work use the image of the leading candidates for the national executive position, during the established two years, have gained media attention and do not refer to satirical deepfakes, comical or caricatured, it is important to clarify the following.

In the case of El Salvador, although it denotes the proliferation of deepfakes, those available in digital media and social networks are satirical or for entertainment purposes, acting more as a form of political promotion, and without disregarding their value for the relevant political analysis, they do not meet the general conditions proposed for the present study.

On the other hand, disinformation is present and on the rise in Paraguayan election campaigns. However, AI was not a prominent tool used to perpetrate this phenomenon; nevertheless, it was invoked to cast doubt on the authenticity of a video belonging to a candidate. For this reason, although authorities later confirmed the video's veracity, the case is not analysed using the proposed variables, but it is mentioned due to its relevance to the present study.

Ecuador

Ecuador had to hold early presidential elections in 2023 because Guillermo Lasso, then President of Ecuador, activated the cross-dissolution mechanism under article 148 of the Ecuadorian National Constitution. For this reason, the elections would choose who should complete the presidential term until 2025.

In this electoral contest, the main candidates facing off were Daniel Noboa for the National Democratic Action coalition, Luisa Gonzalez for the Citizen Revolution Movement, Christian Zurita for the MC25, Jan Topic for the PSC, Otto Sonnenholzner for Avanza and Yaku Pérez for the DSI.

After no candidate won in the first round, a runoff election was held on 15 October 2023, between the two candidates with the most votes: Daniel Noboa and Luisa González. The candidate from Acción Democrática Nacional (ADN) ultimately won with 51.83% of the vote.

It is worth noting that, following the applied search techniques, the fieldwork indicates that while fake news proliferated during the election campaign, the use of AI, such as deepfakes, was not a prominent vehicle for disinformation. In this regard, after reviewing digital media, news videos and searching social networks like Facebook, Twitter, TikTok and Instagram, only one instance of a deepfake circulated during this election campaign was found.

- Deepfake (content): During a television interview, candidate Luisa González promotes adding fentanyl to the list of permitted drugs in Ecuador. She appears to say:

We must work on the permanent solution, which is access to a healthcare system that doesn't allow a human being to die from addiction. We must include fentanyl in the list of controlled substances. Let's understand that this is a public health problem; meanwhile, we must reach agreements to restore healthcare and education as a public right, of high quality and free of charge. And on that, we will all agree.

- The same video was also reportedly circulated with the logo and graphic design of the Ecuadorian newspaper *EXTRA*.

- **Date:** The deepfake was shared from 7 October 2023. On X (formerly Twitter) it appears exposed on that date, while on Facebook and TikTok, according to the screenshots and links provided by the Fact-checking company AFP, it would be from 8 October.
- **Account/profile or media outlet that exposed it:** Difficulty in identifying the account responsible for viralising the content. They are anonymous profiles. Nevertheless, the manipulated video was circulated and is accompanied by the hashtag: ‘#RedDeTuiterosDemocraticos’ which corresponds to a hyperlink that directs the user to publications related to the same topic – the rejection of Gonzalez and the support for Daniel Noboa.
- **Interactions:** This video was shared over 7,000 times on social media. It was shared on Facebook, TikTok and X (formerly Twitter). On Twitter, it has 363,000 views, 1,100 comments and 1,100 likes and reposts.
- **Deepfake figure:** Luisa Gonzalez, candidate for the Citizen Revolution movement.
- **Context in which it was introduced:** It was introduced on social media eight days before the runoff election in which he would face Daniel Noboa for the presidency of Ecuador.
- **Effects:** On the one hand, fact-checking companies like AFP have debunked that the deepfake may be genuine. However, this type of communication does not achieve the same level of virality as a disinformation deepfake. Furthermore, the news outlet *EXTRA* refuted the headline attributed to it and denied having published the video.
- **Type:** The deepfake originated from a fragment of an authentic interview recording, in which the candidate’s gestures and movements are real, but the audio was manipulated. Her voice was altered to simulate that she was promoting fentanyl use.

- Connection to the campaign: The deepfakes were circulated in the days leading up to the election (15 October 2023). In some cases, they were accompanied by hashtags clearly supporting electoral opponents.
- Source: AFP – Fact-Checking Company.¹⁹²

Figure 18. Ecuador's candidate Luisa González deepfake¹⁹³



Paraguay

The Paraguayan general elections were held on 30 April 2023. The electoral contest involved Santiago Peña for the Colorado Party, Efraín Alegre for the Authentic Radical Liberal Party, Paraguay Cubas for the National Crusade Party, Euclides Acevedo for the New Republic Movement and Jose Luis Chilavert for the Youth Party, with the first three being the main candidates.

By a wide margin, Santiago Peña won the elections (42.74% votes), followed by Efraín Alegre (27.28%) and Payo Cubas (22.91%).

¹⁹² AFP Factual, 'El video de la candidata Luisa González promoviendo el fentanilo en Ecuador es un montaje' (AFP Factual, 23 October 2023) <<https://factual.afp.com/doc.afp.com.33Y768Y>> accessed 30 June 2025. To visit X platform tweet <<https://x.com/mafercedeno97/status/1710755616560451823?s=20>> accessed 30 June 2025.

¹⁹³ *ibid.*

According to the research conducted, there was a considerable increase in disinformation on social media and digital platforms during the election campaign. In this regard, while the use of AI has not been a central feature of the campaign, it is worth noting that the OAS Electoral Observation Mission pointed out:

The Mission also received information that there was no debate between the candidates with the most support in the polls. However, a debate between the two candidates, generated with artificial intelligence tools, circulated on social media, which demonstrates the emergence of new technologies with the capacity to both inform and misinform very quickly in electoral campaigns.¹⁹⁴

In this regard, it's worth noting that the debate that circulated was more of a caricatured, computer-generated representation of the two candidates featured. Nevertheless, after the election results were announced, they were challenged, primarily by Payo Cubas, and later by Alegre. In other words, narratives of electoral fraud began to proliferate, taking advantage of the Paraguayan socio-political context.

This led to protesters gathering in front of the Electoral Court headquarters. The situation escalated with the arrest of Payo Cubas and more than 200 other protesters. It was in this context that a video of former presidential candidate Cubas urging the protesters to call off the demonstration circulated. Some considered the video staged, while his running mate, Stilber Valdés, claimed it was an AI-generated video – a deepfake.

Although this video was released after the elections, it is important to highlight the role that AI played as a method or excuse to question the epistemological value of what is seen, and how, in electoral contexts, deepfakes appear as a point of weakness in the trust in images or videos as sources of knowledge. They can be used by candidates to denounce or cast doubt on what is said or done in these productions. This phenomenon is what has been called the Liar's Dividend, which will be discussed later.

¹⁹⁴ OAS MOE, 'Informe Preliminar de la Misión de Observación Electoral de la OEA en Paraguay' (OEA Press, 2 May 2023) <<https://www.oas.org/fpdb/press/Informe-Preliminar-Paraguay-2023.pdf>> accessed 1 July 2025.

Figure 19. Paraguay 's candidate Payo Cubas video



Video circulated on the TikTok platform

Source: *El Nacional Newspaper* (16 May 2023)¹⁹⁵

< <https://vm.tiktok.com/ZMS5KACtm/> > accessed 1 July 2025.

Guatemala

Guatemala held general elections in 2023. 22 political parties participated in the elections, and the main candidates vying for the presidency were: Bernardo Arévalo for the Semilla party, Sandra Torres for the National Unity for Hope (UNE) party, Manuel Conde for the Vamos party, Armando Castillo for the Vision with Values (VIVA) party and Zury Ríos for the Valor-Unionista party, among others. In the first round, none of the candidates obtained the required majority, necessitating a second round of voting on 20 August 2023. Bernardo Arévalo and Sandra Torres competed for the presidency in this runoff. The Semilla candidate won the election with 60.91% of the vote.

Considering the variables and parameters explained at the beginning of the case analyses, it should be mentioned that in the Guatemalan case there is an absence of publicly available evidence of deepfakes to compare with what is presented in some media.

¹⁹⁵ El Nacional, 'Dudas sobre video de "Payo": ¿Obligado a grabar o hecho por inteligencia artificial?' (*El Nacional*, 16 May 2023) <<https://elnacional.com.py/politica/dudas-sobre-video-payo-obligado-grabar-o-hecho-inteligencia-artificial-n47549>> accessed 1 July 2025.

In other words, while the Guatemalan electoral process was characterised by an increase in disinformation and attacks against political parties, candidates and actors related to the campaign, especially on social networks, these were carried out through digital manipulation techniques but not through deepfakes, that is, through techniques based on AI.

As indicated in the seventh report of the Electoral Observation Mission of Guatemala (MOE-GT), the mission, conducted by seven civil society organisations between 15 July and 5 August, documented 15 campaigns of attack, disinformation and defamation against candidates. These posts were disseminated through TikTok, X (formerly Twitter), and Facebook, before migrating to other platforms. A common characteristic was identified in this type of deployment:

One of the characteristics of these messages was that they originated from recently created accounts or bots (between May and June) and were published systematically or strategically to reach closed networks such as WhatsApp or Signal. Another characteristic was that they were generated from accounts of specific actors, such as former candidates, activists, and political operatives.¹⁹⁶

This indicates that while the presence of deepfakes in the current digital ecosystem could not be confirmed – despite their potential circulation during the election period – the use of AI to amplify and increase disinformation through the planned use of bots is noteworthy. Furthermore, the topics addressed included accusations of corruption, links to criminal groups and false government proposals.¹⁹⁷ That is, there are certain narratives that operate in the flow of disinformation in countries, in this case Guatemala.

The type of digital manipulation in the Guatemalan election campaign, such as cheapfakes and shallowfakes are observable in the count of five falsehoods identified by EFE Verifica (fact-checker) during the election period. Thus, for example, a video of presidential candidate (and later elected President of

¹⁹⁶ Misión de Observación Electoral de Guatemala, 'Séptimo Informe. Agosto de 2023' (2023) 9 <<https://dialogos.org.gt/wp-content/uploads/2025/07/Informe-07-MOE-Gt.pdf>> accessed 1 July 2025.

¹⁹⁷ *ibid.*

Guatemala) Bernardo Arévalo was circulated in which he can be heard saying: ‘We are going to take land from those who have it to distribute it to those who don’t have it’ but far from being a deepfake, it actually corresponded to an interview given on a Guatemalan radio station, in which, on the contrary, a part of the speech was omitted, that is, it was ‘cut short’.¹⁹⁸

On the other hand, there is a scarcity of articles or blog posts mentioning the use of AI tools in the post-election period to commit cybercrimes by exploiting the increased information related to Guatemalan political figures. These attacks use the identities of people like Bernardo Arévalo to ‘validate’ or support an oil investment platform called ‘Oil Profit’. This scenario is similar to the use of deepfakes in the context of the 2024 Mexican elections.

However, the fieldwork did not allow us to verify whether these deepfakes had actually proliferated after the 2023 elections, so they were not studied under the stipulated variables.

Dominican Republic

In this election, eight candidates vied for the presidency, the main contenders being: Luis Abinader of the Modern Revolutionary Party, who was seeking re-election, Leonel Fernández of the People’s Force and Abel Martínez of the Dominican Liberation Party. Ultimately, on 19 May 2024, Luis Abinader emerged victorious, and his term was renewed for a second time until 2028.

In the context of the campaign, a manipulated audio recording, that is, a deep voice recording, went viral:

- Deepfake (content): Audio with the voice of Luis Abinader in which he states that it is necessary to wait until after the elections for the Dominican government to take measures to ‘help’ Haiti by welcoming several million of its citizens in August with temporary permits. In other words, the message concerns easing immigration restrictions for Haitian nationals. He can be heard saying:

¹⁹⁸ Daniella Gastón Allen, ‘Falsos Planes de Expropiación y vínculos guerrilleros, mentiras de la Campaña Guatemalteca’ (*EFE Verifica*, 18 August 2023) <<https://verifica.efe.com/falsos-planos-de-expropiacion-y-vinculos-guerrilleros-mentiras-de-la-campa-na-guatemalteca/>> accessed 1 July 2025.

I understand the pressure they're under from that side, the pressure they're under at their own border. We're going to help them, but not now. It can't be now. I can't just go out and say, 'Yes, come here!' ... Can you imagine me coming out and saying that? now' ... After May, or rather after August, then we'll see how it works, maybe temporary permits would be better ... But they can rest assured.

- Date: There is no exact data on when it was released. The media outlets consulted agree on the date and correspond to 14 May 2024.
- Account or media outlet that exposed it: It is not determined who started its circulation or viral spread in the WhatsApp interface. Circle by means such as screen recording in which the audio is played.
- Interactions: Not recorded.
- Deepfake figure: Luis Abinader.
- Context in which it was introduced: Circulation five days before the elections held on Sunday 19 May 2024.
- Effects: Media coverage in both digital newspapers and television. Furthermore, the co-coordinator of the PRM (Modern Revolutionary Party) presidential campaign and senator for the National District accused the opposition of 'launching a dirty and desperate offensive with AI'.
- Despite the viral spread of this audio, Luis Abinader won the elections.
- Type: AI-generated audio (Deepvoice).
- Connection to the campaign: No specific actors have been identified in the circulation of the audio. With regard to political contestation, it is worth highlighting two aspects that become relevant in electoral contexts:

On the one hand, as mentioned, the audio circulated five days before the presidential elections of 19 May. On the other hand, the content of the manipulated audio is linked to a sensitive and contentious issue for Dominican society: the handling of Haitian migration. In this sense, it is important to contextualise that in the Dominican Republic, in the words of sociologist Haroldo Dilla Alfonso, there is a feeling of ‘anti-Haitianism’. This is not a secondary element of Dominican political culture, but an ‘organising component’. This discourse operates from the hatred inherited from the Trujillo regime, which perceives Haitians as cultural, political and biological aggressors, and on the other hand focuses its attention on Haitian poverty.¹⁹⁹

This context was proven, for example, before the Inter-American Court of Human Rights in the 2014 case of *Expelled Dominican and Haitian v Dominican Republic*.²⁰⁰ There, it was determined that there is a context in which the Haitian population and people born in Dominican territory of Haitian descent suffer discriminatory treatment.

The above does not mean endorsing said scenario but rather presenting it in order to understand the content of the deepfake as a political element in light of the context briefly described. The narrative that is put into circulation has a relevant political dimension.

- Source: Broadcast of the programme ‘El Show’ aired by Color Visión Channel 9 on 14 May 2024 (minute 1.03 to 1.59).²⁰¹ Similarly, El Avance Media denies the existence of a deep voice on the X platform.²⁰²

¹⁹⁹ Haroldo Dilla Alfonso, ‘República Dominicana: cuando la xenofobia se institucionaliza’ (*Nueva Sociedad*, 2019) <<https://nuso.org/articulo/república-dominicana-cuando-la-xenofobia-se-institucionaliza/>> accessed 2 July 2025.

²⁰⁰ Inter-American Court of Human Rights *Case of Expelled Dominicans and Haitians v Dominican Republic* (2014) Judgment of 28 August 2014.

²⁰¹ Color Visión Canal 9, ‘The alleged audio recording of President Luis Abinader | The Midday Show’ (YouTube, 14 May 2024) <https://youtu.be/v539REFeBZE?si=dHz7XB-z1WmK_LU9t> accessed 2 July 2025.

²⁰² El Avance Media (*X (formerly Twitter)*), 14 May 2024 <<https://x.com/elavancemedia/status/1790401755408122100?s=51>> accessed 2 July 2025.

Figure 20. Dominican Republic's candidate Luis Abinader *deepfake*²⁰³

Photo of the capture that circulated on WhatsApp and gained media attention



Panama

Panama held presidential elections on 5 May 2024. Eight presidential candidates participated in the election campaign, the main ones were: José Raul Mulino for 'Realizando Metas' party, Ricardo Lombana for 'Movimiento Otro Camino' party, Martín Torrijos for 'Partido Popular' party and Rómulo Roux for 'Cambio Democrático' party. Also on the ballot were Zulay Rodriguez as an independent candidate, José Gabriel Carrizo for the Partido Revolucionario Democrático and Maribel Gordón and Melitón Arrocha, both also independent candidates.

Finally, with 34.23% of the votes, Jose Raul Mulino was elected.

²⁰³ El Avance Media (X (formerly Twitter)(n 202).

Regarding the proliferation of AI tools, it should be noted that few cases have been identified, most of which corresponded with cheapfakes and shallowfakes of an ironic or satirical nature, in which attempts were made to delegitimise candidates for reasons related to their private lives. Nevertheless, one case of deepfake was identified:

- Deepfake (content): The deepfake shows candidate Romulo Roux presenting the return of mining as a campaign proposal. He is seen and heard saying:

I proposed clear topics and solutions in the debate. The mine is returning, and that will be the first decree that I will sign. Safety is guaranteed by the mine, therefore all personnel will coordinate with them. We need more mines so that young people can study. With the exploitation of all mineral products, we will generate the jobs that are needed. I am Rómulo Roux. Mining is back!

