Oficina de Ciencia y
Tecnología del Congreso
de los Diputados

Report C

Disinformation in the digital age

A complex threat for democracies

· *This document is translated from Spanish, where the distinction between the terms disinformation and misinformation is not applicable.*

CONGRESO DE LOS DIPUTADOS

FECYT INNOVACIÓN

Researchers, scientists and experts consulted (in alphabetical order)

- **Arteaga, Félix.** Lead researcher, Elcano Royal Institute, Spain.
- **Arroyo Guardeño, David.** Senior scientist, 'Leonardo Torres Quevedo' Institute for Physical and Information Technologies (ITEFI), Spanish National Research Council, Spain.
- **Badillo Matos, Ángel.** Lead researcher, Elcano Royal Institute, Spain. Tenured professor, Salamanca University, Spain.
- **Cano Orón, Lorena. PhD,** Assistant professor, University of Valencia, Spain.
- **Cardenal Izquierdo, Ana Sofía.** Tenured professor, Universitat Oberta de Catalunya (UOC), Spain.
- **Carrillo, Nereida.** Associate professor, Autonomous University of Barcelona (UAB), Spain. Co-founder, Learn to Check, media education project association.
- **Corredoira y Alfonso, Loreto.** Tenured professor, Complutense University, Spain. Jean Monnet Chair (2020-2023). Co-director of the Observatorio Complutense de la Información and Dir-Politics Group.
- **Ecker, Ullrich K.H.** Professor, University of Western Australia, Australia.
- **García, David.** Professor, University of Konstanz, Germany. Faculty member, Complexity Science Hub Vienna, Austria.
- **González Bailón, Sandra.** Professor, University of Pennsylvania, United States.
- **Innerarity, Daniel.** Professor of political philosophy, Ikerbasque researcher, University of the Basque Country/Euskal Herriko Unibertsitatea (UPV/EHU). Chair of Artificial Intelligence and Democracy, European University Institute of Florence, Italy.
- **Jiménez Cruz, Clara.** Chair, European Fact-Checking Standard Network and International Fact-Checking Network, Europe. Director, Maldita.es, Spain.
- **Magallón Rosa, Raúl.** Tenured professor, Charles III University, Spain.
- **Majó-Vázquez, Silvia.** Postdoctoral researcher, Reuters Institute, Oxford University, United Kingdom. Assistant professor, Vrije Universiteit Amsterdam, the Netherlands.
- **Rosso, Paolo.** Professor, Valencia Polytechnic University, Spain.
- **Rubio Núñez, Rafael.** Professor, Complutense University, Spain. Co-director, Observatorio Complutense de la Información SN-Disorders.
- **Salaverría, Ramón. Professor,** University of Navarra, Spain. Member, MSI-RES Committee of Experts on **Increasing Resilience of Media, Council of Europe.**
- **Wagner, Astrid.** Senior scientist, Institute of Philosophy, Spanish National Research Council (CSIC). Spain.

TEAM C

Alfonso Cuenca. Clerk to the Spanish Parliament. Director of studies, analyses and publications of the Lower House of the Spanish Parliament.

Ana Elorza. Oficina C Coordinator at the Fundación Española para la Ciencia y la Tecnología.

Izaskun Lacunza*. Oficina C Coordinator at the Fundación Española para la Ciencia y la Tecnología.

Maite Iriondo de Hond. Scientific and Technological Evidence Officer.

Rüdiger Ortiz-Álvarez. Scientific and Technological Evidence Officer.

Sofía Otero. Scientific and Technological Evidence Officer.

Jose L. Roscales*. Scientific and Technological Evidence Officer.

Cristina Fernández-García. Networking and Communication Officer.

Miguel García Suárez. Intern Technological Officer.

Alesandra Puyuelo Estrada. Intern Technological Officer.

*Contacts for this report

Production method

Reports C are brief documents on subjects chosen by the Bureau of the Congress of Deputies that contextualise and summarise the available scientific evidence on the analysed subject. They also inform about areas of agreement, disagreement, unknowns, and ongoing discussions. The preparation process for these reports is based on an exhaustive bibliographical review, complemented with interviews of experts in the field who subsequently conduct two review rounds of the text. Oficina C conducts this process in collaboration with the management team of the Spanish Parliament's Lower House Documentation, Library and Archive service.

To produce this report the Oficina C referenced 492 documents and consulted 18 experts on the subject. Of this multi-disciplinary group, 83% come from the field of social sciences (international relations, political science, communication science, sociology, law), and 17% from physical sciences and engineering (computer science, information science, telecommunications engineering). 56% work in Spanish institutions or centres, whereas 44 % have affiliations with at least one institution outside Spain.

Oficina C is the editorial supervisor of this report.

Summary C

The report in 5 minutes

Relevance

Internet and digital developments have brought many advances, economic and social benefits. They also offer a new social and informational context that has enabled an unprecedented amplification of disinformation and its effects, which are a clear threat for democratic systems. This is an issue of national security that reaches critical levels at times of great social importance, such as during a public health crisis, electoral processes, or armed conflicts. Disinformation may have negative repercussions on public assets such as health, or erode democratic processes, institutions and fundamental rights such as the right to information. Management of this phenomenon is complex because certain rights, such as freedom of speech, might be restricted if the necessary caution and precision are not exercized. There is generalized public concern about this issue and clear symptoms of the public's defencelessness exist.

This report explores the causes and impacts of this phenomenon in depth, as well as the mechanisms that could help combat it.

Disinformation

According to the European Commission, disinformation refers to verifiably false or misleading information that is created, presented and disseminated for economic gain or to deliberately deceive the public, which could cause public harm. It may pursue economic gain, have ideological and electoral objectives, or be motivated by geopolitical interests. Disinformation forms part of the set of actions typical of 'hybrid threats' by means of which third countries attempt to exploit vulnerabilities of the European Union. Apart from international instigators, there are also national instigators such as ideological, religious, economic or other advocacy groups who may cause comparable harm.

In practice, disinformation adopts multiple guises and is not always easy to identify. With far-reaching narratives, many of an international scope, disinformation fabricates messages that replace the truth with verisimilitude, mixing false and veracious content. The messages are flexible, adapted to location and current affairs in order to pervade any subject of social relevance, incidence or confrontation that may arise. To achieve success, the messages do not need to generate false beliefs, it is enough for them to cause confusion and create distrust or amplify bias and prejudice. The aim is to produce structural or profound changes in the public sphere rather than immediate results from any given piece of false news. Instigators use techniques such as affective attraction, simplistic or incomplete views, repetition or the use of artificial intelligence, as occurs in deepfakes, to increase their effectiveness. Among targets of disinformation in Spain, the main issues are politics, electoral processes and certain social challenges like migration.

Focal point

The causes and effects of disinformation are deeply rooted in the new informational context, which is heavily mediated by internet and the array of geopolitical, economic, technological, social and personal factors that modulate our relationship with information both on and offline. This is a multifaceted, multi-factor phenomenon steeped in a series of overlapping and mutually reinforcing dynamics, which means it is not always possible to establish clear relationships of cause and effect. Although experts agree on the risks, the complexity of this phenomenon makes it difficult to comprehensively analyse the impact of false and misleading information.