- Date: 26 February 2024.
- Account or media outlet that exposed it: The deepfake appears to be linked to the account @hablapanamanoticias. The audio appears to be from that account.
- Interactions: Its circulation was only verified on Instagram where it received few 'likes' (124). No other type of interactions, shares or plays were recorded.
- Figure of the deepfake: Romulo Roux.
- Context in which it was introduced: During the beginning of the election campaign.
- Effects: No significant effect was recorded. There is no record of it being denied.
- Type: Deepfake using video and voice. However, it's worth noting that the lip movements don't match the audio of the deepfake. The manipulation of the speech of the person pretending to be Roux is evident in the out-of-sync mouth

movements. Furthermore, the audio contains syntax errors, as can be seen in the transcript above. It's likely that the level of convincing this deepfake could generate wasn't what its creator and distributor intended, which explains why it didn't go viral and only garnered 124 'likes' on Instagram.

- Connection to the campaign: Despite the errors or signs that indicate it's a deepfake, the video was released after the campaign had begun but well before the election dates. However, the deepfake also establishes a narrative around a contentious issue in Panamanian society.

The theme explored in the deepfake narrative is the establishment of mining operations. In October 2023, Panama experienced a major mining conflict following the passage of Law 406, which established the concession agreement between the state and Minera Panamá, a subsidiary of First Quantum Minerals (FQM). The response included protests, road blockades, mass demonstrations and even shortages in stores.

Amid the social unrest, on 28 November, the court declared the aforementioned law establishing the contract with the mining company unconstitutional. Ultimately, the FQM subsidiary closed, representing a severe economic blow to the country. Despite this, most presidential candidates maintained the need for its closure during their campaigns.

On the other hand, the deepfake takes advantage of presidential candidate Roux's professional background, given that as a lawyer he belonged to one of the firms (Morgan & Morgan) sponsoring the mining company. This last aspect contributes to the perceived authenticity of the deepfake.

Understanding this context already and regarding the effects of pointing out a widespread consequence following this type of technology, a journalistic study indicated that publications circulating on social networks used AI tools (those mentioned above) to spread misinformation about campaign proposals or private life, which led the Panamanian Electoral Tribunal to understand that the content they were infringing. Rules in the Electoral Code prohibit the unauthorised use of political party symbols or the personal image of third parties, so it ordered platforms like META to remove such content. Despite this, the platform did not deactivate them.

As mentioned, although in the specific case of deepfake this action by the court is not recorded, perhaps due to the little interaction it received on Instagram, it is still relevant information to have parameters of the effectiveness or not in the compliance with orders issued by state courts by the platforms where AI productions circulate.

— Source: The example was received by Martin Pablo Villar.²⁰⁴

As is evident, the Panamanian deepfake indicates rather the irruption or attempt to act as a disinformation tool, while at the same time leaving a valuable element of analysis such as the use of local narratives to achieve epistemological validation by the electorate.

*Figure 21. Panama's candidate Romulo Roux deepfake*²⁰⁵
Deepfake Rómulo Roux on Instagram in the account 'hablapanamanoticias'.



²⁰⁴ Martin Pablo Villar, 'El Rol de la Inteligencia Artificial Generativa (IAG) en la Desinformación Electoral' (Friedrich Naumann Foundation 2024).

²⁰⁵ Hablapanamanoticias Instagram account <https://www.instagram.com/reel/C31cK-imsjOR/?utm_source=ig_web_copy_link> accessed 2 July 2025.

Mexico

The Mexican federal elections were held on 2 June 2024. The presidential candidates were Claudia Sheinbaum Pardo for Morena, Xóchitl Gálvez Ruiz for the Fuerza y Corazón por México coalition (National Action Party, Institutional Revolutionary Party and Party of the Democratic Revolution) and Jorge Álvarez Máynez for the Movimiento Ciudadano party. Ultimately, with 59.76% of the vote, Claudia Sheinbaum was elected President of Mexico.

Regarding the use of generative AI in the 2024 elections, it is important to mention that at least 44 cases of this type of technology being used for political purposes were recorded.²⁰⁶ However, as in Panama, many of the cases identified as deepfakes were actually cases of shallowfakes and cheapfakes (more accessible and simpler editing tools). The report prepared by the Friedrich Naumann Foundation in conjunction with the EON Institute recognises five cases due to their extensive media and social media coverage,²⁰⁷ among them is the following deepfake case, whose characteristics meet the general parameters established for the analysis of this thesis:

- Deepfake (content): The deepfake shows Claudia Sheinbaum inviting people to invest in a financial platform. In the video, she is seen in the Mexico City government offices. In the background, a white screen with black lettering reads ‘Financial Platform 2024’. She is heard saying: ‘Those 4,000 pesos will change your financial future forever; your success is just a few clicks away’.
- Date: There is no data on when it was disseminated and/or circulated. The news articles consulted date from 25 January 2024, meaning there is agreement on the date it gained media attention; however, then-candidate Claudia Sheinbaum denied the deepfake on 24 January. Therefore, a time-frame approximately corresponding to those dates will be considered.

²⁰⁶ Claudia May del Pozo and Daniela Rojas Arroyo, ‘IA Generativa y su influencia en las elecciones de México 2024’ (Friedrich Naumann Foundation 2024).

²⁰⁷ *ibid.*

- Account or media outlet that circulated it: There is no record of who originated it or who disseminated it. The circulation of the video is recorded as ‘agenda topic’ of the newspapers in January 2024.
- Interactions: Not recorded.
- Figure of the deepfake: Claudia Sheinbaum.
- Context in which it was introduced: Closing period of the Pre-Campaign in Mexico. The elections took place on 2 June 2024.
- Effects: Candidate Sheinbaum denied the video in a Facebook Live broadcast, on 24 January 2024. She said: ‘It’s a video they’re making with artificial intelligence because you’ll see it’s my voice, but it’s a fraud’.
- Type: Deepfake manipulated video and voice.
- Connection to the campaign: Several considerations are necessary on this point. Firstly, it is important to note that this deepfake dissemination has been identified as part of a trend in the 2023–2024 Mexican electoral cycle: the use of generative AI for cybercrimes.²⁰⁸ In other words, deepfakes were designed and created using the identity of a presidential candidate (in this case) to exploit her image and the trust or validation she can inspire in the electorate to commit cybercrimes, such as phishing scams.
- The latter coincides with a context of increase in this type of crime in Mexico.²⁰⁹

²⁰⁸ May del Pozo and Rojas Arroyo (n 206) 12.

²⁰⁹ *ibid.*

- It is also worth mentioning that this case was not an isolated incident; additional cases like the one presented here have been detected and circulated, all sharing the same narrative: Claudia Sheinbaum inviting people to invest in nonexistent financial platforms, oil and other things. Furthermore, images of other public figures have also been used in this type of generative AI application.²¹⁰
- On the other hand, although there is a five-month gap between the date of dissemination and election day, it should be noted that the deepfake was disseminated during the period almost at the beginning of the inter-campaign period, a phase of the electoral process during which political parties resolve any differences regarding the internal selection of their candidates. This period is not without limitations for those running for office; they are prohibited from engaging in premature campaigning to ensure fairness in the electoral contest. Examples of prohibitions include calls to vote and proselytising; the dissemination of political propaganda for informational (non-electoral) purposes is permitted. In this regard, it is worth analysing how these types of deepfakes operate in light of these limitations, until it is verified whether they are indeed attempts at cybercrime. At the same time, it undermines the electorate's trust, as cases like this are disseminated throughout the entire electoral process.
- Source: Anna Lagos.²¹¹ Screenshot is shown below in Figure 22.

²¹⁰ May del Pozo and Rojas Arroyo (n 206).

²¹¹ Anna Lagos, 'Claudia Sheinbaum, víctima de deepfake: su imagen fue utilizada en un video fraudulento generado por IA' (*WIRED*, 25 January 2024) <<https://es.wired.com/articulos/claudia-sheinbaum-victima-de-deepfake-su-imagen-fue-utilizada-en-un-video-fraudulento-generado-por-ia>> accessed 3 July 2025. Similarly Christopher Calderón, 'Mexico Elections 2024: Deepfakes and fake news win in the elections' (*El Financiero*, 4 June 2024) <<https://www.elfinanciero.com.mx/elecciones-mexico-2024/2024/06/04/deep-fakes-y-fake-news-marcaron-el-escenario-electoral/>> accessed 3 July 2025; Channel 22 Mexico, 'A lie detector capsule about deepfakes involving the President of Mexico' (*Facebook*, 6 November 2024) <<https://www.facebook.com/Canal22Mexico/videos/c%C3%A1psula-detector-de-mentiras-sobre-los-deep-fake-en-que-involucran-a-la-president/864321199161687/>> accessed 3 July 2025.

Figure 22. Mexico's candidate Claudia Sheinbaum deepfake



Venezuela

The presidential elections in Venezuela were held on 28 July 2024. Ten political figures vied for the presidency. The candidates with the highest voter support were Nicolás Maduro of the Great Patriotic Pole Simón Bolívar (who was already serving as president) and Edmundo González Urrutia of the Democratic Unity Roundtable (MUD).

In the context of the elections, the National Electoral Council (CNE) declared Nicolás Maduro the winner, despite not having published the tally sheets with the detailed results. Meanwhile, the opposition, led by María Corina Machado and Edmundo González, maintains that they were the true winners at the polls.

In the Venezuelan electoral process, disinformation was commonplace. Within this ecosystem of (dis)information, narratives circulated through AI tools:

- Deepfake (content): The Strategic Operational Commander of the Venezuelan Armed Forces, through his official account on X (formerly Twitter), published a series of screenshots of a manipulated video. The screenshots showed a whiteboard with writing on it that displayed 'Machado's plan to eliminate the Armed Forces'. These captures were accompanied by a quote from Hugo Chavez, Supreme and Eternal Commander of the Bolivarian Revolution:

There will be those who try to exploit these difficult times to continue their efforts to restore capitalism, neoliberalism, and ultimately destroy our nation. But they will not succeed. Faced with these new challenges – however great they may be – the response of all patriots, all revolutionaries, all of us who feel our nation in our very bones, as Augusto Mijares would say, is unity, struggle, battle, and victory. Today we have a homeland! And whatever happens, under any circumstances, we will continue to have a homeland.

However, it was found that in the original video of Edmundo Gonzalez and Corina Machado, which had been recorded during an event at the Metropolitan University facilities in an academic context, the blackboard was blank.

- Date: 22 June 2024. While these screenshots were not published in the context of an official campaign, it should be noted that these limits have been relaxed de facto, especially due to the emergence of new technologies. It should also be considered that by that time, the candidacies of the political opponents had already been registered with the National Electoral Council (CNE) (21 March).
- Account or medium that put it into circulation: Strategic Operational Commander of the Armed Forces GJ Domingo Hernandez Lárez in his account of X.
- Interactions: The tweet, accompanied by the images, received 732,600 views, 17,000 comments, 12,000 retweets and 11,000 likes. Additionally, readers, using the tool provided by X to add context to the tweet, indicated that the shared information was false.
- Deepfake figures: Maria Corina Machado and Edmundo González.
- Context in which it was introduced: One month before the presidential elections held on 28 July 2024.

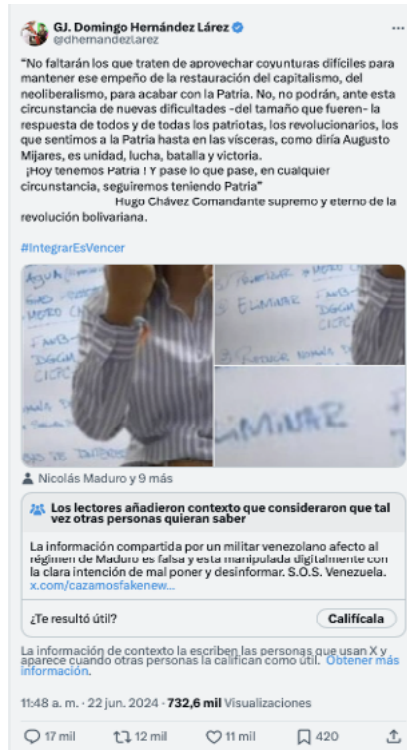
- Effects: Readers of the Twitter post ‘added context’ using the platform’s built-in function, revealing the digital manipulation. It was also debunked by fact-checkers such as ‘Fake News Hunters’.
- Typology: Deepfake by image in which the context in which the opposition figures of Nicolás Maduro were found is modified by the use of AI.
- In this sense, it is important to mention that this case shows how techniques of this type allow not only altering or creating faces and voices, but also altering elements in a scene to make it seem authentic and thus determine the context in which the political references are embedded.
- Connection to the campaign: On this occasion, the deepfake was circulated by a high-ranking member of the Armed Forces, who reports directly to the National Administration, headed by President Nicolás Maduro, who was also running for re-election. It should be added that the content was disseminated after the registration of presidential candidates and a month before the elections.
- Source: Tweet in X by GJ Domingo Hernández Lárez.²¹² Similarly, the example is cited as an example of deepfake in Rubio, García Rossó and Sánchez.²¹³

²¹² @dhernandezlarez (*X (formerly Twitter)*, 22 June 2024) <<https://x.com/dhernandezlarez/status/1804526869993103837>> accessed 3 July 2025.

²¹³ Rafael Rubio Nuñez, Roelisabell García Rossó and Vanessa Sánchez, ‘Desinformación: como entenderla, combatirla y protegerse de sus efectos’ (Friedrich Naumann Foundation for Freedom Regional Office for Latin America 2024) <https://www.freiheit.org/sites/default/files/2024-12/03dic_paper_desinformacion_versioncorta02.pdf> accessed 16 January 2026.

Figure 23. Venezuela's opposition deepfake disseminated by Domingo Hernández Lárez X account

Screenshot of the tweet belonging to Domingo Hernández Lárez (22 June 2024)



Uruguay

The Uruguayan general elections were held on 27 October 2024 (first round) and 24 November of that year (second round). The presidential candidates vying for the highest office in the Uruguayan executive branch were: Yamandú Orsi for the Broad Front, Álvaro Delgado for the National Republican Coalition Party, Andrés Ojeda for the Colorado Party, Gustavo Salle for Sovereign Identity, Guido Manini Ríos for the Open Cabildo Party and Pablo Mieres for the Independent Party. After the first round, none of the candidates obtained the required majority of votes, so a second

round was held between the two candidates with the most votes (Y Orsi and A Delgado) to determine the winner. Finally, on 24 November, the Uruguayan people elected Yamandú Orsi as President of the Oriental Republic of Uruguay.

For further context, it is important to mention that on the same date as the presidential elections, 27 October, two other issues were also put to a plebiscite: pension reform and the power for the state to carry out nighttime raids, which is prohibited by the Uruguayan Constitution. Both proposals were rejected.

In Uruguay, as in other cases presented previously, while there was a high level of disinformation in the electoral context, the presence of deepfakes that gained media attention was incipient. Nevertheless, one of the cases that has been identified is described below:

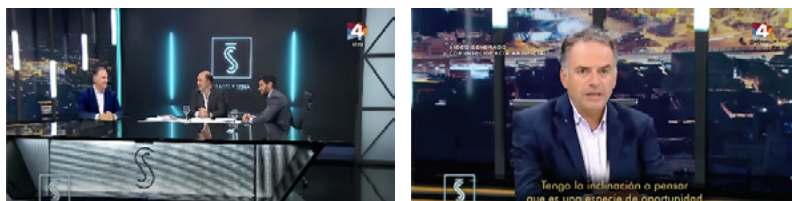
- Deepfake (content): In the broadcast of the programme ‘Santo y Seña’ on Channel 4, during the prime time schedule, the voice and image of Yamandú Orsi (presidential candidate for the Broad Front) were recreated using AI after the candidate declined to appear on the programme. At the table, Orsi was ‘seated’ to debate Andrés Ojeda, who was present in the studio and was one of the leading presidential hopefuls (after Orsi and Delgado) and the candidate for the Colorado Party, one of Uruguay’s traditional political parties. It should be noted that while the ‘deepfake Orsi’ presentation included a label indicating its AI origin, it was barely legible.
- Date: 13 October 2024.
- Account or media that put it into circulation: It took place in the programme ‘Santo y Seña’ of Channel 4, without the authorisation of the candidate.
- Interactions: The level of views of the live programme is not recorded, however, on YouTube the broadcast of the programme has 10,259 views as of the date of consultation.
- Figure of the deepfake: Yamandú Orsi.
- Context in which it was introduced: Days before the first round that took place on 27 October 2024.

- Effects: It had media repercussions because in the deepfake: a) statements not made by the candidate were disseminated, b) the ‘interview’ took place in the form of a debate with another candidate (Andrés Ojeda-Colorado Party) who was physically present in the studio, c) the danger is noted that fragments of the ‘Fake Orsi’ may make it go viral on social media as if it were true and to start fake news campaigns. Despite this, Yamandú Orsi won the elections.
- Type: Deepfake by image and by voice.
- Campaign connection: It took place during the closing campaign events prior to the first round of the presidential election, following the candidate’s refusal to appear on a programme. A debate on public agenda issues was held between candidates from opposing political parties, one per person. It is another in person, which implies not guaranteeing fairness in the electoral contest, since one of the candidates has no factual possibilities of contradicting what the political opponent proposes in the debate.
- This case highlights, on the one hand, the use of this AI tool by traditional media and journalism, and on the other hand, the lack of consent for its use. Not only was a deepfake created using a candidate’s identity just days before the end of the campaign as retaliation for rejecting an invitation to a television programme, but the candidate was also exposed to potential editing or segmentation of what was made to be said (either in audio or video) during prime time on Uruguayan television. This has the potential to negatively impact the candidate’s campaign and misinform the electorate, while also violating access to truthful information and a level playing field in the electoral contest, as the supposed candidate was subjected to a debate on campaign issues with another candidate present in the television studio.

- Source: 13 October 2024 broadcast of the programme ‘Santo y Seña’.²¹⁴ Tweet posted by journalist Ignacio Alvarez, host of ‘Santo y Seña’.²¹⁵

Figure 24. Uruguay’s candidate Yamandú Orsi deepfake

Screenshots from the programme broadcast



Ecuador 2025

The Ecuadorian general elections of 2025 were held on 9 February. 16 candidates ran for president, most notably: Daniel Noboa, seeking re-election, representing the National Democratic Action Party (ADN), Luisa González for the Citizen Revolution Movement (RC), Leonidas Iza for the Pachakutik Plurinational Unity Movement (MUPP) and Andrea González Nader by Patriotic Society Party ‘January 21’ (PSP).

Because no candidate received enough votes on 9 February, a runoff election was held between Daniel Noboa and Luisa González on 13 April. The candidate from ADN was re-elected to lead the country as president.

It’s fair to say that between campaign periods, the candidates in the runoff election have continued their agendas and digital campaigns. While campaign events are not permitted between the two periods, digital strategies are being deployed to undermine and distort the candidates’ proposals in the electoral contest.

Thus, it is important to mention that the OAS Electoral Observation Mission found that during the electoral campaign there was a high level of political polarisation and ‘detected the use of disinformation strategies aimed at influencing public perception

²¹⁴ Canal 4, ‘Santo y Seña - Programa del Domingo 13/10/2024’ (YouTube, 15 October 2024) <https://youtu.be/PP1wmV3CqMw?si=2HdyoiDNO9w_iMc> accessed 4 July 2025.

²¹⁵ @igalvar71 (X (formerly Twitter), 3 October 2024) <<https://x.com/igalvar71/status/1841878323204403311>> accessed 4 July 2025.

and discrediting both contenders'.²¹⁶ In this regard, the mission took knowledge about 'the continuous dissemination of previously debunked false information, including through journalists and public figures, as well as the use of artificial intelligence and the advancement of technologies such as deepfake'.²¹⁷

The case presented below took place less than a month before the presidential debate on 23 March, ahead of the presidential runoff election:

- Deepfake (content): The AI-manipulated video alters an interview with Luisa González broadcast on the Teleamazonas television channel in 2023, making her say: 'The best thing for the country is to abandon dollarisation'. This manipulation was disseminated by the account @BancadaADN accompanied by the narrative:

Ecuador doesn't want any more lies. The constitutional amendment project presented by the Correa administration is just another trick to introduce an alternative currency to the dollar. Now that we're in campaign season, they say they won't dedollarize, but that has been their intention for the last 15 years.

- Date: 27 February 2025 on X by the ADN Bank account.
- Account or media that put it into circulation: The ADN Caucus (@BancadaADN), account of X (former Twitter) of the legislative caucus of National Democratic Action that follows the leadership of Daniel Noboa, while he has run as representative of said party.
- Interactions: This video received 8.8k views, 174 shares and 263 likes on an official account with 5802 followers.
- Deepfake figure: Luisa Gonzalez.

²¹⁶ OEA MOE, 'Informe Preliminar de la Misión de Observación Electoral de la OEA en Ecuador para la segunda vuelta de las elecciones presidenciales' (OEA 2025) 9 <https://www.oas.org/fpdb/press/2025_MOE_Ecuador_Segunda_Vuelta_Preliminar_ESP.pdf> accessed 4 July 2025.

²¹⁷ *ibid.*

- Context in which it was introduced: During the campaign.
- Effects: Debunked by fact-checker company: Lupamedia.
- Typology: Deepfake video and voice from an old interview that had circulated in 2023 and from whose manipulation various deepfakes have been created.
- Connection to the campaign: The circulation of this deepfake from the X account of the ruling party's congressional bloc occurred during the pre-campaign period leading up to the presidential runoff. As previously stated, despite the established electoral calendar, the fight for the popular vote does not cease until election day. In this context, the National Assembly had to vote on the proposal to form an ad hoc commission to address the constitutional amendment recognising the dollar as Ecuador's official currency, given that its economy relies on it. The bill was presented by assembly members from the Citizen Revolution movement, whose president is Luisa González. This situation, combined with the overall electoral context, created fertile ground for the deepfake to circulate.
- It's worth noting that one striking element is that the disinformation circulated through the deepfake in question originated from an account associated with the legislative bloc of the ruling political party, ADN (@BancadaADN). In other words, as in the Venezuelan case, it is the state agents themselves who are pulling the strings, making this type of AI-facilitated disinformation available to the algorithms. In this particular case, it involves the profile on X of a member of the very body responsible for approving Ecuadorian legislation.
- On the other hand, it should be highlighted that the narrative about 'moving away from dollarisation' also circulated during the 2023 presidential campaign, even though this policy was not established as a government plan by the candidates of the Citizen Revolution party to the National Electoral Council.

— Source: Fact Checker Lupa Media²¹⁸ and Original tweet which disseminated disinformation.²¹⁹

*Figure 25. Ecuador's candidate Luisa Gonzalez deepfake*²²⁰
Screenshot of the tweet from the ADN legislative bloc account (27 February 2025)



Based on the above, it is necessary to point out that the established time frame enables the comparison of the Ecuadorian presidential elections in 2023 (extraordinary) and 2025.

²¹⁸ Lupa Media, '¿Luisa Gonzalez propuso salir de la dolarización?' (*Lupa Media*, 27 February 2025) <<https://lupa.com.ec/verificaciones/luisa-gonzalez-dolarizacion-adn/>> accessed 4 July 2025.

²¹⁹ @BancadaADN (*X (formerly Twitter)*, 27 February 2025) <<https://x.com/bancadaadn/status/1895174409351987657?s=51>> accessed 4 July 2025.

²²⁰ *ibid.*

These two assumptions, under the application of the analysis and data search criteria mentioned above, show that although disinformation was a constant in both elections, the development and consolidation of deepfakes as a means for this purpose is only noticeable in the 2025 elections.

During the 2023 Ecuadorian elections, only one deepfake has been identified among those recorded in traditional media or social networks and which gained widespread media attention. This deepfake was accessed through the archives of Lupa Media (an Ecuadorian fact-checker).

On the other hand, the increased proliferation of deepfakes in the 2025 election campaign is verified in a report by the aforementioned fact-checker, which details the manipulation of the electoral campaign in Ecuador:

During the 2025 presidential elections, disinformation was deployed in a sophisticated manner, taking advantage of key moments in the campaign and highly emotional issues. Its circulation was strategic, with narratives designed to polarize, misinform, and manipulate public opinion.²²¹

The study verified 247 viral content pieces, most of which focused on candidates Daniel Noboa and Luisa González, while 19.6% of said content was manipulated or created with AI.

Aspects highlighted in the report include the orchestrated nature of the phenomenon: its circulation did not respond solely to spontaneous dynamics: ‘in many cases coordinated patterns, well-structured digital ecosystems and amplification tactics with clear political objectives were identified’.

This prominence of deepfakes in the electoral contest is also observable, for example, in the reports issued by the OAS Electoral Observation Mission in both 2023 and the 2025 presidential elections. While for the 2023 elections it makes no allusion to the presence of disinformation through deepfakes, for the year 2025, the mission notes this particularity of the electoral campaign (it should be mentioned that it takes up the monitoring carried out by Lupa Media mentioned above).

²²¹ Lupa Media, ‘Disinformation on demand: how the electoral campaign was manipulated in Ecuador’ (*Lupa Media*, 9 April 2025) <https://drive.google.com/file/d/114zSU8g7X2_DCV-TD-ti138dKGAIUE0Z/view> accessed 4 July 2025.

Furthermore, it is important to note that as the analysis of various media outlets and social networks progressed, not only were deepfakes found that use a candidate's image to manipulate what they say or do, but also deepfakes that use AI to alter real news programmes to transmit false information, or that create avatars through deepfake technology designed to emulate presenters or journalists presenting reports (real or manipulated). These can use graphics and logos from real newspapers (as in the Ecuadorian case in 2023) or those created by AI.

The latter has potential interference, especially in those regions or countries where news from foreign agencies is consumed: this was detected in countries such as Ecuador (by Lupa Media) or Venezuela.²²²

3.3 Conclusions of Chapter 3

The cases presented demonstrate that deepfakes, far from being a mere possibility, has entered the electoral landscape in Latin America and the United States. The role of generative AI in Argentina in 2023 is noteworthy, as it allowed for the large-scale testing of deepfakes' production and circulation. Some have exhibited embryonic characteristics of technological sophistication, which has thus far allowed for their identification as falsified. Among these deepfakes, audio deepfakes are the most difficult to verify.

On the other hand, one observable aspect in all the elections of the countries analysed is the massive proliferation of satirical or parodic deepfakes. However, this does not mean that the number of deepfakes used for electoral purposes is trivial or less relevant; quite the contrary.

So far, at least one case of deepfake has been identified with the aim of discrediting a political opponent. This must be analysed in light of all manipulated content that may have been lost in the digital ecosystem, as well as content that could not be debunked or verified due to a lack of technology for that purpose, and content that may not have been made public.

²²² Lupa Media, 'News Reports Altered with Deepfakes: The Disinformation Strategy in the Election Campaign' (*Lupa Media*, 2 April 2025) <<https://lupa.com.ec/explicativos/deepfake-noticieros-desinformacion-campana-electoral-2025/>> accessed 4 July 2025.

Among the number of political deepfakes that attempted to delegitimise or discredit electoral opponents, the following characteristics stand out: the organised nature of the dissemination of deepfakes on popularly used platforms such as X, mainly, and TikTok, Instagram and Facebook.

On the other hand, it has been verified that in America deepfakes are not only produced by partisan users, but also by politicians themselves, as well as representatives of state powers in the exercise of said functions (as in the case of the ADN legislative bloc of Ecuador in 2025) and owners of the platforms on which these deepfakes are dumped (Elon Musk).

Finally, it is worth highlighting the role that narratives play in deepfakes, characterised by a large local component that may elicit some kind of reaction from the electorate.

In light of these findings, the next chapter proposes to establish, together with the existing theoretical framework, an exposition on the dynamics under which deepfakes operate in the information ecosystem, particularly in electoral contexts, and will attempt to explore the regulatory reaction of countries to this situation in light of a human rights approach.

4. Deepfakecracia and human rights

Where politicians fail, deepfakes succeed.

This chapter aims to conceptualise the features observed in the empirical analysis presented in Chapter 3, in order to determine the features of the American deepfakecracia, which means highlighting the characteristics that have been identified in the use of deepfakes as disinformation tools during the presidential elections studied.

Based on this characterisation, the aim is to determine what the legislative reactions have been, especially in Latin American countries, taking into account that their participation in AI governance is not made visible by academia in a systematic way to date.

Finally, it will be highlighted whether the human rights approach is present in these legal responses and whether the human rights paradigm in force to date is sufficient to address the scenario posed by the irruption of AI through deepfakes in the electoral field.

4.1 American deepfakecracy: Electoral deepfakes

The expansion and accessibility of generative AI, while it can be used for multiple purposes such as educational or artistic ones, brings with it significant concerns for democracy in terms of electoral transparency and citizens' access to information.

AI-generated disinformation was identified as one of the main risks according to the World Economic Forum's 'Global Risks 2024' report, which mentions the potential of generative AI to radically redefine how election campaigns are conducted and how it can distort public perception during elections.²²³ Furthermore, the UN Secretary-General's expert advisory body on AI, in its 2024 Final Report, identified as a primary concern the potential harm AI can cause to information integrity (for example, through disinformation).²²⁴

It should be added that, while techniques for discrediting political opponents or deliberately spreading false information have always existed, the techno-social context is not the same. Since the emergence of the internet and the expansion of social networks as the main channel for disseminating and consuming news, the acceleration in its spread and reach has introduced a new dimension to disinformation. This, coupled with AI techniques, places trust in democratic institutions and the traditional rules of the electoral process at a critical crossroads.

According to the 2024 Latinobarometer survey, both political parties and Congress, institutions characteristic of democracy, have experienced democratic disillusionment, even though more people believe Congress is necessary (53%) than political parties (50%). In line with this, the same report states that four out of ten citizens criticise political parties, Congress and the opposition.²²⁵

To further complicate the context in which deepfakes are introduced as a disinformation tool, the basic institutions of democracy – political parties, congress, the judiciary and the government – have declining levels of public trust. Likewise, electoral institutions have suffered a decline from 2006 to 2024, an aspect that must be analysed in light of new technologies. In this regard, it is

²²³ World Economic Forum 'Global Risks Report 2024' 19th ed. (2024) <https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf> 100.

²²⁴ United Nations Secretary-General, High-Level Advisory Body on Artificial Intelligence, 'Governing AI for Humanity. Final Report' (United Nations 2024), 29 <https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf> accessed 8 July 2025.

²²⁵ Corporación Latinobarómetro, 'Latinobarómetro 2024: La Democracia Resiliente' (2024), 48.

worth adding that the Latinobarómetro suggests that social media enjoys even more trust than the government,²²⁶ despite the fact that 76% of Latin Americans believe that false information circulates on social media.

Thus, it is reasonable to question how the emergence of deepfakes operates in a critical regional context. Does the deliberate use of these technologies undermine democratic experience, increase its fragmentation or is it simply ignored in a scenario marked by the constant circulation of fake news?

Based on this, the analysis presented in Chapter 3 of this thesis was proposed. It demonstrated that all countries that have held presidential elections since 2023 have experienced, at least once during the electoral campaigns, disinformation amplified by the use of deepfakes, as these carry narratives and epistemological challenges tailored to the national context in which they are disseminated.

These deepfakes would fall under an analytical category that is still under construction: electoral deepfakes. This refers to AI-generated deepfakes disseminated during election campaigns, whether in digital or traditional media. These may serve a purely satirical or propagandistic purpose, but when used as disinformation tools, they serve the purpose of deception, distortion of narratives, defamation, discrediting or threatening the political opponent.

It should also be noted that deepfakes appear in the digital ecosystem combined with other types of digital productions such as *shallow fakes* and *cheapfakes*. All of these techniques can be used to spread disinformation; however, the difference lies in the fact that deepfakes are based on generative AI, while the aforementioned refer to edits made with basic digital editing tools, generally produced manually with applications like Photoshop. Examples of these techniques include slowing down videos, interrupting dialogue or taking quotes out of context to convey a message that was not the original.

²²⁶ Corporación Latinobarómetro (n 225) 59.

Now, continuing with the topic of electoral deepfakes, it's important to mention that fieldwork reveals an increase in their proliferation during presidential elections, specifically in the 2023 and 2025 elections in Ecuador. Furthermore, when comparing the 2023 Ecuadorian election campaign with the Argentine campaign, the latter clearly demonstrates the emergence of deepfakes as tools for disinformation and political propaganda.

It should be noted that while the pre-eminence of this tool in the digital ecosystem took the form of satire or mockery, the role of those pursuing an electoral goal did not diminish. Its introduction into the digital ecosystem undermines epistemic trust²²⁷ and over-contaminates public discourse.

4.1.1 Local narratives

An important element in the circulation of American deepfakes is that they acted as vehicles for disseminating local narratives. This is the case, for example, of the deep voice recording of President Abinader circulated on WhatsApp, in which he speaks on a topic that could anger Dominican voters.

This confirms, following the deepfake production model outlined in Chapter 2, the human element behind them, as well as the use of false information tailored to the local context, which aims to elicit emotional reactions from users (affective bias). It's important to clarify that this doesn't mean ignoring the existence of overarching and general narratives, such as allegations of electoral fraud. These narratives are presented in a techno-social medium, such as social media platforms.

The importance of this element lies in the fact that narratives are ways of telling a story. They have a performative character. They allow us to construct and reconstruct events for others. At the same time, it is an exercise of power: it can create or tell stories from a particular perspective and interest. This is what happens with the narratives transmitted by deepfakes. It is not something innocuous; rather, ideas travel with them.

²²⁷ John Twomey and others, 'Do Deepfake Videos Undermine Our Epistemic Trust? A Thematic Analysis of Tweets That Discuss Deepfakes in the Russian Invasion of Ukraine' (2023) 18(10) PloS one 17 <<https://doi.org/10.1371/journal.pone.0291668>>.

4.1.2 Technopolarity

In addition to narratives tailored to the local sociopolitical context, deepfakes have been identified that exploit the identities of presidential candidates to take advantage of their perceived credibility and defraud people through cybercrimes. Furthermore, the use of avatars or the creation of fictitious television programme hosts, posing as foreign news agencies to circulate false information about the candidates, has been identified.

On the other hand, it is important to highlight not only the local contexts that are exploited as input to transmit false data, but also the international technological context based on AI and social networks: currently the agora par excellence where public debate takes place has migrated onto the digital space. This in turn requires thinking about how this dialogue takes place, how it can be improved, what its weaknesses are and what threatens it.²²⁸

The ubiquity of digital social media platforms and the advancement of artificial intelligence have posed new challenges to the integrity of information in electoral processes, such as the proliferation of deepfake videos, the increase in digital political violence against candidates, disinformation campaigns through micro-segmentation aimed at potential voters through recommendation algorithms, as well as the use of technology to perpetrate digital political violence based on gender, for example.²²⁹

In the platformisation of democratic process, the electoral game is not only witnessing the reconfigurations brought about by generative AI, but also a paradigm shift in the very foundation of the international system. Alongside globalisation and technological advancements, a new power centre has emerged: technology companies and/or magnates.