Social networks, messaging services and large digital platforms have changed the way in which information flows and reaches the public. On one hand, they have given rise to an information explosion of varying quality that makes it difficult to identify truthful content and generates uncertainty. On the other, they are the main source of information that simultaneously blurs the information flow: anyone can create, transmit and share content. In this disintermediation, the classic interpreters of reality, such as mainstream journalism, television or political elites, lose importance. However, they still maintain a decisive role in the amplification of the true or false information that circulates on internet and, therefore, on its impact.

Among many interrelated dimensions, experts indicate the importance of the geopolitical framework in the growing use of what are increasingly effective disinformation operations employed by some countries as tools to destabilize. In terms of systems, they also highlight a reduction of trust in democratic institutions. The loss of social and economic wellbeing or an increase in inequality or dissatisfaction, among other aspects, can become cracks in the fabric of states and their societies, leaving them more vulnerable to false information. Mainstream journalism, with weakened professional structures and wilting public trust are losing their effectiveness to curb the threat. In more social terms, the post-truth context, increasing affective polarization and the circulation of conspiracy theories promotes social fragmentation and reinterpretation of the relationship of society with falsity and veracity: at one extreme we find acceptance of a lie and on the other, negation of objective evidence. On a personal level, these realities converge with multiple factors, such as cognitive bias and certain socio-affective factors that can predispose a All of these elements converge, creating a strong resistance among people to the rectification

Image FOTCIENCIA Fragmentos de memoria ©Gabriel Castilla Cañamero

of their mistaken beliefs or to acceptance of the falsity of information close to their heart.

Finally, the technological context is defined by a digital business model, which is a hurdle to the neutrality and plurality of the information users receive. This model seeks to gain a user's attention and monetize it by means of advertising, taking advantage of the latest developments in artificial intelligence and big data analytics that underpin the platforms. Foremost among their techniques are algorithmic screening systems and targeted or personalized advertising and propaganda. These are tools that have an enormous capacity to amplify the impact of false and misleading information. Of particular concern is the lack of algorithmic transparency, which limits our understanding of their role. A final technological challenge is the lack of knowledge about the flow of false information on private messaging networks or about substantial advances in the use of artificial intelligence, for instance, using bots or deep fakes to spread false contents.

On the horizon

Experts highlight that dealing with disinformation requires both coordination and a combination of many instruments and measures to mitigate its short-term effects, in addition to structured strategies enabling us to fight it in the long term. Experts call on all of the agents involved, from institutions and actors in the spheres of information and politics as well as large digital platforms and online businesses, to exercise responsibility and cooperate in order to avoid an exploitation of uncertainty and false o misleading information, adopting the checks and measures to do so. The general goal is resilience, in addition to the digital and media literacy of society as a whole. To achieve this aim, public policies can take a wide range of steps, including regulatory measures, to reinforce the role of both the main actors involved and the public themselves.

Democratic institutions are facing the structural challenge of fostering a dialogue with the public that reinforces trust and adapts to the new informational context. The role of journalists should also be strengthened, promoting their capacities and resources, independence, transparency and plurality as a measure to mitigate disinformation.

Fact-checking agencies also have an important, positive role to play in society when it comes to monitoring and refuting false information, a role that can also be supported by other actors.

Public resilience to false and misleading information can be reinforced with media literacy plans. There are many proposals in this area and psychology is progressing in the development of effective mechanisms to neutralize false information at the level of individuals. In social terms, another suggestion is to promote an ethical framework guiding the behaviour of people, or any other agent involved, towards the rejection of disinformation, promoting a redesign of the architectures of social networks and digital platforms to make the flow of false information more difficult.

The regulatory framework and policies of the EU promote measures aimed at defending and strengthening democracy, and at consolidating mechanisms that fight disinformation in a systemic way. These measures include attributing responsibilities to large digital platforms, demonetizing content, extending media plurality and freedom, and moderating online electoral content. The recently enacted EU Digital Services Act specifies responsibilities concerning disinformation in the digital ecosystem, among other aspects. Indeed, large social networks and digital platforms play an essential role in moderating the flow of false information and should therefore be considered necessary allies in the fight against disinformation. Although many measures have been taken to fight disinformation, experts still note major challenges in this field, such as the need for greater transparency.

While Spain is progressing in step with the EU framework, experts stress the importance of consolidating a national strategy that integrates the various dimensions of disinformation to structure the development of public policies in this field. Among other aspects, this should include a digital and media literacy plan. Transparency, accountability and international cooperation are likewise essential elements in any actions involving cooperation between the public and private sectors, and civil society.

Disinformation in the digital age

Introduction

Disinformation has negative repercussions on public health and safety. It undermines democracy and certain fundamental rights we cannot take for granted that must be protected.

Internet and digital progress form a pillar of economic and social development[1]. They represent a great opportunity with many social benefits, allowing the public greater, more transversal possibilities to explore and access information. However, as with any other technology, when used with malicious or inappropriate intent, there may be new threats or an amplification of threats that already existed offline[2]. So, although disinformation is not a new phenomenon, its spread and the seriousness of the risks it involves in the digital age are.

Information plays a central role in society; it is the raw material that forms knowledge, as well as democratic debate and decision-making[3–6]. If information is false or misleading, it perverts or can even block democratic processes. Disinformation is therefore included among the main threats to the democratic system, as it can negatively affect public assets such as health, the economy and national security. It can even erode the rule of law and fundamental rights, influencing or delegitimizing election results with misleading or false information[3,5,7–10]. Although digital development has expanded the scope of false information, this is a complex phenomenon with systemic causes rooted in institutional trust, the economy, and the channels and technologies that mediate the information we consume[6,11–16].

Disinformation is a concern for much of the general public, who also exhibit signs of vulnerability due to the difficulty of recognizing the truthfulness of sources or of perceiving manipulation[17]. The effects are cumulative and endure over time since false and misleading information is hard to rectify: it is difficult to accept that one's own beliefs or opinions are wrong. On the other hand, disinformation has no frontiers and also operates in privately managed spaces, such as social networks, which means that its management requires both international and public-private sector cooperation[11,18]. Another element of this challenge is that the public is at the forefront in this war[19]: their minds are the territory to be conquered in this cognitive battle of disinformation. So digital and media literacy is a key step towards resilience.

The European Commission has taken steps against disinformation, making it clear that democracy cannot be taken for granted; it must be nurtured and actively protected[20]. In Spain, public authorities recognize the threat that disinformation represents for the State[21,22].

An evolving conceptual framework

Disinformation and other information disorders

Disinformation is a specific phenomenon referring to misleading or false information with malicious intent. However, it presents diverse forms and its identification, causes, effects are complex and multifaceted.

There is no general consensus on exactly what disinformation is, although there are many suggested classifications[11,23,24]. One of the most widespread classifications is that of information disorders, which includes three distinct but interrelated concepts[11,25,26]:

Misinformation (*información errónea* in Spanish): This is false, shared inadvertently and without malicious intent, either because someone has been deceived, due to honest belief or to carelessness.