Thus, technopolarity becomes evident in these technological actors who embody their own economic and political interests, challenging the power of the state and, along with it, the rules of the democratic game. In this regard, the High-Level Advisory Body on AI has also expressed that one of the greatest concerns

²²⁸ UNESCO, 'Inteligencia Artificial y Democracia' (2024) <https://unesdoc.unesco.org/ark:/48223/pf0000389736_spa> accessed 6 July 2025.

²²⁹ *ibid* 11.

is the harm related to inequalities stemming from the disparity in the control and ownership of AI technologies, such as the concentration of wealth and power in the hands of a few, such as corporations.²³⁰

Thus, to state the obvious, technopolarity also invites us to rethink the paradigm of the international community fundamentally starring or by states that gave rise to human rights protection systems, and along with them the current paradigms that only reinforce state obligations.

4.1.3 The liar's dividend

Another aspect that has become evident is the use of these tools as an argument to claim AI-generated disinformation and avoid accountability for scandals or political media coverage. This implies that politicians or other public figures within the disinformation and distrust ecosystem are exploiting deepfakes for their own benefit to avoid giving explanations. This phenomenon, considered an indirect effect of disinformation, is known as '*the liar's dividend*'.²³¹

In other words, this phenomenon involves claiming that real information is false, erroneous and misleading, in this case generated by AI.

This case of 'disinformation about disinformation' was exemplified by the alleged deepfake in Paraguay, where a presidential candidate had called for an end to the protests in front of the electoral authority, after having disseminated narratives of fraud that had sparked those demonstrations. His running mate claimed that the video had been manipulated by AI. Later, police authorities refuted this claim and stated that the video had been personally recorded by the former presidential candidate.

4.1.4 Informed consent

Various regulations in Latin America refer to the presence of informed consent as an element to enable sanctions or criminal repression.

²³⁰ United Nations Secretary-General, High-Level Advisory Body on Artificial Intelligence (n 224).

²³¹ KJ Schiff, DS Schiff and NS Bueno, 'The Liar's Dividend: Can Politicians Claim Misinformation to Evade Accountability?' (2025) 119(1) American Political Science Review 71; Robert Chesney and Danielle K Citron, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security' (2019) 107(6) California Law Review 1753.

In this sense, informed consent clearly implies part of access to a person's information, which, in light of the topics addressed in these pages, means that the person whose biometric data, image or voice, are being manipulated should know in detail what the deepfake will be used for, on what media it will be disseminated, what it will be made to say or do, and how. In other words, complete and detailed information must be provided to avoid violating any of the rights of the person in question.

Informed consent would be useful in cases of deepfakes with a cultural, educational or awareness-raising purpose. That is, in cases where the deepfake does not serve harmful purposes against third-party candidates or the democratic system itself. This element becomes critical when, despite the refusal, an ultra-deep manipulation (deepfake) is still carried out, especially in the digital ecosystem, since there would be difficulties in establishing mechanisms for its effective control and compliance given the massive number of users.

This last element was highlighted in the case of Uruguay, in which, faced with the refusal of one of the presidential candidates to appear on a television set, the host of the programme decided to create a deepfake of the candidate and expose him to a debate in prime time on Uruguayan television.

4.4.5 Deepfakes and gender

The factual analysis in Chapter 3 shows that deepfakes have been used as a weapon in electoral contexts ('weaponisation'). But from a gender perspective, it takes on particular nuances. This last point indicates that when the target of this weapon are women, even more if they are those who work in the public sphere, it takes on a particular form of execution: During election campaigns, this technology has been used without consent to impersonate them in order to undermine their campaigns by attacking them or spreading false information about their private lives.

Of the cases discussed in the previous chapter, the deepfakes targeting Myriam Bregman (presidential candidate in Argentina) and Kamala Harris (presidential candidate in the United States) exemplify this point. On the one hand, Myriam Bregman is depicted as hypersexualised and being sexually assaulted by one of her political opponents, Javier Milei, during the Argentine elections. On the other hand, Kamala Harris was subjected to intense

scrutiny regarding the beginnings of her political career, which resulted in two deepfakes portraying her in a hypersexualised manner. The accompanying descriptions alluded to supposed ‘libertine beginnings’ with the aim of undermining her credibility in politics.

In both cases, deepfakes are used not to manipulate or falsify an aspect of reality, but to create and tell a story that never happened. The distinguishing feature in these cases lies in the gender stereotypes, norms and existing power structures that are reproduced and materialised in deepfakes, where women are objectified, some physical aspect is highlighted and they are hypersexualised, even reaching the point of pornographic deepfakes or AI-manipulated images where forms of abuse are perpetrated.

It is worth noting that although two cases are presented under this particular type of attack by means of deepfake, it should be mentioned that this is partly due to the limitations of the object of analysis, that is, the restriction on presidential candidacies, which implies recognising that in general terms the countries analysed had an average of only one to two presidential candidates.

That said, it doesn’t mean that the link between deepfakes and gender is non-existent or scarce; rather, there are studies that compile cases of this type of AI manipulation used against female politicians, both candidates and those holding positions of varying hierarchy below the presidency, such as mayoress, congresswomen and gubernatorial candidates.

For example, in Mexico, a candidate for the state of Morelos was targeted by a deepfake in which she threatened to cut social programmes if citizens did not vote for her.²³² Similarly, during municipal elections in Paraguay, a deepfake was circulated in which a candidate announced her defeat. In Brazil, during the re-election campaign of the mayor of Bauru, a deepfake was circulated showing her naked. In the same country, federal deputy Tabata Amaral was the victim of pornographic deepfakes during her electoral campaign.²³³

²³² Derechos Digitales and others, ‘Submission to the UN Special Rapporteur on Freedom of Opinion and Expression: Thematic report on “Freedom of Expression and Elections in the Digital Age”’ (2025) 1 <https://www.tedic.org/wp-content/uploads/2025/02/Submission_Foe-and-Elections_Derechos-Digitales-Karisma-TEDIC-IPANDETEC.pdf> accessed 17 January 2026. In the same sense see Claudia May del Pozo and Daniela Rojas Arroyo, ‘IA Generativa y su influencia en las elecciones de México 2024’ (Friedrich Naumann Foundation 2024).

²³³ Derechos Digitales and others (n 233) 2.

In the United States, approximately 35,000 cases have been identified on pornographic websites featuring 26 members of the US Congress, 25 of whom were women. According to this study, female members of Congress are 70 times more likely to appear as victims of non-consensual, AI-generated intimate images than men.²³⁴

This phenomenon has been observed in other countries, such as Spain, where deepfakes of this kind have been found to target female leaders of political forces or symbols of equality policies.²³⁵ Similarly, in Pakistan, the Minister of Information for Punjab, Asma Bukhari, was also a victim of this type of AI-driven digital violence aimed at damaging her credibility.²³⁶

It is also worth noting that if the concept of women politicians is analysed from an extensive perspective, this problem includes journalists, leaders, candidates, representatives, militants or activists.

Particularly in the case of female politicians considered in this work – the presidential candidates – the boundaries between disinformation, politics and pornography, the three main areas in which image manipulation occurs become blurred.²³⁷ This manipulation must be understood within a power struggle in which the ideology of hate, misogyny and false narratives converge.²³⁸

All of this is an extension of the digital violence women suffer online, which intensifies when they are exposed in public life and through gender misinformation. The most common forms that these deepfakes take through hypersexualisation and the development of non-consensual pornography are: face replacement, changing one person's face for another and the replication of facial expressions, which often lend realism to the copy in audio-visual pieces.²³⁹ Decontextualisation is also a contributing factor.

²³⁴ E Rosalie Li, Benjamin Shultz and Nina Jankowicz, 'Deepfake Pornography Goes to Washington: Measuring the Prevalence of AI-Generated Non-Consensual Intimate Imagery Targeting Congress' (American Sunlight Project 2024) <<https://static1.squarespace.com/static/6612cbdf9a9ce56ef931004/t/67586997eae5c6ae3b-b5e24/1733847451191/ASP+DFP+Report.pdf>> accessed 17 January 2025.

²³⁵ Almudena Barrientos-Báez, María Teresa Piñero Otero and Denis Porto Renó, 'Imágenes falsas, efectos reales. Deepfakes como manifestaciones de la violencia política de género' (2024) 82 Revista Latina de Comunicación Social 1, 20.

²³⁶ Saba Chaudhary, 'Women's Politicians Bearing the Brunt of Deepfake' (*Media Diversity Institute*, 17 February 2024) <<https://www.media-diversity.org/womens-politicians-bearing-the-brunt-of-deepfake/>> accessed 17 January 2026.

²³⁷ Barrientos-Báez, Piñero Otero and Porto Renó (n 236) 20.

²³⁸ *ibid* 13.

²³⁹ *ibid* 18.

The goal of this type of AI manipulation is to gradually silence and discredit women, damage their public image and credibility, humiliate them and affect how they are perceived. This perpetuates misogyny and maintains existing power structures.

Notwithstanding the foregoing, given the complexity and scope involved in explaining this topic, it is suggested that it be addressed in subsequent studies.

4.2 What is the reaction to the American deepfakecracy? Attempts at national and international regulations

Based on the cases analysed, the emergence of deepfakes in electoral contexts as a disinformation tool in the American continent is observed almost simultaneously with the rest of the world.

Furthermore, in the analysed time period, the fieldwork reveals evidence of an increase in the proliferation of this type of AI production (particularly if we take into consideration the electoral case of Ecuador in 2023 and 2025).

For this reason, it is reasonable to question whether the countries of America have reacted legislatively to this reconfiguration of electoral rules and whether the human rights approach has been present in these reactions.

The questions are not trivial, since most of the literature consulted on the subject only highlights the pre-eminence of three distinct systems or trends in AI regulation²⁴⁰ that seem to be vying for control over the direction of future codifications.²⁴¹ These trends are led by the United States, the European Union and China.

In this way, the studies focus on the legal perception of only one of the countries that make up the American continent, excluding the legal consequences that this phenomenon entails for Latin America.

²⁴⁰ Gergely Ferenc Lendvai and Gergely Gosztönyi, 'Deepfake y Desinformación- ¿Que puede hacer el derecho frente a las noticias falsas creadas por deepfake?' (2024) 41 IDP 1.

²⁴¹ Matias Gonzalez Mama, Agustín Perez Aledda and Lina Palacios, 'Global AI Governance: Who regulates, with what approach, and for whom?' (Research Article No 67 Center for Studies in Freedom of Expression [CELE] 2025).

Another perspective that gets lost in the whirlwind caused by AI and the race among states to establish a regulatory framework for it is that of human rights. The latter often takes a back seat.²⁴²

For this reason, this paper aims to briefly analyse the national regulatory response specifically regarding electoral deepfakes in the Americas. Following this, it will map international human rights instruments that serve as a legal framework obligating states to adopt measures to mitigate this problem, or whether, on the contrary, the human rights paradigm is insufficient to address the incursion of AI-manipulated content into electoral contexts.

The three regulatory trends

As mentioned, the main regulatory trends are those of the United States, the European Union and China. This coincides with an important point to highlight: these are the three regions where the development of AI systems and design is overwhelmingly concentrated.²⁴³ Indeed, Stanford University's AI Index Report recognised 109 machine learning models as the most influential in the AI ecosystem, of which 61 were created in the United States, 25 in Western Europe and 15 in China.²⁴⁴

This gap means that most policies and practices on AI have been developed in and for the Global North,²⁴⁵ which clearly implies that this disparity is expressed at the level of global AI governance.

Specifically, Lendvai and Gosztanyi focus their analysis on these three approaches applied to deepfake regulation.²⁴⁶ Thus, the United States is characterised by fragmented regulations and a liberal approach, that is, it is based on a *laissez-faire* of deepfakes. The regulations aim for a framework that strongly promotes innovation, demands low levels of transparency and resorts to measures *ex post*, through litigation and state laws.

²⁴² Gonzalez Mama, Perez Aledda and Palacios (n 241) 1.

²⁴³ Gordon LaForge, Robert Muggah and Gabriella Seiler, 'Bridging the AI Governance Divide' (T20 Policy Brief Task Force 05 2024) <https://www.t20brasil.org/media/documentos/arquivos/TF05_ST_05_Bridging_the_AI_gov66cdcbf06f991.pdf> accessed 17 January 2026.

²⁴⁴ Nestor Maslej and others, *The AI Index 2024 Annual Report* (AI Steering Committee, Institute for Human-Centered AI, Stanford University 2024) ch 3; *ibid*.

²⁴⁵ Gonzalez Mama, Perez Aledda and Palacios (n 241).

²⁴⁶ Lendvai and Gergely Gosztanyi (n 240).

On this last point, it is worth mentioning that by the end of July 2024, 151 projects related to deepfakes and deceptive media in the electoral context had been registered.²⁴⁷ The state regulatory landscape is diverse; some absolutely prohibit the use of deepfakes and deceptive media in electoral campaigns, while others allow it under certain conditions. In some states, such as Minnesota and Texas, the use of deepfakes in the 90 and 30 days prior to elections, respectively, to influence them is considered a crime. At the federal level, there is only one law that prohibits the non-consensual distribution of sexually explicit images, whether real or AI-generated (Take it Down Act).²⁴⁸

For its part, the European Union establishes comprehensive regulations through the AI Act, approved in May 2024 by the European Council, emphasises the evaluation of AI technology under a risk-based system (unacceptable, high risk, limited and minimal) and pays close attention to transparency and clear guidelines on disinformation. Particularly regarding deepfakes, it does not define them precisely, but refers to ‘ultra-impersonations’ and establishes the obligation to publicly indicate their artificial origin.²⁴⁹

This is linked to recital 62, which considers the level of risk of the AI system according to the context and establishes that ‘those AI systems intended to be used to influence the outcome of an election or a referendum, or the electoral behavior of natural persons in the exercise of their vote in elections or referendums’ must be classified as high risk systems, and are therefore subject to the provisions established by the AI Act in this regard.²⁵⁰

Finally, China’s national regulation, enacted in 2022, is characterised by its strict nature, requiring deepfake service providers to establish guidelines and procedures for identifying false or harmful information. The objectives focus on data security, content management and transparency. The approach is based on

²⁴⁷ Lawrence Norden, Niyati Narang and Laura J Protzmann, ‘States take the lead in regulating AI in elections-within limits’ (*Center for Justice*, 7 August 2024) <<https://www.brennancenter.org/our-work/research-reports/states-take-lead-regulating-ai-elections-within-limits>> accessed 17 January 2026.

²⁴⁸ Take it Down Act 2025 (US) <<https://www.congress.gov/bill/119th-congress/senate-bill/146>> accessed 16 June 2026.

²⁴⁹ Artificial Intelligence Act. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 [2024] OJ L 2024/1689

²⁵⁰ *ibid*, recital 62.

controlling these technologies: when deepfake technology is detected as being used to create, replicate, publish or disseminate false information, service providers must take measures to mitigate it and maintain records of such incidents.

So far, we have outlined the main candidates in algorithmic governance to date, based on a considerable portion of the consulted literature. But what about most of the Americas, specifically Latin America? Is it simply a recipient of the regulatory modeling of other regions? This last point is understood as ‘one of the mechanisms through which the laws of a country or region migrate to other latitudes. It is a relatively common global phenomenon that allows us to predict certain dynamics, such as the prevalence of core countries or the advantages of those who act first’.²⁵¹ Alternatively, has it attempted to regulate deepfakes under its own specific circumstances?

4.2.1 Latin American national regulations

The following briefly highlights Latin American national regulations on deepfakes in electoral contexts.

Argentina

Since 2018, there have been attempts to regulate robotics and AI.²⁵² However, regarding the regulation of deepfakes, the first regulations were introduced in the province of Mendoza and Ciudad Autonoma de Buenos Aires, both in 2024. The Mendoza bill establishes a ban on the use of deepfakes in electoral processes, the creation of a registry of digital service providers that manipulate audiovisual content, and also ambiguously proposes fines, temporary or permanent disqualifications from participating in electoral processes, ‘and any other measure deemed necessary’ for those who fail to comply with its provisions. Meanwhile, the bill introduced in the Buenos Aires City legislature mandates the labelling of AI-manipulated content and sets out a series of guiding principles, including the protection of fundamental rights, without further details.

²⁵¹ Matias Gonzalez Mama and Ramiro Álvarez Ugarte, ‘Regulatory Modeling. Tracking the Influence of the Digital Services Act in Latin America’ (Research Paper No 63, Centro de Estudios en Libertad de Expresión (CELE) 2025).

²⁵² Favio Farinella, ‘Regulación de la Inteligencia Artificial en Argentina’ (Sistema Argentino de Información Jurídica 2024) <<https://www.saij.gob.ar/DACF240120>> accessed 17 January 2026.

At the national level, the presentation of projects on the criminal regulation of deepfakes with pornographic connotations has also been a major focus of the legislative scene since 2024.

A more general bill, introduced by national deputy Juan Fernando Brügge in July 2024, proposes penalising, through the addition of article 139 to Title IV on Crimes Against Civil Status (Criminal Code), the unauthorised use, generation or modification of AI-generated images, videos and audio recordings of a person with the intention of making them appear real, or that cause harm or ‘potential or actual damage’. It also proposes penalising those who create pornographic deepfakes and those who conceal or edit the identification of the artificial origin of such production. The bill establishes a harsher penalty when the victims are minors.

A key distinguishing feature is the introduction of consent as a defining element, the absence of which justifies the sanctioning of the conduct. Furthermore, it makes exceptions for content concerning matters of public interest or parody related to public officials or individuals in the public eye, thus aligning with the protection of speech according to current freedom of expression standards.

However, it wasn’t until 2025, after deepfakes of Mauricio Macri and Silvia Lospennato circulated announcing their withdrawal from the race during the electoral blackout period for the Buenos Aires City legislative elections, that a regulatory initiative to modify the National Electoral Code was introduced. The proposal establishes limits, procedures and mandatory labelling for social media and news outlets. Specifically, the bill aims to require digital platforms to implement technical mechanisms to detect and remove misleading content within 48 hours, and also mandates the provision of tools for reporting false information.

Similarly, on 26 June 2025, a bill was introduced to amend the National Electoral Code to establish as grounds for disqualification from candidacy those individuals involved in electoral manipulation using AI. This bill extends the prohibition not only to authors, co-authors or accomplices in the creation of deepfakes but also to those who knowingly participated in the dissemination, reproduction, financing or viral spread of such content. This regulation, which aims to be comprehensive, could face

limitations, as the origin of deepfakes is often difficult to trace due to repeated sharing, and it also overlooks the fact that the identity of those contributing to the viral spread of this type of content is often unknown.

Uruguay

In the Republic of Uruguay, there is no specific regulation. However, a bill has been introduced that aims to specifically regulate the creation or dissemination of misleading content through images, sounds or videos, with the purpose of causing reputational damage to a candidate or spreading significant misinformation related to the electoral campaign.

The crime is considered complete if the act described in the law is committed within the period of three months before and one month after the date of the national elections. The law establishes a penalty of six months to two years imprisonment.

The second paragraph reads: 'It will not constitute criminal conduct when it is evident that the content is a satire or parody, or when it is clearly identified that the content has been artificially designed or when other actions are verified in order to avoid its confusion with reality'.

Finally, it exempts from responsibility for content generated by third parties or users, the media, administrators of internet sites, digital platforms or similar, when notified of the falsity, the existence of doubts about its authenticity is admitted or where appropriate, it is removed immediately.

In this regard, it is worth noting the evident use of the state's punitive power to try to deter the generation of disinformation through deepfakes. However, the legal definition of the crime presents a series of inconsistencies in light of the principle of legality, according to which the criminal offense must be specific, precise and concrete. This last point clearly involves the terms used.

Thus, the bill does not define what will be understood by 'generating and disseminating' 'notorious disinformation'. This raises questions about when such a phenomenon is notorious and when it is not, and what makes it notorious (likes, shares, exposure on how many platforms or media outlets).

It also refers to the ‘generation of images, sounds, or videos’, meaning that it does not limit the criminal classification solely to those generated by AI, but uses a more ambiguous form of wording, which would allow to include other scenarios besides deep-fakes, for example memes, cheapfakes and shallowfakes. The continuation of its second paragraph is imprecise.

The foregoing must be analysed taking into account what the Inter-American Court of Human Rights has repeatedly held regarding the drafting of criminal offences: they must adhere to strict and unambiguous terms that define punishable conduct, such that there is a clear definition of the criminalised conduct, establishing its elements and allowing it to be distinguished from non-punishable behaviours or illicit conduct sanctioned with non-criminal measures.²⁵³ Similarly:

ambiguity in the formulation of criminal offenses generates doubts and opens the field to the discretion of the authorities, which is particularly undesirable when it comes to establishing the criminal responsibility of individuals and punishing it with penalties that severely affect fundamental rights, such as life or liberty.²⁵⁴

On the other hand, it does not refer to the type of responsibility that a candidate has during those elections when that person is the one who spreads the deepfake. Do a user and a candidate for public office have the same type of responsibility?

It does not talk about prevention as education, nor does it talk about assistance to victims, it does not set a deadline for removing the content, nor does it talk about consent.

²⁵³ María Luisa Piqué, Artículo 9. Principio de Legalidad y de Retroactividad’ in *Convención Americana de Derechos Humanos y su proyección en el Derecho Argentino* (1st edn, La Ley 2013) 169.

²⁵⁴ Inter-American Court of Human Rights. *Case of Kimel v Argentina* Judgment of 2 May 2008. Merits, Reparations and Costs. Series C No 177 para 63.

Paraguay

Currently, there is no specific legislation on deepfakes. As a parliamentary initiative, a bill was introduced on 12 June 2025, concerning the ‘classification and punishment of the creation, dissemination, and use of deepfakes and manipulated content’.²⁵⁵ It is currently under treatment.

This project, unlike the Uruguayan one, is not specifically limited to electoral deepfakes, but proposes a broader and more comprehensive scope of application, both in terms of the subject matter covered and the territorial scope of the cases it encompasses. In this regard, the bill states in its explanatory memorandum that one of the uses of deepfakes is to ‘Manipulate public opinion: Especially during election periods, by spreading serious misinformation, delegitimizing candidacies or sowing social chaos’.

This project uses the criminal justice system to try to discourage the dissemination of this type of content, while at the same time not only repressing deepfakes, but also other types of manipulated content that do not ‘reach the level of sophistication of a deepfake’.

Furthermore, they highlight the intention to harm as elements of the criminal offence, which places it under the scope of intentional crimes and introduces the concept of informed consent (previously defined in the text of the project).

The project refers to electoral deepfakes as one of the alleged malicious uses when ‘manipulation public opinion in electoral processes, referendums or any other democratic process, affecting the freedom of the vote or the transparency of the process’ (article 6 point 6 of the Bill).

It also establishes that the dissemination of deepfakes by public officials in the exercise of their duties or by abusing them constitutes an aggravating circumstance (article 7 point 5). Furthermore, it is worth mentioning that the bill includes preventative regulations focused on digital literacy and inter-institutional cooperation with the technology industry. At the same time, it emphasises strengthening international cooperation in the fight against deepfakes through the signing of multilateral agreements and mutual assistance in international investigations.

²⁵⁵ Honorable Chamber of Deputies of Paraguay Bill that defines and penalizes the creation, dissemination and use of deepfakes and manipulated content (12 June 2025) Exp No D2585277 <<https://silpy.congreso.gov.py/web/descarga/expediente-178212?preview>> accessed 6 July 2025.

Chile

Currently, there is no specific law governing deepfakes. Like other states, bills have been introduced in 2024 with the intention of regulating AI and its applications. One of them²⁵⁶ addresses the establishment of limits on AI to protect fundamental rights and recognises privacy and transparency as human rights (articles 3 and 4 of the draft). At the same time, it establishes the responsibility of developers, operators and controllers of AI systems for any damages or harm they cause. In this sense, the draft aims for a general regulation of AI systems.

On the other hand, although it is not an attempt at specific regulation on electoral deepfakes, on 17 June 2025 the Chilean Senate presented a bill²⁵⁷ where it establishes the obligation of labelling by all AI content generators that use personal data of any kind for the creation or modification of AI content for dissemination purposes.

This requires compliance with a series of characteristics and imposes penalties for non-compliance or alterations of the seal or metadata that identify the AI content. These penalties are those provided for the crime of computer forgery (article 5 of Law No 21,459) and it is also subject to civil liabilities arising from the malicious or fraudulent use of the AI content.

These projects are mentioned because of their connection to electoral deepfakes; however, Chile has not introduced any bills on this matter, only those related to pornographic deepfakes. While it maintains a human rights perspective, it doesn't address the impact on, for example, the transparency of the democratic process or system, though it does mention public trust.

²⁵⁶ Chilean Chamber of Deputies Bill for the Protection of Human Beings from Artificial Intelligence (3 September 2024) <<https://www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmID=17732&prmBOLETIN=17112-19>> accessed 17 January 2026.

²⁵⁷ Chamber of Deputies of Chile Bill, initiated by motion of the Honorable Senators Bianchi, Castro González, Castro Prieto, Kusanovic and Walker, Regulating the Obligation to Establish a Clear and Traceable Seal for Content Generated with Artificial Intelligence, in the Terms Indicated <<https://www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmID=18265&prmBOLETIN=17618-19>> accessed 17 January 2026.

Brazil

Brazil, for its part, has addressed the use of electoral deepfakes within the regulatory framework of electoral propaganda. Its incorporation into the legal system was achieved through Resolution No 23,732 issued by the Superior Electoral Court.

Article 9 of said resolution prohibits the use, in electoral propaganda, of fabricated or manipulated content to disseminate false or decontextualised information with the potential to harm the integrity of the electoral process. Similarly, it prohibits the use of deepfakes to harm or favour a candidate, even with authorisation (consent) to create, replace or alter the image or voice of a living, deceased or fictitious person. Failure to comply with these provisions constitutes abuse of political power and misuse of the media, subject to the revocation of a candidate's electoral registration.

Campaigns can only use AI tools on election-related content in specific situations such as improving image quality or producing graphic elements of visual identity, cartoons and logos.

Regarding the role of platforms, the resolution establishes that those who allow the paid promotion of political content must provide an easily accessible repository of all published ads, detailing the ad broadcast period, the number of people reached and the segmentation criteria defined by the advertiser.²⁵⁸ Furthermore, they must take immediate measures to cease promoting, monetising and providing access to content that could affect the integrity of the electoral process, regardless of whether or not its circulation is notified.

An important aspect to highlight is that it introduces joint civil and administrative liability for application providers when they do not remove content and accounts during the electoral period in cases of risk such as: a) anti-democratic information and acts, b) notoriously false or seriously decontextualised facts, c) acts of incitement to violence against members of the electoral justice system and d) hate speech and dissemination of AI-generated content that does not respect the labelling forms established in the resolution.

²⁵⁸ Beatriz Farrugia, 'Regulating the use of AI for Brazilian elections: what's at stake' (DFRLab, 29 May 2024) <<https://dfriab.org/2024/05/29/regulating-the-use-of-ai-for-brazilian-elections-whats-at-stake/>> accessed 17 January 2026.

While the regulations represent progress in the specific regulation of AI in the electoral context, they use some terms that are not defined in the resolution, such as ‘disinformation content’ or ‘decontextualised fact’, which could lead to difficulties in their application.²⁵⁹ Furthermore, a limitation of the regulation in question concerns the entities subject to it: it does not stipulate policies on generative AI content created by users not affiliated with parties or candidates; therefore, it also does not stipulate whether such cases are subject to monitoring or sanctions.

Peru

Recently, through Law No 32314²⁶⁰ on 29 April 2025, the Peruvian state incorporated modifications into the Criminal Code and the Cybercrime Law, regarding the use of AI in the commission of crimes. The recent legislation imposes aggravated penalties for those who use technologies such as deepfakes for illicit purposes.

The law emphasises the production of sexually explicit deepfakes involving minors and the use of deepfakes to commit fraud. It also introduces an aggravated penalty for cybercrimes committed using AI or similar technologies.

Thus, this new regulation in the normative framework does not specifically address the issue of electoral deepfakes, nor does it refer to specific responsibilities of public officials, candidates or social platforms regarding the production or dissemination of deepfakes in electoral contexts.

Colombia

Regarding the problem specifically related to deepfakes, Colombia has promoted a bill that aims to add an aggravating circumstance to the crime of identity theft.²⁶¹ When identity theft is carried out using AI, the issue is again addressed in a general way. This means that legislative efforts do not directly target the use of deepfakes in electoral contexts, but rather propose a more

²⁵⁹ Farrugia (n 258).

²⁶⁰ Congress of the Republic of Peru Law No 32314 Law That Amends the Criminal Code, Legislative Decree 635, and Law 30096, Law on Computer Crimes, to Include the Use of Artificial Intelligence in the Commission of Crimes (29 April 2025) <<https://www.normaslegalesonline.pe/clp/contenidos.dll/Legislacion/1807201/1809258/1809858/1809860>> accessed 17 January 2026.

²⁶¹ Congress of the Republic of Colombia Bill No 225 of 2024 Senate whereby an aggravating circumstance is modified and established in Article 296 of Law 599 of 2000, Colombian Penal Code (2024) <<https://forogpp.com/wp-content/uploads/2024/02/pl-225-24-falsedad-personas-ia.pdf>> accessed 17 January 2026.

ambiguous range of regulations under the umbrella of criminal law. Furthermore, it does not propose a comprehensive plan to combat this type of problem, nor does it address all the actors involved in the creation and dissemination of electoral deepfakes.

The project's rationale acknowledges that one purpose of the malicious use of deepfakes is the transmission of disinformation, and that one of the legal rights that is being protected, in addition to personal integrity, is also public trust.

Mexico

Meanwhile, in Mexico, a bill has been identified for implementation in the Federal District (Mexico City) in which identity theft is repressed through AI technologies from the perspective of criminal regulation.

In the rest of Latin America, no specific legislative initiatives have been identified (as in Bolivia, Venezuela, Nicaragua, Honduras and the Dominican Republic), or those that exist focus on the regulation and promotion of AI in general, but not on the use of deepfakes as disinformation tools in electoral contexts (Ecuador, Panama, Costa Rica and Guatemala). This search was conducted considering bills submitted for consideration and approved on the official websites of the legislative bodies.

The analysis concludes that in Latin America, despite having experienced the use of deepfakes as disinformation tools in recent presidential elections almost simultaneously with other regions of the world, national regulations are still in their infancy, as they involve legislative initiatives. In other words, there are no specific regulations in force regarding electoral deepfakes.

These initiatives show a certain tendency to resort to regulation subject to the criminal sphere, where prison sentences are involved, and in a few other cases, actions are discouraged by the imposition of fines.

In this regard, it should be noted that, as mentioned in Chapter 2, AI is constantly evolving, and as such requires a broad regulatory definition, or at least one that encompasses potential changes. Under the legislative logic of a criminal law, this would violate the principle of legality, as it does not provide a concrete, precise and detailed description of the criminal offence.

However, this work has shown that in the ecosystem of disinformation, ergo of electoral deepfakes, various actors are involved: users, public officials, political parties, traditional media and social media platforms.

In this respect, few cases offer a comprehensive approach to the problem involving all stakeholders. Furthermore, the American trend focuses on sanctions *ex-post*, but not in a comprehensive approach also considering prevention such as digital education or the obligation of states to implement public detection measures developed by themselves or in conjunction with other technology companies.

Nor has it been found that the regulations respond to local contexts and needs.

4.2.2 International regulations

International regulatory initiatives are characterised by a lack of universal standards and fragmented developments, leading to inconsistencies in AI governance. The international landscape is characterised by non-binding (soft law) and general initiatives, namely: the OECD Recommendation on Artificial Intelligence (2019), which establishes industry-oriented principles based on human-centred values;²⁶² the G20 Principles on AI, inspired by the aforementioned instrument (2019);²⁶³ and the G7 Hiroshima Process (2023), a set of principles and codes of conduct for organisations developing advanced generative AI systems.²⁶⁴

Under the UN framework, the first steps toward understanding AI from a human rights perspective had already been taken in 2018 with the UN General Assembly Resolution A/73/348 on the promotion and protection of freedom of opinion and expression.²⁶⁵ However, a significant milestone occurred in 2021 with the

²⁶² Gonzalez Mama, Perez Aledda and Palacios (n 241).

²⁶³ G20, 'AI Principles' (Annex to the G20 Ministerial Statement on Trade and Digital Economy, June 2019) accessed 16 June 2026

²⁶⁴ G7, 'Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems' (30 October 2023) <https://g7g20-documents.org/database/document/2023-g7-japan-leaders-leaders-annex-hiroshima-process-international-guiding-principles-for-organizations-developing-advanced-ai-system> accessed 16 June 2026.

²⁶⁵ United Nations General Assembly (UNGA), 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression' (2018) UN Doc A/73/348 <<https://docs.un.org/en/a/73/348>> accessed 17 January 2026.

UNESCO Recommendation on the Ethics of AI.²⁶⁶ This recommendation addresses issues such as the need to create risk assessment and mitigation procedures that incorporate a proportionality analysis of AI use according to context. It also addresses issues such as environmental impact and the inclusion of a cross-cutting gender perspective. The recommendation proposes that states should use AI to improve access to information and knowledge and that they must also guarantee freedom of expression and access to information in relation to the automatic generation, moderation and preservation of content (preventing the dissemination of disinformation) and the transparency of operators.²⁶⁷

Furthermore, in 2023, an AI Advisory Body was established under the auspices of the General Secretariat, which issued an interim report that same year ('Governing AI for Humanity') acknowledging the significant absence of the Global South in discussions on AI governance. The final report was published in September 2024.

In the latter, it acknowledges not only the benefits but also the risks associated with AI, including the increased creation and spread of disinformation, damage to the integrity of information, and the impact on freedom of expression through information bubbles, etc. It is particularly interesting to consider that experts understand that a key social impact of AI revolves around disinformation and trust.

Ultra-forgeries, voice clones, and automated disinformation campaigns pose a specific and serious threat to institutions and democratic processes such as elections ... as well as for democratic societies and social trust in general, among other things through the manipulation and interference of foreign information. The development of closed-loop information ecosystems, reinforced by AI and utilizing personal data, can have profound effects on societies, potentially making them more prone to accepting intolerance and violence towards others.