Malinformation (*información dañina* in Spanish): This can be real or false, is not always verifiable, and is shared with the specific intent to cause harm. It may include opinion, personal or other types of information, which is stolen or exposed without permission[25].

Disinformation (*desinformación* in Spanish): This is false information with malicious intent. The European Commission defines it as verifiably false or misleading content that is created, presented and disseminated for economic gain or with the intention to deliberately deceive the public, which may cause public harm[11]. This harm includes its capacity to negatively affect public debate and democratic processes, particularly in an electoral context, as well as impacting negatively on public assets like health and safety, the economy, or security among others[22,27–30].

In practice, disinformation can be difficult to identify[24,25,31,32]. For instance, it is not normally possible to know the intent of the person who issues the false information: it may have its origin in an intent to harm, but its dissemination may be amplified by people who share it without harmful intent[5]. In fact, it takes advantage of the public's (or other agents') good intentions for its dissemination[33]. So erroneous and false information tend to fall under the term 'phenomena of disinformation'[34–36]. Along these lines, in Spanish the word *desinformación* (disinformation) is usually used as a synonym of misinformation, *información errónea* in Spanish, regardless of the issuer's intention[37]. From a legal perspective, only certain extreme practices, mainly connected with malinformation, are included in the Criminal Code[25] (**Key point 1**).

The complexity of the disinformation phenomenon is accentuated when we consider its nature, causes and effects, whose scope are still not clearly understood[4,24,38]. Evidence indicates that the problem has amplified due to the changes that digital development implies in the way information flows and reaches the public[11–14,23,39–42], which is heavily defined by social networks[5]; however, this is a multifaceted problem. In addition to the technological dimension, there is the question of trust in democracy and its institutions, the role of journalism as a guarantor of democracy, geopolitics, the economy, as well as social, individual or cognitive considerations[4,5,33,43,44].

There are extreme types of disinformation that are typified as crimes in the Criminal Code, but most disinformation operates within legal bounds.

Key point 1. Information disorders and the Criminal Code

In 2020, in light of the serious consequences of disinformation that arose in the context of the COVID-19 health crisis, and with the aim of guiding legal actions[45], the Spanish State Attorney General's Office identified several offences typified in the Criminal Code[46] that certain specific forms of disinformation could constitute[47]:

Hate crime: In Spain, data indicate this type of crime is on the rise in recent years (1,724 cases in 2022) with racism/xenophobia (369 cases) and sexual orientation/gender identity (466) being the most common. These are followed by sex/gender discrimination and ideology or religion[48].

Disclosing and divulging secrets: When disinformation is accompanied by disclosure of personal information, etc.

Crime against the moral integrity of a person: In cases that affect a particular person.

Public disorder: Related to false information about terrorist attacks, catastrophes or other incidents which cause alarm, situations of danger for society or require mobilization of the emergency services.

Slander and libel: Slander is 'an action or expression that harms the dignity of another person, discrediting their reputation or attacking their self-esteem', whereas libel is the accusation of having acted or expressed an opinion 'with knowledge of its falsehood or with reckless disregard for the truth'.

Crimes against public health, fraud and impersonation of a professional: Fake therapies, false methods of detecting diseases etc.

Crimes against consumers and the market: This covers all crimes in the Criminal Code related to markets and consumers, of which there are many types, which punish false information in the context of markets and/or consumers.

Narratives for disinformation

Narratives are constructed from far-reaching, flexible messages that combine different degrees of truthful and false information, which increases their verisimilitude. They adapt to local vulnerabilities, reinforcing individual beliefs and prejudices.

Strictly speaking, not all disinformation is false since the concept includes the idea of being misleading. So, it is not simply a matter of what is commonly known as fake news[34] (false information, or more simply, lies)[4,11,27,34,40,41]. The distortion of information presents in varying degrees of false and manipulated content, which can enable the generation of far-reaching disinformation narratives that are often subtle, which makes them more difficult to detect[24,31,34,49-53]. These degrees range from 100% false content that is created in order to deceive, to different degrees of alteration, such as a modified context, establishing false connections or sources, the use of parody or 'hahaganda'[54], as well as manipulated or unrelated images or videos[25,31,55]. Also related with propaganda are the systematic alteration of information, details or sources for a specific purpose, for instance by means of suppressing or overexposing content, or using it in conspiracy theories[5,7,56,57].

Using this wide range of techniques, narratives usually intermix truthful content with false or manipulated content so that verisimilitude replaces the truth[6,24]. These narratives also evolve over time, including subjective elements adapted to the local context and current affairs[4]. For instance, a global narrative that seeks to delegitimize electoral processes may interpret any topical news story as proof of a process being manipulated or null. The Spanish Department of National Security has identified some international narratives of Russian origin, such as 'evil elites against the people', 'traditional values under threat', 'sovereignty and national values under threat', highlighting that they are based on reinforcing extremist political and social movements[54].

Therefore, any subject, incident or confrontation of social importance is susceptible to exploitation by a disinformation narrative that flows at international level[43]. Success resides less in the ability to deceive than in sowing doubt, confusion, distrust or indifference towards a subject, institution or democratic process[10,24,34,50-53] or in fostering the inability to recognize a set of commonly agreed facts that describe a reality[58]. In fact, misleading or false information rarely challenges the beliefs of the recipient, rather it aligns with their ideas and feelings to activate and reinforce their own prejudices or mistaken belief[59]. This tactic seeks to achieve long-term changes in the public that are structural or profound in nature, rather than immediate outcomes from a specific item of false news.

Scope and relevance in the digital age

A matter of national security

Digital developments amplify this phenomenon and make it easier to achieve economic, ideological or geopolitical objectives.

This is a matter of national security that can reach critical levels and have enormous social relevance.

Internet has enabled the rise of new threats to states and amplification of the scope, opportunities and means available to some threats that already existed, in addition to their impact offline. This is the case of disinformation[11-15], which, alongside others, is considered within hybrid threats[60,61]. This constitutes a systemic threat with the potential to destabilize states and democratic processes[7], which makes it a matter of national security.

This is the context in which operations to influence information and foreign interference take place[4,62]. The former covers the use of diverse manipulative tactics and information disorders of different types, while the latter includes coercive efforts that may be used in combination with other actions such as cyberattacks, financial or other types of pressure[4,60]. In a broader understanding of the term 'threat' NATO coined the term cognitive warfare[63].

- False or fake news: The scientific community considers the term 'fake news' inappropriate, mainly due to its lack of precision (the term is contradictory) and because it has become assimilated into political discourse to point the finger at information or media that contradict the agent's point of view. If the information were only false, the problem could be simply addressed with fact checking processes.
- Hahaganda: This consists of camouflaging disinformation and manipulation with humour: ridiculing and humiliating politicians and political institutions to undermine credibility and trust in the chosen target.
- Hybrid threats: These are actions undertaken by state or non-state actors, which seek to exploit the vulnerabilities of the European Union (EU) for their own benefit, using a coordinated combination of measures (diplomatic, military, economic and technological) without engaging in formal warfare.