²⁶⁶ United Nations Educational, Scientific and Cultural Organization (UNESCO), 'Recommendation on the Ethics of Artificial Intelligence' (UNESCO 2021) <<https://unesdoc.unesco.org/ark:/48223/pf0000381137>> accessed 15 January 2026.

²⁶⁷ Gonzalez Mama, Perez Aleda and Palacios (n 241) 11.

Protect the integrity of representative government institutions and processes requires robust systems for verifying and detecting ultra-counterfeiting, as well as rapid procedures for detecting and removing content that may deceive and cause harm or social division or that incite war propaganda, conflicts and hate speech People who are not public figures must be protected against the creation of hyper-forgeries of their image for fraudulent, defamatory, or abusive purposes. Sexualized hyper-forgeries are especially concerning for women and girls and can constitute a form of gender-based violence.²⁶⁸

Also on 21 March 2024, the UN General Assembly adopted Resolution 78/265 ('Seizing the opportunities of a Safe AI and reliable for sustainable development'). The most remarkable aspect of this is that it was adopted by all UN member states.²⁶⁹ It emphasises that AI is an evolving field and that governance must address the needs of both developed and developing countries. It stipulates that states must refrain from or cease using AI systems that are impossible to operate without protecting human rights. The report highlights AI solidarity and its criteria, such as transparency, digital literacy to combat disinformation and equity, among others.

Finally, the Global Digital Compact, adopted in 2024 as an annex to the Future Pact, should be added to this list. It calls on companies and developers to adopt safeguards in the training processes of AI models, identify AI-generated material and certify the authenticity of the content and its origin, as well as using labels, watermarks and other techniques. It specifically refers to the obligation to mitigate potential harms, such as hate speech and discrimination in AI-generated content'.²⁷⁰ It recognises the urgency of addressing and countering misinformation and disinformation while also raising the need to ensure information integrity, that is, that it is reliable, relevant and accurate.²⁷¹

²⁶⁸ United Nations Secretary-General, High-Level Advisory Body on Artificial Intelligence (n 224) 40 (emphasis added).

²⁶⁹ Gonzalez Mama, Perez Aleda and Palacios (n 241).

²⁷⁰ *ibid* 20.

²⁷¹ United Nations General Assembly, 'Global Digital Compact' (2024) UN Doc A/79/L.2.

Meanwhile, under the Organization of American States, there is an initiative of several member states: the Inter-American Framework for Data Governance and Artificial Intelligence (MIG-DIA). As such, it includes cross-cutting guidelines on respect for human rights, multi-stakeholder dialogue, citizen participation, environmental sustainability and updating regulatory frameworks on privacy and data protection. It also emphasises the importance of promoting the alignment of national policies with international standards and fostering regional dialogue to consolidate national positions in international forums.

Finally, although the Working Group on Business and Human Rights has identified that states are gradually migrating to binding instruments, it should be mentioned that a fragmented context of soft law still prevails, against which the European Union Artificial Intelligence Regulation, mentioned in the previous section, stands as a general binding legal instrument on AI.

This demonstrates the efforts to outline AI governance in all the aforementioned instruments. In some of them, human rights appear as a guiding principle, but in a secondary role. Logically, they are addressed in a general way in the instruments belonging to the universal protection system.

This, in light of the topic studied in this thesis, allows us to ask:

4.3 Where do human rights stand in the face of disinformation by deepfakes in electoral contexts within the Inter-American human rights system?

This study has demonstrated that deepfakes have appeared in the American presidential elections as a disinformation tool, generally using the image of candidates with the intention of deceiving the electorate. This seeks to highlight a potential harm that goes beyond the honour of a candidate or political party, threatening the erosion of the democratic fabric through a direct attack on electoral integrity, the preservation of which is an obligation for states under the American Convention on Human Rights (ACHR).²⁷²

²⁷² Inter-American Court of Human Rights *Capriles v Venezuela* Judgment of 10 October 2024. Preliminary Objection, Merits, Reparations and Costs, paras 106–107.

In turn, the proliferation of electoral deepfakes can become a multi-faceted tool when they circulate gender stereotypes (such as the cases of Myriam Bregman and Kamla Harris), class stereotypes (such as the Bullrich deepfake), are used as a means to commit other crimes (such as the Claudia Sheinbaum deepfake) or violate the right to fair competition by subjecting a presidential candidate through an AI representation to a televised debate without their consent (Yamandú Orsi deepfake), and even serve to circulate highly conflictive narratives for society (the cases of Rómulo Roux or Luis Abinader).

This makes it clear that the current electoral landscape in the Americas, shaped by AI and social media, presents a paradigm shift from the traditional relationship between the state, political parties and citizens. Now, platforms that, together with technological innovations like generative AI, are creating new dynamics and rules of engagement must be considered part of this new model.

Therefore, it must be questioned whether the current human rights paradigm is suited to this new scenario. In this regard, it should be noted that while the thesis's central issue can be approached from various thematic perspectives, to respect the length limit, the legal analysis will attempt to be as concise as possible, in accordance with the standards of the Inter-American system of human rights, considering that most of the countries studied here are members.

Thus, it is worth noting that, given the fragmented and mostly soft law regulations, there is no specific international instrument applicable to the Americas region that regulates deepfakes, much less in electoral contexts. In this regard, it should not be forgotten that AI, as a scientific advancement, is protected by the right to science under various instruments of the Inter-American system, particularly article 26 of the ACHR.²⁷³

The Inter-American Court of Human Rights has also referred to the protection afforded by the International Covenant on Economic, Social and Cultural Rights (ICESCR),²⁷⁴ whose monitoring body, the Committee on Economic, Social and Cultural Rights, in its General Comment No 25, establishes that the right to science includes the right to participate in scientific progress and enjoy

²⁷³ Inter-American Court of Human Rights Advisory Opinion 32/25 of 29 May 2025 'Climate Emergency', 170, para 471.

²⁷⁴ International Covenant on Economic, Social and Cultural Rights (adopted 16 December 1966, entered into force 3 January 1976) 993 UNTS 3 (ICESCR).

its benefits, without discrimination.²⁷⁵ Likewise, this commentary reiterates basic principles such as non-discrimination in the enjoyment of the benefits of science, the consideration of vulnerable groups in light of these advances and recognises that in certain contexts science and its applications may affect economic, social and cultural rights, and therefore may be subject to limitations provided they are determined by law, promote ‘the general welfare in a democratic society’, and such restriction is compatible with the nature of the right being restricted.

In light of this, it should be noted that electoral deepfakes (studied here) as a disinformation tool are produced with the aim of deceiving the electorate, which directly undermines access to information, a necessary condition for people to be able to freely exercise their decisions in electoral contexts.

In this regard, it is worth noting that the Inter-American Court of Human Rights, in the contentious case of *Claude Reyes and others v Chile*, has recognised for the first time the right of access to information held by the state as a human right protected by international law.²⁷⁶ The case confirmed that the right to access public information was enshrined in article 13 of the ACHR, relating to freedom of expression, and considered that from the two dimensions, individual and social, of freedom of expression, a series of rights protected by said article were derived, among which was the right and freedom to seek, receive and impart information and ideas of all kinds:

²⁷⁵ Inter-American Court of Human Rights Advisory Opinion 32/25 of 29 May 2025 ‘Climate Emergency’, 170, para 472.

²⁷⁶ Inter-American Court of Human Rights *Claude Reyes v Chile (Funds, Reparations and Costs)* Judgment of 19 September 2006, Series C No 151.

77. With regard to the facts of the present case, the Court considers that Article 13 of the Convention, by expressly stipulating the rights to 'seek' already ... The right to 'receive' information protects the right of every person to request access to information under the control of the State, with the exceptions permitted under the Convention's restrictions. Consequently, this article protects the right of individuals to receive such information and the State's positive obligation to supply it so that the person can have access to that information or receive a reasoned response when, for any reason permitted by the Convention, the State may limit access to it for the specific case.²⁷⁷

It also considers that:

In this regard, it is important to highlight that there is a regional consensus among the States that make up the Organization of American States (hereinafter 'the OAS') *on the importance of access to public information and the need for its protection*. This right has been the subject of specific resolutions issued by the General Assembly of the OAS. In its latest Resolution of June 3, 2006, the OAS General Assembly 'urged States to respect and enforce access to public information *to all people* and [a] promote the adoption of legislative or other provisions that were necessary to ensure their recognition and effective application.'²⁷⁸

In this way, the Inter-American Court of Human Rights highlights the importance and presence of this right in various documents of the Inter-American system. Furthermore, since it derives from freedom of expression, access to information is also a fundamental component of the exercise of democracy, in accordance with article 4 of the Inter-American Democratic Charter:²⁷⁹

79. The Inter-American Democratic Charter highlights in its article 4 the importance of 'transparency of government activities, probity, the responsibility of governments in public management, respect for social rights and freedom of expression and of the press' as fundamental components of the exercise of democracy.

Furthermore, the close relationship between democracy and freedom of expression is highlighted, as well as the need for the community to be sufficiently informed to exercise its choices. The court reiterates what was stated in the *Advisory Opinion 5/85 on 'Mandatory Registration of Journalists'* and in cases like *Ricardo Canese v Paraguay* and *Herrera Ulloa v Costa Rica*:

²⁷⁸ *Claude Reyes v Chile* (n 276) para X (emphasis added).

²⁷⁹ *ibid* para 79.

freedom of expression is a fundamental element upon which the existence of a democratic society is based. It is indispensable for the formation of public opinion. It is also a condition *without which no* political parties, trade unions, scientific and cultural societies, and in general, all those who wish to influence the community, can fully develop. It is, in short, a condition for the community ... *When exercising your options, be sufficiently informed.* Therefore, it is possible to state that a society that is not well informed is not fully free.²⁸⁰

This implies that access to (public) information is not only a right in itself, but also has an instrumental character. Subsequently, the Inter-American Court of Human Rights has addressed the content of this right and its relationship with other prerogatives from a public perspective, that is, as information held by the state.

For example, in a case prior to the express recognition of access to public information in *Claude Reyes*, in *Palamara Iribarne v Chile* establishes that the duty to keep confidential ‘does not cover information relating to the institution or the functions it performs when it has been made public’.²⁸¹

On the other hand, in *Gomes Lund and Others v Brazil*, the court specified the requirement of acting in good faith and diligently when access to information is linked to the right to the truth.²⁸² In the judgment *Case of the inhabitants of La Oroya v Peru*, in addition to developing corporate responsibility regarding human rights, the court addresses access to information and political participation.²⁸³ While it does so in relation to the environment, it also refers to elements such as the principles of publicity and transparency in public administration so that people can exercise democratic control over state actions and access to information without needing to demonstrate a direct interest or personal impact, thus strengthening the standard it has put forward.

²⁸⁰ *Claude Reyes v Chile* (n 276) para X.

²⁸¹ Inter-American Court of Human Rights *Case of Palamara Iribarne v Chile* Judgment of 22 November 2005. Merits, Reparations and Costs. Series C No 135 para 77.

²⁸² Inter-American Court of Human Rights *Case of Gomes Lund and others* (*‘Guerrilha do Araguaia’*) *v Brazil* Judgment of 24 November 2010. Preliminary Objections, Merits, Reparations and Costs. Series C No 219 para 211.

²⁸³ Inter-American Court of Human Rights *Case of inhabitants of La Oroya v Perú* (Preliminary Objections, Merits, Reparations and Costs) Judgment of 27 November 2023, Series C No 511.

Furthermore, it elaborates on the obligation of active transparency that implies the state's duty to provide information, on its own initiative, that is necessary and relevant for individuals to exercise other rights, especially regarding the right to life, personal integrity and health. In particular, it maintains that access to environmental information must be affordable, effective and timely.

On the other hand, in *Civil Association Active Memory v Argentina* the Inter-American Court refers to the right of access to information regarding state intelligence activities.²⁸⁴ In this regard, it reiterates the OAS regional consensus on the importance of this right. The court also states that in cases of human rights violations, state authorities cannot invoke mechanisms such as state secrecy or confidentiality of information, or reasons of public interest or national security, to withhold information requested by the judicial or administrative authorities in charge of the investigation. Specifically, the court refers to access to classified information and its proper archiving and preservation as essential elements for guaranteeing the right to the truth.

However, it should be emphasised, due to its connection with the thesis topic, that in *Capriles v Venezuela*,²⁸⁵ the Inter-American Court of Human Rights, in line with what it held in the Advisory Opinion No 28/21 on indefinite presidential re-election,²⁸⁶ It establishes that the democratic principle inspires and guides the interpretation and application of the ACHR, particularly the exercise of the political rights under article 23 of that instrument. In this regard, it establishes that the state must guarantee, through positive measures, that every person formally entitled to political rights has the *real opportunity* in order to exercise them, that is, the state must provide the conditions and mechanisms for them to be exercised effectively.

²⁸⁴ Inter-American Court of Human Rights *Case of Asociación Memoria Activa v Argentina* Judgment of 26 January 2024. Merits, Reparations and Costs.

²⁸⁵ Inter-American Court of Human Rights *Case of Capriles v Venezuela* Judgment of 10 October 2024. Preliminary Objection, Merits, Reparations and Costs.

²⁸⁶ Inter-American Court of Human Rights. *La figura de la reelección presidencial indefinida en el contexto del Sistema Interamericano de Derechos Humanos* (Opinión Consultiva OC-28/21) Corte IDH Serie C No 43 (7 de junio de 2021).

Following this argument, the court considers that article 13 of the ACHR (which, it should be remembered, protects access to public information) reinforces the state's obligations to achieve fairness in the electoral contest, with respect to access to the media and election propaganda and simultaneously, introduces the *OAS Guide to Good Practices in Electoral Matters*, with a view to deepening the idea of equity in access to the media.²⁸⁷

It also recognises the obligation of states to guarantee the integrity of the electoral process in light of articles 23, 24 and 13 of the ACHR. This means that, at a minimum, transparency must be guaranteed throughout the electoral process, along with opportunities for citizens to access information about electoral campaigns.

In this regard, and taking note of various international instruments and bodies that have developed the content of political rights, the Court considers that the obligation to preserve electoral integrity requires States to guarantee, at a minimum, and in accordance with domestic law, the following:

- A — transparency throughout the electoral process, particularly in the financing of campaigns and in the results counting phase, as well as the participation of witnesses, prosecutors and/or observers belonging to political parties and/or civil society, and the presence of independent national and international observers;
- B — opportunities for those competing for public office to publicize their proposals through traditional and digital media, and for citizens to have access to information about election campaigns.²⁸⁸

Furthermore, it is worth mentioning that the court finds a close link between the population's access to information in the electoral context and the right to access information:

²⁸⁷ Inter-American Court of Human Rights. *La figura de la reelección presidencial indefinida en el contexto del Sistema Interamericano de Derechos Humanos* (n 286).

²⁸⁸ Inter-American Court of Human Rights *Case of Capriles v Venezuela* Judgment of 10 October 2024. Preliminary Objection, Merits, Reparations and Costs para 107

The Court recalls that the obligation to guarantee the public's access to information within the framework of an electoral process is closely related to the right of access to information, which is protected by Article 13 of the American Convention. Therefore, this Court considers that the coverage by public media must adhere to a principle of equity.²⁸⁹

However, from everything stated regarding access to information, it follows from the Inter-American interpretive framework that: 1) the Inter-American Court of Human Rights recognises access to information held by the state as a human right and 2) although in the *Capriles* case the court expressly acknowledges the current communication context, considering social media platforms and networks, it refers exclusively to those of an institutional nature, that is, public ones.

However, in its Advisory Opinion on Climate Emergency (OC-32/25), the Inter-American Court of Human Rights reinforces the interpretation presented thus far, maintaining that 'Article 13 of the American Convention on Human Rights not only protects the expression of ideas, but also the fundamental right of society to receive truthful information from diverse sources, which is essential for informed and democratic decision-making'. It further adds that States must take appropriate measures to guarantee the integrity of public information.²⁹⁰

Also important is the consideration of paragraph 528 in which the Inter-American Court of Human Rights mentions:²⁹¹

Considering the risks arising from disinformation and other forms of information manipulation on climate matters, the Court calls on civil society, the media and other actors in the information field to assume an active role in the generation and dissemination of reliable content on climate change.

This suggests that while the developments and recognitions outlined in this area are important, in the field of political and digital rights, current times demand an understanding that there is a 'profound change in information and content consumption

²⁸⁹ *Case of Capriles v Venezuela* (n 288) para 125.

²⁹⁰ Inter-American Court of Human Rights Advisory Opinion 32/25 (n 273) para 526.

²⁹¹ *ibid* para 528.

habits, driven by technological changes occurring both on the network itself, due to its constant increase in speed, and in the type of “devices” people use to access information’.²⁹² This is even more relevant when considered in light of a context in which disinformation has emerged strongly in recent years during elections.²⁹³

This phenomenon, namely disinformation, has evolved over time, especially with the expansion of internet use and the emergence of large intermediary actors who concentrate a large part of the flow of information that circulates there.²⁹⁴ In this last group, platforms belonging to private companies are protagonists of the contemporary technopolarity.

It only remains to add that the panorama is furthered by the use of AI, which reinforces what is called an ‘epistemic crisis’, understood as the inability to establish common ground for addressing public debate and to determine with certainty what is true and what is false, as a consequence of the “breakdown” of trust in democratic institutions of the press and politics .²⁹⁵

It should be noted that the UN General Assembly, in its Resolution A/77/287, has already stated that ‘States have a crucial role to play in combating the effects of disinformation, both in terms of their own actions and their duty to protect against human rights abuses by third parties, including commercial enterprises’.²⁹⁶

For its part, the Inter-American Court of Human Rights, in its recent Advisory Opinion on Climate Emergency, has recognised the duty of states to adopt progressive measures to counter climate disinformation.²⁹⁷

Therefore, it must be considered that disinformation in electoral contexts, especially when it uses AI-generated tools such as deepfakes, has implications for other human rights, such as access to information and the exercise of political rights enshrined in the ACHR. This must be articulated with a criterion of dynamic interpretation of the ACHR and interpretative interaction, under

²⁹² Inter-American Commission on Human Rights Office of the Special Rapporteur for Freedom of Expression, ‘Disinformation, pandemic, and human rights’ (2023) (OEA/Ser.L/V/II), 18.

²⁹³ *ibid.*

²⁹⁴ *ibid.* 17.

²⁹⁵ *ibid.* 37.

²⁹⁶ United Nations General Assembly, ‘Report of the Secretary-General. Countering disinformation to promote and protect human rights and fundamental freedoms’ (2022) UN Doc A/77/287, para X.

²⁹⁷ Inter-American Court of Human Rights Advisory Opinion 32/25 (n 273) para 526.

which the convention is understood as a living instrument adapted to the evolution of the times and, in particular, to current living conditions, with a view to establishing broad protection of the human person.²⁹⁸

4.4 Conclusion of Chapter 4

In short, this paper has attempted to expose, on the one hand, the characteristics of American deepfakecracia in the recent presidential elections of 2023–2024. Furthermore, the review of national and international legislative attempts reveals the legislative fragmentation that exists on this matter, while the human rights approach appears as a secondary consideration.

Finally, in light of inter-American standards, the existence of a new electoral paradigm is raised, which reconfigures the traditional relationship between the state, political parties and citizens, as well as public debate, has been transformed. Today, digital platforms and intermediaries represent key players in this reconfiguration. The information ecosystem has changed, along with the rules of the electoral game and the ways in which political and public discourse is conducted. New AI tools, such as deepfakes, are used not only to disseminate messages but also to employ tactics to discredit political opponents and democratic institutions.

For this reason, it is argued that a new inter-American paradigm of human rights is necessary, where standards recognise the new habits and actors of the digital information system or at least reinforce and specify their obligations regarding compliance with human rights.

²⁹⁸ Josefina Comune and Natalia Luterstein, 'Artículo 29. Normas de Interpretación' in *Convención Americana de Derechos Humanos y su proyección en el Derecho Argentino* (1st edn, La Ley 2013).

5. Conclusion

Chapter 1 of this work outlines the main characteristics of the current information ecosystem, as well as the ways in which disinformation operates in digital spaces. Furthermore, the first part concludes with a description of the multitude of actors involved in this disinformation phenomenon, whose methods of interaction evolve in tandem with technological advancements. The chapter's objective is to understand how the phenomenon of disinformation intersects with deepfakes in electoral contexts.

Chapter 2 then presented what AI is and how we arrived at the concept of deepfakes. It emphasises that this is a scientific innovation and, as such, aligns with the changes discussed in Chapter 1.

Chapter 3 fulfils the second specific objective of the thesis by compiling non-representative cases of deepfakes that used the image or identity of presidential candidates in countries across the Americas that held elections for the highest executive office. It then explores the narratives employed and their implications in local contexts. This chapter makes it clear that deepfakes have arrived as tools of disinformation and are here to stay.

Finally, the last chapter addresses what is meant by electoral deepfakes for the purposes of this work, the definition of which is given by the aspects evaluated in the cases of the previous chapter, as well as presenting in a concrete way some particularities observed in the aforementioned case studies.

On the other hand, a gap is identified regarding the development and regulation of electoral deepfakes between the United States and the rest of Latin America, despite the latter having experienced the presence of electoral deepfakes almost simultaneously. In this regard, the legislative impact of electoral deepfakes in Latin American countries is analysed within a context where these countries are vying to impose a model of AI governance.

However, this thesis asks: What is the human rights impact of electoral deepfakes, particularly within the Inter-American system? It is considered that disinformation disseminated through deepfakes during elections can be addressed from multiple perspectives and international standards, which would exceed the scope of this work. For this reason, priority is given to the standards invoked by the Inter-American Court of Human Rights regarding the right of individuals to access information.

From the latter, it can be concluded that the court has strengthened the standard on access to information held by the state, that is, from a public approach or perspective, which is not sufficient for the new electoral paradigm posed by platforms and generative AI.

In this sense, the *Capriles v Venezuela* case and the recent Climate Emergency Advisory Opinion imply a step towards this redefinition of the electoral landscape, insofar as the current communication context is acknowledged, as well as morigera. The right to access information is limited to that held by the state, and an appeal is made to other actors in the flow of information circulation in light of disinformation (albeit of a climatic and environmental nature).

Bibliography

Books and Articles

- Alanazi S and Asif S, 'Explorando la tecnología deepfake: creación, consecuencias y contra-medias' (2024) 6 Human-Intelligent Systems Integration 49 <<https://doi.org/10.1007/s42454-024-00054-8>>
- Anguita P and others, 'El fenómeno de la desinformación: Revisión de experiencias internacionales y en Chile' (Comité Asesor contra la Desinformación; Ministerio de Ciencia, Tecnología, Conocimiento e Innovación 2023)
- Anka R, 'Chapter 3: Responsible AI' in Maslej N and others, The AI Index 2024 Annual Report (AI Steering Committee, Institute for Human-Centered AI, Stanford University 2024)
- Association for Civil Rights (ADC) and others, 'Disinformation on the Internet in electoral contexts of Latin America and the Caribbean' (2019)
- Avolio M, Qué es la Inteligencia Artificial y cómo impacta en nuestras vidas . ed Grijalbo 2024
- Barrientos-Báez A, Piñero Otero MT and Porto Renó D, 'Imágenes falsas, efectos reales. Deep-fakes como manifestaciones de la violencia política de género' (2024) 82 Revista Latina de Comunicación Social 1
- Caiafa CF and Lew SE, '¿Qué es la Inteligencia Artificial?' (CONICET 2020)
- Calvo E and Aruguete N, Fake News, Trolls and Other Charms (ed Siglo XXI 2020)
- Chataut R and Upadhyay A, 'Introduction to Deepfake Technology and its Early Foundations' in Deepfakes and Their Impact on Business (IGI Global Scientific Publishing 2025)
- Chesney R and Citron DK, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security' (2019) 107(6) California Law Review 1753
- Comune J and Luterstein N, 'Artículo 29. Normas de Interpretación' in Convención Americana de Derechos Humanos y su proyección en el Derecho Argentino (1st edn, La Ley 2013)
- Corporación Latinobarómetro, 'Latinobarómetro 2024: La Democracia Resiliente' (2024)
- del Pozo CM and Rojas Arroyo D, 'IA Generativa y su influencia en las elecciones de México 2024' (Friedrich Naumann Foundation 2024)
- della Porta D, 'Comparative Analysis: Case-Oriented Versus Variable-Oriented Research' in Approaches and Methodologies in the Social Sciences: A Pluralist Perspective (CUP 2008)
- Farrugia B, 'Regulating the use of AI for Brazilian elections: what's at stake' (DFRLab, 29 May 2024) <<https://dfrlab.org/2024/05/29/regulating-the-use-of-ai-for-brazilian-elections-whats-at-stake/>> accessed 17 January 2026
- Feinmann JP, Filosofía Política del poder mediático (1st edn, Planeta 2013)
- Gómez MJ and others, 'Diez preguntas frecuentes y urgentes sobre Inteligencia Artificial' (Fundación Sadosky and Program. AR 2024) <<https://program.ar/wp-content/uploads/2024/08/Diez-preguntas-frecuentes-y-urgentes-sobre-Inteligencia-Artificial.pdf>> accessed 17 January 2026

- Gonzalez Mama M and Álvarez Ugarte R, 'Regulatory Modeling. Tracking the Influence of the Digital Services Act in Latin America' (Research Paper No 63, Centro de Estudios en Libertad de Expresión (CELE) 2025)
- Gonzalez Mama M, Perez Aleda A and Palacios L, 'Global AI Governance: Who regulates, with what approach and for whom?' (Research Article No 67 Center for Studies in Freedom of Expression [CELE] 2025)
- González-Bailon S and De Domenico M, 'Bots are less central than verified accounts during contentious political events' (2018) 118(11) *Proceedings of the National Academy of Sciences* e2013443118 <[www.pnas.org/doi/epub/10.1073/pnas.2013443118](https://doi.org/10.1073/pnas.2013443118)> accessed 18 January 2026
- Goodfellow I and others, 'Generative Adversarial Nets' (University of Montreal 2014)
- Graphika CF, 'Actors, Behaviors, Content: A Disinformation ABC. Highlighting Three Vectors of Viral Deception to Guide Industry and Regulatory Responses' (Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression 2019)
- Kirschbaum I and Martínez F, 'Desinformación en la Argentina: una perspectiva multisectorial' (Centro de Estudios en Tecnología y Sociedad (CETyS) UdeSA 2024) <<https://dspaceapi.live.udea.edu.ar/server/api/core/bitstreams/e6b7b460-6004-4191-b9c6-7de88f8c13ca/content>> accessed 18 January 2026
- Kitchin R and McArdle G, 'What makes Big Data, Big Data? Exploring the Ontological Characteristics of 26 Datasets' (2016) 3(1) *Big Data & Society* 1 <<https://doi.org/10.1177/2053951716631130>> accessed 17 January 2026
- Kleis Nielsen R and Ganter SA, The Power of Platforms: Shaping Media and Society (OUP 2022) <<https://doi.org/10.1093/oso/9780190908850.001.0001>>
- Labuz M and Nehring C, 'On the Way to Deep Fake Democracy? Deepfakes in Election Campaigns in 2023' (2024) 23 *European Political Science* 454
- Lendvai GF and Gosztönyi G, 'Deepfake y Desinformación- ¿Que puede hacer el derecho frente a las noticias falsas creadas por deepfake?' (2024) 41 IDP 1
- López López PC, Mila Maldonado A and Ribeiro V, 'La desinformación en las democracias de América Latina y de la península ibérica: de las redes sociales a la inteligencia artificial (2015-2022)' (2023) *Uru: revista de comunicación y cultura* 69
- López Sotelo J, *Deep Learning. Theory and Applications* (1st edn, Alpha 2021)
- Maslej N and others, 'The AI Index 2024 Annual Report (AI Steering Committee, Institute for Human-Centered AI, Stanford University 2024)
- Nielsen R.K and Ganter S.A. 'The Power of Platforms: Shaping Media and Society' (2022) *Oxford Studies in Digital Politics* -University Press. <<https://doi.org/10.1093/oso/9780190908850.001.0001>> Accessed 18 January, 2026
- Ortiz Dominguez M, 'Artificial neural networks' (2025) 12(23) *Ingenio y Conciencia Boletín Científico de la Escuela Superior Ciudad Sahagún* 38
- P. Anguita, I. Bachmann, L. Brossi, C. Elórtégui, M.J. Escobar, P. Ibarra, J.C., F. Padilla y P. Peña . "El fenómeno de la desinformación: Revisión de experiencias internacionales y en Chile" (2023) *Comité Asesor contra la Desinformación*; Ministerio de Ciencia, Tecnología, Conocimiento e Innovación- P 3P.
- Pawelec M, 'Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions' (2022) 1 *Digital Society* 1 <<https://doi.org/10.1007/s44206-022-00010-6>> accessed 17 January 2026
- Piqué ML, 'Artículo 9. Principio de Legalidad y de Retroactividad' in *Convención Americana de Derechos Humanos y su proyección en el Derecho Argentino* (1st edn, La Ley 2013)
- Ramos-Zaga F, 'Deepfake: Analysis of its technological and legal implications in the era of Artificial Intelligence' (2024) IX (27) *Global Law. Studies on Law and Justice* 359 <www.scielo.org.mx/pdf/dgedj/v9n27/2448-5136-dgedj-9-27-359.pdf>
- Rubio Nuñez R, García Rossó R and Sánchez V, 'Desinformación: como entenderla, combatirla y protegerse de sus efectos' (Friedrich Naumann Foundation for Freedom, Regional Office for Latin America 2024) <www.freiheit.org/sites/default/files/2024-12/03dic_papier_desinformacion_versioncorta02.pdf> accessed 16 January 2026

Russell SJ and Norvig P, *Inteligencia Artificial. Un Enfoque Moderno* (2nd edn, Pearson Educación SA 2004)

Schiff KJ, Schiff DS and Bueno NS, 'The Liar's Dividend: Can Politicians Claim Misinformation to Evade Accountability?' (2025) 119 (1) *American Political Science Review* 71

Sierra Caballero F and Sola Morales S, 'Golpes mediáticos y desinformación en la era digital. La guerra irregular en América Latina' (2020) 17 *Community and Society* Guadalajara e7604 <www.scielo.org.mx/scielo.php?script=sci-arttext&pid=S0188-252X2020000100121&lng=es&nrm=iso> accessed 17 January 2026

Silmovich A, 'Desinformación política e Inteligencia artificial en la campaña presidencial argentina de 2023' in Oskar Almazá-Lopez, Begoña Cabanes-Cacho and Eva Blumbury (coords), *IA, educación y medios de comunicación: modelo TRIC* (2024). Dyckinson ed.

Suarez AA, 'Politización de la desinformación en contextos de información devaluada. El caso de Latinoamérica' (2022) 4(17) *Revista Internacional de Comunicación y Desarrollo* <<https://doi.org/10.15304/ried.4.17.8793>> accessed 17 January 2026

Tönberg P and Chueri J, 'When Do Parties Lie? Misinformation and Radical-Right Populism Across 26 Countries' (2025) 31(2) *The International Journal of Press/Politics* 367 <<https://doi.org/10.1177/19401612241311886>> accessed 17 January 2026

Twomey J and others, 'Do Deepfake Videos Undermine Our Epistemic Trust? A Thematic Analysis of Tweets That Discuss Deepfakes in the Russian Invasion of Ukraine' (2023) 18(10) *PloS one* 17 <<https://doi.org/10.1371/journal.pone.0291668>> accessed 17 June 2026

United Nations Educational, Scientific and Cultural Organization (UNESCO), 'Recommendation on the Ethics of Artificial Intelligence' (UNESCO 2021) <<https://unesdoc.unesco.org/ark:/48223/pf0000381137>> accessed 15 January 2026

Villar MP, 'El Rol de la Inteligencia Artificial Generativa (IAG) en la Desinformación Electoral' (Friedrich Naumann Foundation 2024)

Vovkodav V, Sendino Marcos L and Cordero Redondo A, 'Fake news y su impacto en la sociedad: una revisión de la literatura' (2022) 13(1) *Pangea Revista de la Red Académica Iberoamericana de Comunicación* 45 <<https://revistapangea.org/index.php/revista/article/view/220>> accessed 18 January 2026

Waisbord S, 'Truth is What Happens to News' (2018) 19(13) *Journalism Studies* 1866 <https://is.muni.cz/el/fss/podzim2019/ZURb1608/um/readings/week_10/Truth_is_What_Happens_to_News.pdf> accessed 18 January 2026

Wardle C and Derakshan H, 'Information Disorder: Toward an interdisciplinary framework for research and policy making' (Council of Europe Report DGI(2017)09 2017) <<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>> accessed 18 January 2026

Official documents

Documents of the Inter-American Commission on Human Rights (IACHR)

Special Rapporteurship for Freedom of Expression of the Inter-American Commission on Human Rights, 'Report: The Right of Access to Information' (2009)

— with contributions from the Department of Electoral Cooperation and Observation and the Department of International Law of the General Secretariat of the Organization of American States, 'Guide to guarantee freedom of expression in the face of deliberate disinformation in electoral contexts' (OEA/Ser.D/XV.22 2019)

Special Rapporteurship for Freedom of Expression of the Inter-American Commission on Human Rights, 'Disinformation, pandemic, and human rights' (2023) (OEA/Ser.L/V/II)

— 'Annual Report of the Office of the Special Rapporteur for Freedom of Expression' (OEA/Ser L/V/II, Doc 39 2024) 32 <www.oas.org/en/iachr/expression/reports/IA2024RELEEN.pdf> accessed 18 January 2026

United Nations

International Covenant on Economic, Social and Cultural Rights (adopted 16 December 1966, entered into force 3 January 1976) 993 UNTS 3 (ICESCR)

United Nations General Assembly, 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression' (2018) UN Doc A/73/348 <<https://docs.un.org/en/a/73/348>> accessed 17 January 2026

United Nations General Assembly Human Rights Council, 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan. Disinformation and freedom of opinion and expression' (2021) UN Doc A/HRC/47/25

United Nations General Assembly, 'Report of the Secretary-General. Countering disinformation to promote and protect human rights and fundamental freedoms' (2022) UN Doc A/77/287

United Nations Secretary-General, High-Level Advisory Body on Artificial Intelligence, 'Governing AI for Humanity. Final Report' (United Nations 2024) <www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf> accessed 8 July 2025