Despite being considered a type of warfare, in common with other structural threats, disinformation campaigns deliberately operate in the grey area below the threshold of war and usually within the law[59,64]. As a whole, they shape and foster narratives of disinformation[59]. Likewise, part of the challenge of halting disinformation lies not so much in detecting and neutralizing the false information, but rather in being able to connect the information to the large-scale narratives and their objectives in the long or medium term[65,66].

Different ends for the same means

This is not only an external threat. The objectives and social actors connected with disinformation also exist at domestic or national level, and it is these actors who are of greater concern because they are usually better perceived by the general public[17,67]. The main motivations behind disinformation campaigns are often[7,9,12]:

- **Economic**: Derived from the interests of the instigator, ranging from fraud, unfair competition or the intent to strengthen or protect a specific sector/economic activity, to capitalize on disinformation as a business model[8,9,67,68].

- **Ideological**: Here, the main motivation is to influence the results of electoral processes using disinformation[64,67,69,70]. In Spain, the Spanish Department of National Security[67] identifies the main ideological motives as discrediting governments, political parties or candidates, and undermining public confidence in the integrity of the electoral process.

- **Geopolitical**: This seeks to destabilize democracy or provoke a disproportionate reaction, questioning the democratic nature of a country or its international prestige to generate division and uncertainty about socially important subjects, among others issues[9,13,67]. It is an attempt to mould new social values that can subvert the foundations of democratic societies[4].

Motivations often converge and intermix external and national social actors to such an extent that it is often not possible to identify which is which[4]. Despite this, consensus exists that disinformation strategies multiply during situations of great social importance[24,71,72]. For instance, the group of experts within the Spanish Department of National Security framework highlight[4] certain studies on the role of disinformation in the 2016 USA presidential elections,[73] the French elections in 2017[74], the Brexit vote[75,76], the armed conflict in Ukraine[54] and, more recently, in the conflict in the Near East[77,78]. Particular attention should be paid to the infodemic[79], that arose around COVID-19, with misleading large-scale narratives at international level regarding its origin, prevention or the danger of vaccinations[53,80-83]. In the case of Spain, studies and reports exist that show the role of disinformation in areas such as COVID[84-86], electoral processes[67,70,87-89], the issue of Spanish regions and Catalonia[59,90] among others[66,70,86,91].

A new social and informational context

Social networks, messaging services and large digital platforms have changed the way in which information flows and reaches the public thanks to an explosion of new channels and possibilities that form the basis of informational disintermediation.

Experts agree that the digital ecosystem, understood as the large social networks, platforms and search engines alongside private messaging, has been decisive in the democratization of and global access to information, its acceleration and exponential growth. Today, they are considered the main means of communicating information[11-15], particularly social networks and messaging services[24,92]. However, evidence suggests that the speed and scope of false information spread by these means surpasses that of truthful information and amplifies the false content[93]. The content, channels and actors involved in how society is informed have also multiplied. In consonance, the classic roles of issuing agent and receiver have become blurred into 'prosumers'[12]. These aspects facilitate decentralization: there is no single source, but rather many, normally coordinated, sources with a multi-directional online flow of disinformation[10]. This new context accelerates the process of disintermediation, in which the traditional interpreters of reality, such as mainstream journalism, political or other professional or institutional actors, are replaced in the digital ecosystem.

· Infodemic: Term coined during the COVID-19 epidemic to describe the surfeit of information on the subject.
· Prosumer: A consumer of a product or service, in this case information and diverse content, who at the same time participates in producing it.

The technology itself and any user, real or not, anonymous or known, become the prescribers of content[6,94]. Moreover, the general interest in news, understood as journalistic information, has continuously decreased in recent years and is a minority interest among internauts, which weakens their ability to halt disinformation[95,96].

Agent: instigators and distributors

Actors may be national or foreign, or both if there are shared interests. The digital ecosystem makes it difficult to identify them.

Internet amplifies the possibilities of hiding identity, generating trust and eluding the checks and control mechanisms aimed at limiting false information[4]. Agents may be state or non-state actors, including proxies of both[4]. State actors may be governments and/or their associated structures acting on their own or foreign populations. Some studies indicate that around 81 countries use social networks to disseminate propaganda and disinformation based on well-trained cyber groups or 'cyber troops'[97]. The volume of business their activity represents is estimated at around 9 million euros for 2020[97]. Non-state actors include corporations, lobbies, marketing agencies and ideologically-based advocacy groups, like political parties or other formations, such as religious or ethnic groups, among others[4,7,70,71,98]. The interests of state and non-state actors may converge, the distributing agents may be related with the instigators or not, and may act deliberately or unintentionally, which means that any person, organization or institution is susceptible to 'broadcasting' and receiving disinformation[99].

It is important to note that this wide variety of actors normally operates in a coordinated manner, made easier by the digital ecosystem, which in turn highlights the systemic nature of the threat [59].

Channels: digital impact and prevalence of classic channels

Social networks, messaging services and search engines have become the main mediators of information, and they foster the flow of disinformation. Political and media elites still have a major role in the mass amplification of false information that circulates on internet.

In today's information ecosystem, new online channels live alongside other information outlets that previously existed offline and their current digital versions[96].

The digital world has enabled an explosion of thousands of mainstream journalist publications[100] and other types, such as public social networks and private messaging services, to transmit information with new forms and formats, like news aggregators, blogs, podcasts and a long etcetera. Social networks and large digital platforms are the main route of access to information and the news, an intermediary role that was traditionally performed by mainstream journalism; however, these platforms and networks do not follow any professional principles[6,43,101]. They have particular weight among younger people[95] whose use of networks like Tik-Tok is prevalent[96]. For the general public, Facebook and WhatsApp are the most common media to be informed and interact with the news[96]. WhatsApp is particularly preferred for sharing news[96]. This leads to an explosion of informational possibilities, which can be liberating but also overwhelming[6]:

- The digital ecosystem brings an information overdose that covers all types of information both within and outside the journalistic domain, mixing opinion with information, where immediacy is king. While access to all types of content is promoted, the capacity to understand or use the content as knowledge is not[43,102,103].
- Any information can appear to be journalistic or be sponsored by groups, institutions, experts or influential people[24].
- It is quick and cheap to inject information online. In fact, disinformation campaigns may be fed with the creation a network with its own media and resources[18].
- This has a low reputational cost for the instigators, thanks to the difficulty of attribution, the lack of critical awareness of a lie or to greater social acceptance of it[6-8,71].

- Proxies: May be agencies, organisations, etc. There is no apparent, public or direct link with the instigators, but proxies are covertly connected, funded and controlled by them.
- Cyber groups: Are groups or individuals who have financial or other types of backing from governments, political parties or other organizations, with the mission to manipulate public opinion online. Their main objective is to disseminate propaganda and disinformation in order to influence the perception of events, political and social issues or any other subject of interest. The concept of 'cyber troops' is a clear reference to the concept of a hybrid war or other operations to gain influence.
- Information overdose: This term refers to the excess of true or false information available or received on a subject, which can result in cognitive saturation or a difficulty to effectively process and assimilate the information.