United Nations General Assembly, 'Global Digital Compact' (2024) UN Doc A/79/L.2

Cases and Advisory Opinions of the Inter-American Court of Human Rights (IACHR)

Inter-American Court of Human Rights *Case of Claude Reyes and others v Chile* Judgment of 19 September 2006. Merits, Reparations and Costs. Series C No 151

Inter-American Court of Human Rights *Case of Gomes Lund and others ('Guerrilha do Araguaia') v Brazil* Judgment of 24 November 2010. Preliminary Objections, Merits, Reparations and Costs. Series C No 219

Inter-American Court of Human Rights *Case of Palamara Iribarne v Chile* Judgment of 22 November 2005. Merits, Reparations and Costs. Series C No 135

Inter-American Court of Human Rights *Case of Expelled Dominicans and Haitians v Dominican Republic* Judgment of 28 August 2014

Inter-American Court of Human Rights *Case of Asociación Memoria Activa v Argentina* Judgment of 26 January 2024. Merits, Reparations and Costs

Inter-American Court of Human Rights *Case of Capriles v Venezuela* Judgment of 10 October 2024. Preliminary Objection, Merits, Reparations and Costs

Inter-American Court of Human Rights Advisory Opinion 32/25 of 29 May 2025 'Climate Emergency'

European Commission

Code of Practice on Disinformation (European Commission 2018)

Online sources

@BancadaADN (*X (formerly Twitter)*, 27 February 2025) <<https://x.com/bancadaadn/status/1895174409351987657?s=51>> accessed 4 July 2025

@dhernandezlarez (*X (formerly Twitter)*, 22 June 2024) <<https://x.com/dhernandezlarez/status/1804526869993103837>> accessed 3 July 2025

@diegowainstein (*X (formerly Twitter)*, 27 July 2023) accessed 29 June 2025

@GordoDa_ (*X (formerly Twitter)*, 7 November 2023) <<https://x.com/GordoDan/status/1722010056160403556>> accessed 29 June 2025

@igalvar71 (*X (formerly Twitter)*, 3 October 2024) <<https://x.com/igalvar71/status/1841878323204403311>> accessed 4 July 2025

A24.com, 'Votó Mauricio Macri: Habló del temporal y sobre la difusión de videos falsos' (*YouTube*, 18 May 2025) <<https://youtu.be/Al06Kd5ud4Q?si=2ilj7FINKTLqnEB>> accessed 18 January 2026

AFP Factual, 'El video de la candidata Luisa González promoviendo el fentanilo en Ecuador es un montaje' (*AFP Factual*, 23 October 2023) <<https://factual.afp.com/doc.afp.com.33Y768Y>> accessed 30 June 2025

Barojan D, “#TrollTracker: Bots, Botnets, and Trolls” (*DFRLab*, 9 October 2018) <<https://medium.com/dfrlab/trolltracker-bots-botnets-and-trolls-31d2bdbf4c13>> accessed 17 January 2026

Bond S, ‘The FCC says AI voices in robocalls are illegal’ (*NPR*, 8 February 2024) <www.npr.org/2024/02/08/1230052884/the-fcc-says-ai-voices-in-robocalls-are-illegal> accessed 29 June 2025

— ‘A political consultant faces charges and fines for Biden deepfake robocalls’ (*NPR*, 23 May 2024) <www.npr.org/2024/05/23/nx-s1-4977582/fcc-ai-deepfake-robocall-biden-new-hampshire-political-operative> accessed 29 June 2025

Calderón C, ‘Mexico Elections 2024: Deepfakes and fake news win in the elections’ (*El Financiero*, 4 June 2024) <www.elfinanciero.com.mx/elecciones-mexico-2024/2024/06/04/deep-fakes-y-fake-news-marcaron-el-escenario-electoral/> accessed 3 July 2025

Canal 4, ‘Santo y Seña - Programa del Domingo 13/10/2024’ (YouTube, 15 October 2024) <https://youtu.be/PP1wmV3CqMw?si=2H-dyoiDNO9w_iMc> accessed 4 July 2025

Cazadores de Fake News, ‘Operaciones de Influencia en Honduras: Lo que debes saber sobre el dataset publicado por Twitter’ (*Cazadores*, 26 August 2022) <www.cazadoresdefakenews.info/honduras-lo-que-debes-saber-operaciones-de-influencia-twitter/> accessed 17 January 2026

Chaudhary S, ‘Women’s Politicians Bearing the Brunt of Deepfake’ (*Media Diversity Institute*, 17 February 2024) <www.media-diversity.org/womens-politicians-bearing-the-brunt-of-deepfake/> accessed 17 January 2026

Channel 22 Mexico, ‘A lie detector capsule about deepfakes involving the President of Mexico’ (*Facebook*, 6 November 2024) <www.facebook.com/Canal22Mexico/videos/c%C3%A1psula-detector-de-mentiras-sobre-los-deepfake-en-que-involucran-a-la-presi-dent/864321199161687/> accessed 3 July 2025

Chequeado, ‘Home page’ <<https://chequeado.com/>> accessed 29 June 2025

Corral Ignacio, ‘No, Lionel Messi no publicó en sus redes sociales una imagen de un león en apoyo a Javier Milei’ (*Chequeado*, 24 October 2023) <<https://chequeado.com/ultimas-noticias/no-lionel-messi-no-publico-en-sus-redes-sociales-una-imagen-de-un-leon-en-apoyo-a-javier-milei/>> accessed 29 June 2025

— ‘No, Lionel Messi no apoyó públicamente a Sergio Massa, como afirman estas placas falsas’ (*Chequeado*, 4 November 2023) <<https://chequeado.com/ultimas-noticias/son-falsas-estas-supuestas-placas-de-tyc-sports-sobre-lionel-messi-y-un-apoyo-a-sergio-massa/>> accessed 29 June 2025

Derechos Digitales and others, ‘Submission to the UN Special Rapporteur on Freedom of Opinion and Expression: Thematic report on “Freedom of Expression and Elections in the Digital Age”’ (2025) 1 <www.tedic.org/wp-content/uploads/2025/02/Submission_Foe-and-Elections_Derechos-Digitales-Karisma-TE-DIC-IPANDETEC.pdf> accessed 17 January 2026

Detect Fakes <<https://detectfakes.kellogg.northwestern.edu>> accessed 17 January 2026

Dilla Alfonso H, ‘República Dominicana: cuando la xenofobia se institucionaliza’ (*Nueva Sociedad*, 2019) <<https://nuso.org/articulo/república-dominicana-cuando-la-xenofobia-se-institucionaliza/>> accessed 2 July 2025

El Nacional, ‘Dudas sobre video de “Payo”: ¿Obligado a grabar o hecho por inteligencia artificial?’ (*El Nacional*, 16 May 2023) <<https://elnacional.com.py/politica/dudas-sobre-video-payo-obligado-grabar-o-hecho-inteligencia-artificial-n47549>> accessed 1 July 2025

Ewe K, ‘The Ultimate Election Year: All the Elections Around the World in 2024’ (*TIME*, 28 December 2023) <<https://time.com/6550920/world-elections-2024/>> accessed 16 January 2026

Farinella F, ‘Regulación de la Inteligencia Artificial en Argentina’ (Sistema Argentino de Información Jurídica 2024) <www.saij.gob.ar/DACF240120> accessed 17 January 2026

Gastón Allen D, ‘Falsos Planes de Expropiación y vínculos guerrilleros, mentiras de la Campaña Guatemalteca’ (*EFE Verifica*, 18 August 2023) <<https://verifica.efe.com/falsos-planes-de-expropiacion-y-vinculos-guerrilleros-mentiras-de-la-campana-guatemalteca/>> accessed 1 July 2025

Hablapanamanoticias Instagram account <www.instagram.com/reel/C31cKims-jOR/?utm_source=ig_web_copy_link> accessed 2 July 2025

La Nación, 'Qué hay detrás de la imagen falsa de Donald Trump con Taylor Swift como protagonista' (*La Nación*, 20 August 2024) <www.lanacion.com.ar/estados-unidos/que-hay-detras-de-la-imagen-falsa-de-donald-trump-con-taylor-swift-como-protagonista-nid20082024/> accessed 29 June 2025

LaForge G, Muggah R and Seiler G, 'Bridging the AI Governance Divide' (T20 Policy Brief Task Force 05 2024) <www.t20brasil.org/media/documentos/arquivos/TF05_ST_05_Bridging_the_AI_gov66cdcbf06f991.pdf> accessed 17 January 2026

Lagos A, 'Claudia Sheinbaum, víctima de deepfake: su imagen fue utilizada en un video fraudulento generado por IA' (*Wired*, 25 January 2024) <<https://es.wired.com/articulos/claudia-sheinbaum-victima-de-deep-fake-su-imagen-fue-utilizada-en-un-video-fraudulento-generado-por-ia>> accessed 3 July 2025

Latin American Center for Investigative Journalism, 'Digital Mercenaries' (2023) available at <www.elclip.org/mercenarios-digitales/> accessed 27 May 2025

Lupa Media, '¿Luisa Gonzalez propuso salir de la dolarización?' (*Lupa Media*, 27 February 2025) <<https://lupa.com.ec/verificaciones/luisa-gonzalez-dolarizacion-adn/>> accessed 4 July 2025

— 'News Reports Altered with Deepfakes: The Disinformation Strategy in the Election Campaign' (*Lupa Media*, 2 April 2025) <<https://lupa.com.ec/explicativos/deepfake-noticieros-desinformacion-campana-electoral-2025/>> accessed 4 July 2025

Lupa Media, 'Disinformation on demand: how the electoral campaign was manipulated in Ecuador' (*Lupa Media*, 9 April 2025) <<https://drive.google.com/file/d/1I4zSU8g7X2DCV-TD-ti138dKGAIE0Z/view>> accessed 4 July 2025

Li ER, Shultz B and Jankowicz N, 'Deepfake Pornography Goes to Washington: Measuring the Prevalence of AI-Generated Non-Consensual Intimate Imagery Targeting Congress' (American Sunlight Project 2024) <<https://static1.squarespace.com/static/6612cbdf-d9a9ce56ef931004/t/67586997eae5c6ae3bb5e24/1733847451191/ASP+DFP+Report.pdf>> accessed 17 January 2025

Misión de Observación Electoral de Guatemala, 'Séptimo Informe. Agosto de 2023' (2023) 9 <<https://dialogos.org.gt/wp-content/uploads/2025/07/Informe-07-MOE-Gt.pdf>> accessed 1 July 2025

NBC News, 'Listen: Fake Biden Robocall Tells Voters To Skip New Hampshire Primary' (*NBC News*, 22 January 2024) <www.nbcnews.com/video/listen-fake-biden-robocall-tells-new-hampshire-not-to-vote-in-primary-202609733664> accessed 29 June 2025

Nicas J and Cholakian Herrera L, 'Argentina, a testing ground for the use of Artificial Intelligence in political campaigns' (*La Nación*, 15 November 2023) <www.lanacion.com.ar/politica/la-argentina-un-campo-de-pruebas-para-el-uso-de-la-inteligencia-artificial-en-las-campanas-politicas-nid15112023/> accessed: 29 June 2025

— 'Is Argentina the First AI Election?' (*The New York Times*, 15 November 2023) <www.nytimes.com/2023/11/15/world/americas/argentina-election-ai-milei-massa.html> accessed 16 January 2026

Norden L, Narang N and Prottmann LJ, 'States take the lead in regulating AI in elections-within limits' (*Center for Justice*, 7 August 2024) <www.brennancenter.org/our-work/research-reports/states-take-lead-regulating-ai-elections-within-limits> accessed 17 January 2026

OAS MOE, 'Informe Preliminar de la Misión de Observación Electoral de la OEA en Paraguay' (*OEA Press*, 2 May 2023) <www.oas.org/fpdb/press/Informe-Preliminar-Paraguay-2023.pdf> accessed 1 July 2025

OAS MOE, 'Informe Preliminar de la Misión de Observación Electoral de la OEA en Ecuador para la segunda vuelta de las elecciones presidenciales' (OEA 2025) 9 <www.oas.org/fpdb/press/2025_MOE_Ecuador_Segunda_Vuelta_Preliminar_ESP.pdf> accessed 4 July 2025

PERFIL, 'From Soviet to Indiana Jones: The use of Artificial Intelligence in Massa's campaign according to The New York Times' (*PERFIL*, 16 November 2023) <www.perfil.com/noticias/politica/uso-inteligencia-artificial-campana-sergio-massa-segun-the-new-york-times.phtml> accessed 29 June 2025

Ramos E, 'Video does not show Sergio Massa inhaling cocaine: it is a "deepfake"' (*Verificador*, 10 November 2023) <<https://larepublica.pe/verificador/2023/11/10/video-no-expone-a-sergio-massa-inhalando-cocaina-es-un-deepfake-474240>> accessed 29 June 2025

Sánchez Caparrós M, 'Prevenir y controlar la discriminación algorítmica' (Rubinzal-Culzoni 2021) <www.justierradelafuego.gov.ar/wp-content/uploads/2021/07/Sánchez-Caparrós.-Prevenir-y-controlar-la-discriminación-algorítmica.pdf> accessed 17 January 2026

Settles G (*Internet Archive*, 23 August 2024) <<https://archive.org/details/screenshot-2024-08-23-at-5.41.29-pm>> accessed 29 June 2025

Spring M, 'Las imágenes falsas creadas con IA para intentar atraer el apoyo de los votantes negros hacia Trump' (*BBC*, 4 March 2024) <www.bbc.com/mundo/articulos/c3g4l5xgvryo> accessed 29 June 2025

Stryker C and Scapicchio M, '¿Qué es la IA generativa?' (*IBM*, 2024) <www.ibm.com/es-es/think/topics/generative-ai> accessed 17 January 2026

UNESCO, 'Inteligencia Artificial y Democracia' (2024) <https://unesdoc.unesco.org/ark:/48223/pf0000389736_spa> accessed 6 July 2025

Cases

Decision 06/26/2025 about Topic 533 and Topic 987 [Supreme Federal Court of Brazil]

Legislation

Chamber of Deputies of Chile Bill, initiated by motion of the Honorable Senators Bianchi, Castro González, Castro Prieto, Kusanovic and Walker, Regulating the Obligation to Establish a Clear and Traceable Seal for Content Generated with Artificial Intelligence, in the Terms Indicated <www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmID=18265&prmBOLETIN=17618-19> accessed 17 January 2026

Chilean Chamber of Deputies Bill for the Protection of Human Beings from Artificial Intelligence (3 September 2024) <www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmID=17732&prmBOLETIN=17112-19> accessed 17 January 2026

Congress of the Republic of Colombia Bill No 225 of 2024 Senate whereby an aggravating circumstance is modified and established in Article 296 of Law 599 of 2000, Colombian Penal Code (2024) <<https://forogpp.com/wp-content/uploads/2024/02/pl-225-24-falsedad-personas-ia.pdf>> accessed 17 January 2026

Congress of the Republic of Peru Law No 32314 Law That Amends the Criminal Code, Legislative Decree 635, and Law 30096, Law on Computer Crimes, to Include the Use of Artificial Intelligence in the Commission of Crimes (29 April 2025) <www.normaslegalesonline.pe/clp/contenidos.dll/Legislacion/1807201/1809258/1809858/1809860> accessed 17 January 2026

Honorable Chamber of Deputies of Paraguay Bill that defines and penalizes the creation, dissemination and use of deepfakes and manipulated content (12 June 2025) Exp No D2585277 <<https://silpy.congreso.gov.py/web/descarga/expediente-178212?preview>> accessed 6 July 2025



Global Campus of Human Rights

Europe

South East Europe

Latin America-Caribbean

Asia-Pacific

Central Asia

Caucasus

Arab World

Africa

Global Campus of Human Rights

The Global Campus of Human Rights is a unique network of more than one hundred participating universities around the world, seeking to advance human rights and democracy through regional and global cooperation for education and research. This global network is promoted through eight Regional Programmes which are based in Venice for Europe, in Sarajevo/Bologna for South East Europe, in Yerevan for the Caucasus, in Pretoria for Africa, in Bangkok for Asia-Pacific, in Buenos Aires for Latin America and the Caribbean, in Beirut for the Arab World, and in Bishkek for Central Asia.

The Global Campus Awarded Theses

Every year each regional master's programmes select the best master thesis of the previous academic year that is published online as part of the GC publications. The selected GC master theses cover a range of different international human rights topics and challenges.

The present thesis - ***Deepfakecracia: Electoral Deepfakes as Disinformation Tools in the 2023–2024 Presidential Elections*** written by **Agostina Dasso Boonman** and supervised by Ivan Schuliaquer, National University of San Martin – was submitted in partial fulfilment of the requirements for the Master's Programme in Human Rights and Democratisation in Latin America and the Caribbean (LATMA-GC Latin America and Caribbean), coordinated by National University of San Martin (Argentina).

GC Headquarters

Monastery of San Nicolò,
Riviera San Nicolò, 26
I-30126 Venice Lido (Italy)



This document has been produced with the financial assistance of the European Union and as part of the Global Campus of Human Rights. The contents of this document are the sole responsibility of the authors and can under no circumstances be regarded as reflecting the position of the European Union or of Global Campus of Human Rights.

www.gchumanrights.org