Alongside these changes, some experts note a strong relationship between the flow of false information and what has been called the information disintermediation crisis[6,24,59,104-106]. In Spain, half the population receives information that has been algorithmically curated rather than content subject to editorial decision or professional journalistic or institutional criteria that define the quality of the information. Moreover, on some social networks, the primary source of news is not necessarily a professional journalist[95,96].

However, despite these profound changes the traditional elites, represented by mass mainstream journalism, political parties or major institutions, still play an important role[44,70,98,107-109]. The impact and scope of false information exponentially increase when it is collected and disseminated by them. In general terms, television closely follows digital media as the preferred channel for information. This shows the responsibility and importance that these channels and actors still have[70,96,110].

Content

Disinformation relies on general techniques, such as affective attraction or repetition, to increase its influence. Among the main subjects targeted by disinformation in Spain are politics, electoral processes and social challenges like migration.

Any subject can be the target of false information, but trends vary over time and between countries, are conditioned by language as well as local cultural aspects[66,111]. This is why the flow of disinformation is more common between societies that share a language[112,113]. To improve acceptance of false information, messages based on affective attraction are employed, as they foster an emotional response. These may include visual components, and their credibility is enhanced by their straightforwardness, their apparently solid narrative and simple repetition[25,114-116].

IBERIFIER, the digital media observatory for Spain and Portugal, funded by the European Commission and linked to the European Digital Media Observatory (EDMO), in line with the results of recent studies and reports, groups the content of false news in Spain into main subject areas: politics and elections, health, environment, migration, gender, famous people, security and sexuality[66,70,88,89,117,118]. Specifically, the predominant content during the first third of 2023 was about climate change, politics and elections, with narratives about the climate and electoral fraud related to the restriction of rights[89]. On the other hand, false and misleading information about health and science spiked during the COVID-19 pandemic and has prevailed, or even increased, since then[34,86,119]. The EU has recently devoted over a million euros to deepening its understanding of disinformation related to some of these subjects[120].

Receiver

Any individual or institution may be the target of disinformation campaigns tailored to the receiver and the intended goal.

Information empowerment of the individual involves benefits but can also debilitate the capacity to tackle disinformation.

Any person can be a receiver and/or distributor, whether intentionally or unintentionally, of false information. Even so, the understanding is that certain groups exist which are particularly vulnerable, like older adults or people who are in a situation of social exclusion[19,121]. In Spain, it is estimated that 53% of the population experiences daily exposure to false information, compared to 37% of the EU population[17]. Nevertheless, its distribution is not random, but charged with intention[6,122]. Any institution, public or private organization, social collective or individual can be the objective of campaigns designed ad hoc to achieve specific goals, which often play on people's emotions[18]. In addition, the avalanche and immediacy of information, while promoting empowerment and informational self-sufficiency, also pose significant challenges by breaking the classic one-way information flow (as anyone can access 'all' information). In this new context of disintermediation, it is difficult to distinguish information from opinion, to know the level of truthfulness, establish trust in sources, etc., which means that instincts, emotions or personal/social bias easily govern our relationship with information[6]. Taken as whole, this creates a climate that predisposes people to believe dis or misinformation.

· Algorithmic curation: Social networks, digital platforms and search engines offer content that is selected or filtered and ordered in accordance with the criteria of a given algorithm, mainly aimed at tailoring them and holding the user's attention.
· Primary source of news: In Spain, professional journalists and their news brands are the most popular sources of news on Facebook (46%), YouTube (44%), Instagram (42%) and Twitter (57%). At the other end of the spectrum, the general public (44%) and alternative media (35%) dominate as information sources on TikTok.

Contemporary phenomena involved in the rise of disinformation

Disintermediation of information occurs in a social context where many highly interrelated factors facilitate the presence and impact of false information. The complexity and dynamic nature of the phenomenon make it difficult to clearly distinguish its causes and effects.

The disintermediation crisis interacts with the many dimensions that define institutional, political, economic, social or individual reality, giving shape to the phenomenon of disinformation[12,16,33,123]. For these reasons, disinformation is a multifaceted, multi-factor phenomenon that is steeped in a series of overlapping, mutually reinforcing, dynamics. Some authors emphasize the importance of digitalization of the public sphere, with algorithmic filtering, a weakening of the structures of professional news organizations, polarization, a rise in autocracy, loss of trust, or an increase in conspiracy paranoia among others[12,33]. The complexity of the relationships means that some of these factors can constitute both a cause and an effect of disinformation, depending on the focus or sources consulted[7,8,12].

Trust and the democratic framework

Disaffection with democracy as the backdrop

Trust in the democratic institutions, related to public wellbeing, reinforces the resilience of states and their societies against disinformation.

Certain studies report a deep global crisis that affects both the number and the quality of democracies. This is related to an increase in disaffection and distrust of democratic institutions and their guarantors, such as mainstream journalism[6,59,124–126]. This is a vulnerability that weakens the capacity of institutions to fight against disinformation narratives, increasing the public's susceptibility to its effects[26,127] while also acting as a barrier to response [20,127]. Although it is not possible to identify specific, unequivocal causes, experts indicate that there is a structural component, principally related with the effects of and response to successive economic crises and an increase in inequality. There is also a socio-affective element[5,6,58,59,115,124,128] that compounds, for instance, collective or personal feelings of grievance, disillusionment with a general lack of interest in the democratic system, institutions and politics, or dissatisfaction and a search for meaning and identity in the globalized context. All of this engenders individuals predisposed to believe, whether for social or psychological reasons. For instance, during COVID-19, some studies have shown that institutional trust was a key aspect in mitigating the influence of the infodemic and fatigue derived from the situation of exceptional crisis[129–131].

Geopolitics and the international framework

In the geopolitical context, disinformation is increasingly used as a tool to destabilize third countries.

Trust between states also plays a major role. In the last decade, experts stress the growing awareness and use of 'soft powers' in the international context. These are related with the use of culture and communication, including false and misleading information, in the same way as any other weapon in the arsenal of states or geopolitical spheres of influence[42,132]. Likewise, the progress and consolidation of so-called hybrid wars, which include disinformation, create a scenario that is prone to and ripe for this threat[133,134]. International tensions became more acute, particularly during COVID-19[135] and, more recently, after the Russian invasion of Ukraine[4,62,136–138]. Among the best-known actors, the EU[139–141] and many reports[7,9,71,142] point to Russia, specifically the Russian Internet Research Agency, and to a lesser degree, other countries, such as China[71,142,143].

Information mediation and journalism

The economic crisis, job insecurity in the sector and lack of public trust have undermined the capacity of journalists to fight disinformation and the role of journalism as a guarantor of democracy.

Different factors have weakened the role of mainstream journalism as the primary mediator of information and safeguard against disinformation[6,43,49]. Experts and data in Spain indicate a financial and professional crisis in the sector, resulting in reduced spending and increased job insecurity for editorial staff, as well as a trust issue[4,24,39,43,49,95,144,145]. Lack of resources, information overdose and the immediacy of news, alongside competition for attention, advertising, and positioning in the digital ecosystem[43,95,146,147] have stood in the way of accurate information[148,149].

The context has become a breeding ground for the appearance of information with a widely varying range of quality. This may be due to the standards of accuracy of the information, including so-called pseudo-media[150], an increase in sensationalism to obtain attention (click-bait)[151,152], the use of networks as a source[39,43,150,153-155] or the rise of informers[156]. All of this, along with the accidental publication, lack of resources, or intentional dissemination of misinformation[43,49], is related to the crisis of trust experienced by the sector[43,127,157,158] and an increased economic dependence on public authorities and loyal audiences[39,159,160]. As a whole, these aspects contribute to undermining trust and, in contravention of codes of journalism[161], can increase a lack of political or financial independence or the perception of it[39,159,160], polarizing both the media and audiences[95,162,163]. In the case of Spain, some experts connect this perception with a lack of transparency and an arbitrary[163-165] use of institutional campaigns in the media[166].

Social fragmentation

Fragmentation of public debate helps the influence of disinformation. The lack of a culture of social and political debate can promote this phenomenon through affective polarization or the post-truth context.

The growth of phenomena such as post-truth[12,167-169], conspiracy theories[57,170,171], and ideological or affective polarization[172] that can be related to a state's political culture and the culture of debate, affects the progressive fragmentation of public spaces and the capacity of false information to penetrate society[6-8,12,31,33,49,70,171]. Such phenomena have changed the way that society relates to falsity but also to truthfulness. Hence some experts indicate that there is an epistemological or even ontological crisis or alteration in what we understand by 'the truth'[6,12,98,173,174]. On the one hand, these phenomena foster the introduction and acceptance of false or biased messages that directly appeal to emotions, feelings and beliefs. On the other, particularly in the case of conspiracy theories, they change the relationship of society with objective evidence and scientific knowledge itself, to the extent that evidence and knowledge may be completely rejected. As a whole, they impede rational thought and make it easier to accept the narratives of disinformation[175,176], which can hamper or even block social debate, amplifying distrust in institutions[4,33,49]. Within the EU, Spain is considered one of the most polarized countries, not so much in political, but in affective and social terms[24]: even subjects with no apparent ideological weight are now affected by polarization[12,177]. Nevertheless, not all emotional or polarizing content uses false or misleading information.

Affective polarization has increased significantly across various countries. Evidence for its causes is still limited, particularly outside the USA, with varying results observed across different countries and studies[12,101,178,179]. Foremost among the causes is political culture, marked by a polarization of the elites, the proliferation of social networks and hyper-partisan means of communication, or the lack of a culture of debate[12,24,101,178-180] alongside structural factors related with the aforementioned disaffection. This situation converges with socio-affective factors derived from feelings of social dissatisfaction or disaffection with democracy, which favour the diffusion of dis and misinformation[33,115] to fan the flames of general indignation[181,182], incite chaos or appeal to a desire to 'set the world on fire'[183].

Cognition and individual vulnerability

Psychology plays an important role in why false information is shared or believed, but there are other causes that vary depending on the subject of the information.

People play a central role in amplifying false information because they are the main receivers and distributors[15,93]. Understanding the factors and mental shortcuts that intervene in the process of deciding what is true or false, and how beliefs are formed about this[115,184] becomes a matter of great importance[115,185,186]. However, the factors that define susceptibility to, acceptance of and the potential to disseminate false information are diverse, and establishing a general pattern is impossible[15,86,115,187]. Influences can vary depending on the people, their personal or social context, from one country to another, or even on different subjects[86,111,115,123].

· Pseudo-media: Digital media that are created with the aim of exploiting an advertising business model based on the attention economy of the digital ecosystem, without paying attention to criteria of journalistic quality, ethics or any other.
· Post-truth: This is a perception of reality that is linked to or denotes circumstances in which appealing to emotions and personal belief is more influential than objective fact in terms of forming public opinion. This makes it difficult to perceive truthfulness. Post-truth is not only a lie, but also a distortion of the truth loaded, above all, with intent.
· Conspiracy theories: They promote simplistic, self-justified ideas regardless of the real or most likely facts about complex issues. Such theories state that certain events or situations are secretly manipulated by powerful forces with negative intentions. They usually have a set of well identified elements in common.
· Affective polarization: This refers to the emotional distance between the affinity we feel with people whose political ideas we sympathize with, and the rejection we feel towards those whose opinions we do not share.
· Ideological polarization: Is the degree of divergence between people who hold beliefs that are either consciously conservative or progressive on a wide range of socially important subjects.

Demographic factors

There is no clear consensus on the influence of demographic factors.

The findings related to demographic factors are contradictory[15,122]. In the USA, increasing vulnerability with age is a well-confirmed factor, whereas for Europe the data in this regard is not so conclusive[5,188]. In the case of minors, there is particular concern about their exposure to platforms based on audiovisual content, which requires further investigation due to the lack of information[189]. Regarding other demographic factors, the general relationship between educational level, income or gender and the tendency to believe or share disinformation varies depending on the study[15].

Cognitive predisposition to believe and share

Cognitive bias exists and predisposes a person to believe and share false information.

Many cognitive factors are related with the tendency to believe false information[15,115,190]:

- Level of critical thought[115,191]. This may be connected with dual process theory, trusting to intuition rather than thinking[115,192], and with the bias of motivated reasoning itself that, like confirmation bias, favours reinforcing one's own ideas and biases.
- The illusory truth effect[115,186]. In essence, this means the predisposition to believe information that sounds familiar[193] even though it is false. This may occur due to repetition, consistency with previous experience[194], or through simple messages in accordance with the cognitive miser theory and heuristics.
- Cognitive errors. These are mainly related with the way in which sources are perceived and the knowledge or predisposition an individual has about a subject or the false information itself[115,195]. This type of error includes lack of attention to or the accuracy of the information received[185] or excessive trust in a source[196,197].

As for why the false information is shared, when this is done unintentionally, it is usually to share something important in good faith[198] and from the impulse to simply share information on social networks[199] or due to a lack of interest or attention[185,200]. When false information is shared deliberately, evidence shows that this is done mainly for self-interest, to signal belonging to a group[201], to gain notoriety[202] or as a mechanism to cover other psychological needs related with social discontent[181-183].

Socio-affective context

Disinformation may exploit aspects derived from how an individual sees themselves in society and the emotional or affective status derived from it.

Affective aspects amplify the information's power of persuasion[203]. They serve to generate false beliefs and spread them in order to exploit an emotional or moral component within the information[181,182,204]. This is typical of sensationalism, which takes advantage of a person's emotional state[115,205] or attempts to induce one by exploiting emotions such as fear or insecurity. Some studies associate personality traits[15,185], habits and beliefs[187,206] or ideology[200,207] with vulnerability and a tendency to interact with false information. They also highlight the strong influence of mass mainstream journalism, influencers, political elites and experts on the reach of disinformation and public perception[115,208].

Durability and continued effect

There is a large body of scientific evidence about people's resistance to changing their mistaken beliefs or accepting the falsity of information that is in their interest.

People show a great resistance to accepting rectification of false information and consequently changing their mistaken beliefs. These beliefs may last a long time, regardless of the cognitive skill of an individual or rectification of the false believe. This is what is known as the continued influence effect[184,209]. In a more extreme example, continuous rectification may 'backfire' causing a rebound that means people remain unconvinced of the truth and cling on to a false belief[210,211]. While it may seem like an overstated adverse effect, resistance to

· Dual process theory: This concerns analytic thought -based on logical reasoning- and intuitive thought -based on emotions and feelings-, although recent evidence calls into question the relationship between them and misinformation, highlighting their complexity. The cognitive miser theory and heuristics favour simple messages that do not require a lot of processing power to understand.
· Motivated reasoning: We validate what is in line with our ideology and world view. Confirmation bias or cognitive dissonance reinforce this tendency.
· Cognitive miser theory: Refers to a type of bias that leads our perception towards information confirming our own beliefs, thus reducing the mental effort required to process it.

correction and the backfire effect are particularly pronounced when the refutation clashes with beliefs or values that form one's self-identity[115].

Digital governance and business models

The digital business model, based on attention, responds to economic interests that may encourage a lack of privacy, exposure to and circulation of disinformation, regardless of the negative aspects that this may entail for the public.

Geopolitical factors, institutional trust, as well as social and psychological bias that works outside the internet may be reflected online. Experts highlight that internet is not a universal, homogeneous network that is anonymous and neutral. Even if the technologies underpinning it are universal, the standards it operates with vary between countries and platforms[212,213].

There is a connection between areas of the digital economy and disinformation. Specifically, areas that are based on the structural business model of many digital platforms, which monetizes the attention of users by showing them advertisements[5]. In this relationship, the power of users to intervene is very limited[5]. The value of content is its ability to attract attention[13,151], and the value of attention resides in its use to modify behaviour in order to, for instance, sell a specific product[214]. The sale of online advertising space is therefore usually the main source of income for large social networks and many websites[13]. This model fosters the distribution of false or manipulative content capable of attracting attention over and above truthful content[13,101,215,216], by means of affective attraction and characteristic messages that increase acceptance[151] such as sensationalism[151,217,218]. Such content is highly profitable if it is monetized with advertising[7,11,27,68,214].

There is a proliferation of businesses that, in essence, generate false content to make profit in this way[8,219]. Some data indicates that around 200 million American dollars in advertising ends up in domains identified as channels of global disinformation (data for 2019)[220], a figure which is 76 million if we focus on data for the EU (data for 2020) [221]. Global estimates on the propagandistic and disinformative activities of cyber groups associated with states also amounts to another 10 million (data for 2020)[97].

User data is another element of value in monetization: they are the basis of a new economy[222]. Data is compiled and analysed with the aim of directing the personalized content that will most attract attention and the advertising that will be most effective for each user[7]. Personal data and user behaviour can be deliberately revealed with consent, as happens on many social networks, or derived from the technologies monitoring online behaviour. They can also be inferred from the information and interactions with other users, 'friends' in the context of many social networks, or even offline if we consider the internet of things[13,223,224]. Likewise, they can be directly purchased or sold for a specific purpose[225]. Sensitive information like political or religious ideology, which is protected under the General Data Protection Regulation (GDPR) is easily deducible from an analysis of these datasets without the need to infringe the GDPR[13,226-228]. Social networks are, therefore, a risk for data protection and the privacy of the public[5].

With their personal information, citizens finance 'free' services, like the social networks themselves, or other types of freemium without a clear understanding of their downsides and possible outcomes[214]: disinformation and manipulation, psychological harm, addictions, loss of privacy, etc.[5,214,229]. The architecture of social networks and digital platforms fosters this situation. It usually promotes lax privacy configurations to increase user participation and can also be based on configurations, chosen or not, that limit the individual's perception of the information available, the climate of opinion or many other aspects[5].

· Monitoring online behaviour: General behaviour online is continually trawled using technologies like cookies, scripts and tracking pixels, advertisements, CSS/HTML code or even by third-parties who intervene in the platforms and systems we use by means of application programming interfaces (API) or other techniques.
· Freemium: This term is commonly used in the digital services and applications environment to describe a business model where the basic version of a product is available free of charge, with additional characteristics available for an additional cost. It is very common in many mobile applications, digital services etc.
· Architecture: In the context of websites and social networks, this refers to how they are designed and present options and configurations for users. These architectures are design strategies that influence the decisions users make when they interact online.

This business model exploits personal data so that the technologies which govern them, algorithms and targeted advertising techniques show a user the world that, according to calculations, he or she wishes to see[6]. The main motivation, as experts highlight, is private profit, which means that public accountability is limited as are any solutions that might be reached without the need for regulatory intervention[5,230]. Likewise, the structure can be put at the disposal of third parties for purposes other than purely commercial ones, which may include propaganda, an aspect that is of particular concern from an electoral perspective[8,231].

Technologies that can be used for disinformation

The basis is a set of technologies related with big data analytics and artificial intelligence[13]. They are in constant evolution and contribute to the role of internet as a mediator of the information on social networks and major search engines/platforms, as well as underpinning the attention economy. They therefore play an important role in false information distirbution[232], contributing to its generation and dissemination. Foremost among such technologies[13] are:

Algorithms: echo chambers and filter bubbles

The algorithmic curation of information compromises the neutrality and plurality of information users receive. In the search for users' attention, it can foster dispersion of false or misleading information and create conducive environments for disinformation.

The order in which information appears and is shown to each user on social networks and search engines is determined by recommendation algorithms and algorithmic curation. These algorithms can compromise a user's access to neutral, plural information and deliberation on internet[233,234], and this mediation concerns more than half of the Spanish population[96]. Each social network and search engine has its own algorithms, which are in constant development. This indicates that the premise that social networks are representative of public opinion can be considered false[6]. Thus, the internet is strongly controlled by private corporate algorithms designed to maximize earnings by attracting user attention, without necessarily considering the possible psychological or social effects[230]. Indeed, the habitual lack of transparency regarding these algorithms is a limiting factor in the fight against disinformation by users or experts because it is obstructive[169,178,235-238] as it obstructs the identification of bias, our understanding of its influence on social and individual behaviour, and the development of detection and prevention mechanisms.

Attention-based algorithms may foster exposure to false and misleading information by offering content of an impactful or sensationalist nature, or which arouses radicalization and extremism[13,239-241]. This way of offering information has been linked to 'filter bubbles' and 'echo chambers'. Although they are different concepts, both are indicative of a lack of exposure to opinions other than one's own, and the creation of silos of self-referential truths, which increase polarization or block public debate[6,13,179,233,241-244]. Nevertheless, social networks and platforms also give a voice to marginalized and disadvantaged communities[5]. There is also active scientific debate about these phenomena and their effects[178,241,245,246].

On the one hand, emerging evidence shows that echo chambers are less common than had been assumed[178,247,248] and that the proportion of the population that reaches them is a minority[92,207,233] who are already highly polarized[178,247]. Nevertheless, they can have a significant effect on public debate[178]. On the other hand, the effect of algorithmic filtering on the quality of our information diet heavily depends on the diet it is compared to. The latest empirical evidence indicates that algorithmic filtering does not necessarily limit the information diet[178] and disputes some of the negative effects, like polarization[245,246,249] and radicalization[239,240,246,250], among others. There are even studies indicating that it could reduce exposure to false information[245]. Although these results reinforce the idea of factors other than technology as drivers of polarization[179,246], they may also reflect recent changes in the

- Filter bubbles: A phenomenon in which a person is mainly exposed to information and perspectives that reinforce their beliefs and pre-existing points of view because of the personalization and recommendation algorithms of digital platforms and social networks. This happens without any voluntary action by the user.
- Echo chambers: This is an environment or platform where ideas, messages or concepts find a receptive audience, which amplifies their impact. Here, the user has an active role, demanding specific content for different reasons, including ideology, which enables disinformation to circulate more freely.

algorithms of various platforms aimed at increasing moderation[239]. However, the lack of algorithmic transparency makes it enormously difficult to confirm these studies[169,178,235-238].

Misleading profiles: bots, cyborgs, trolls and fake groups

Deceptive accounts and groups are important in the dissemination of false information in the digital ecosystem. They can cause reputational harm to third parties, trigger affective polarization, infiltrate disinformation into the news agenda, or be used as an example of what happens in society.

On social platforms, fake groups and profiles abound, and their main function is to spread disinformation, almost always in an automated, large-scale manner[13]. These profiles may be totally automated systems based on artificial intelligence (bots), semi-automated accounts with human intervention (cyborgs) or completely human users (trolls)[13,41]. They may have the financial backing of an actor who orchestrates the disinformation campaigns[41] and often form fake or hybrid groups (bots and humans interact)[8].

Despite the efforts of platforms and social networks to limit their presence, the malicious use of bots and fake profiles to spread disinformation is growing and experiencing an economic boom under the on-demand services model[251]. Their number and the types of activity exponentially increase just before electoral processes[13,252,253]. Nevertheless, although certain dissent exists[31,93,254,255], verified accounts (recognisable people) seem to have a more important role in dissemination than artificial and fake accounts[31,93,255]. On the other hand, advances in artificial intelligence techniques make it increasingly difficult to distinguish bots from human users[13,31,256,257].

As well as acting as disseminators, fake accounts also act to silence or inflate the popularity of users and content[8], which influences algorithmic filtering[258-262], placing false and misleading information on the political news agenda by stealth[261,263] and promoting apparently spontaneous artificial currents of public opinion (astroturfing)[258,264,265]. These accounts also instigate social tension and affective polarization by means of false flag messages that use extreme terms to stress, and on occasion ridicule, an ideological, social or scientific stance, generating a response and a climate of reactionary opinion[266,267]. In a country the size of Spain, the control of a few hundred accounts[8] is sufficient to make a significant impact on a social network like X (previously known as Twitter)[13].

Data analysis and micro-segmentation.

Micro-segmentation allows management of advertising at the level of individuals or small-groups, but can also be used for propaganda, which can undermine political debate and make it easier to influence election processes using false or misleading information.

Micro-segmentation is a commonly used practice in advertising that allows content to be directed at specific groups of users based on their characteristics, feelings, etc., personalizing the messages they receive[13]. Some recent research has noted its potential to reach a new level of ultra-personalization (nano-segmentation) on networks like Facebook, with campaigns targeted at individuals[268].

The use of micro-segmentation for purposes of disinformation can negatively affect the public as it enables the control and automation of data compilation, as well as selection of the channels and content that will have an impact on users[13,184,269-271]. It can be based on many different tools, some of which are commercially available from the digital platforms themselves, among which are dynamic prospecting[13,272,273], programmatic advertising or psychographics[35]. In Europe, the use of micro-segmentation is restricted under the GDPR[274], although some experts believe this to be insufficient[228,275], and data show that the public reject this practice[5].

- Astroturfing: This is the strategy of online manipulation involving the creation or promotion of a false impression of public support for, or opposition to, a cause, idea, product, individual or policy. Both individuals and groups attempt to make their messages or actions appear organic and spontaneous when they are in reality being orchestrated or covertly funded.
- Dynamic prospecting: This marketing technique involves the real-time automated adaptation of advertising and publicity depending on the information about a user and their attention-based, emotional and economic behaviour. It allows thousands of versions of an advertisement to be generated so as to attract attention in a tailored way. It is based on the compilation of user data, preferences, search history or geographic location to personalize advertising content, which is shown in an individualized manner.
- Programmatic advertising: This is an automated way of buying and selling advertising space online that uses algorithms and technology to facilitate and optimize the processes and objectives of advertising.
- Psychographics: The study and classification of characteristics, attitudes, values, interests and behaviour of a set of people or a specific audience.

An aspect of particular concern is the use of this technology to spread ideological or political content[228,276,277,278] above all electoral content. It can foster emotional polarization and make it difficult to understand the general stance of parties, concealing or stressing only certain postures tailored to each individual. This can undermine public, democratic debate and potentially manipulate votes[5,13,225,226,275,277]. Some experts suggest the need for electoral regulations adapted to such practices and empirical research in the European context to understand who uses these tools and what their effects are, in order to limit their potential influence on the right of freedom of vote[7,225,228,275,277,279]. The EU is currently addressing regulation of segmentation in political advertising[280].

Encryption and private messaging

Private messaging is one of the main vehicles of false information, but encryption -which is necessary for privacy- means it is difficult to identify and mitigate.

Private messaging platforms are the preferred channel for sharing information and news[96]. These platforms include the internal messaging services of social networks and applications developed exclusively for the purpose, like WhatsApp, Telegram, Signal, Skype, etc. Foremost among the advantages they have for false information spread is the greater trust with interlocutors and the opacity that encryption, in most cases, gives them[13]. On one hand, this characteristic is key for privacy and, as many experts note, improves the cybersecurity of communications in accordance with European and Spanish policies[2]. On the other, encryption prevents identification of false information and its associated agents, etc., even by the service provider company, which also limits their liability in this regard[13]. Experts indicate the need to face the challenge of gaining better insight into how disinformation flows in these spaces[281-283].

Generative artificial intelligence

Artificial intelligence can generate false texts, images or videos that are often indistinguishable from truthful content and can even affect fundamental rights. This is an area that may require regulation.

The different techniques included in the concept 'generative artificial intelligence' are notable for their capacity to produce false and manipulated content of increasingly good quality images, audio, video and texts. Although these technologies offer great opportunities and have positive impacts[284], they also involve major risks, such as the potential violation of different fundamental rights[285-287]. There is a wide spectrum of quality in this type of manipulation, from the obvious -such as those used in memes and 'hahaganda', which can also disinform[13]- to large language models, deepfakes and voice cloning (**Key point 2**). They can be used to produce false content that is practically indistinguishable from reality for both human and artificial intelligence. This marks the end of *seeing is believing*[6]. This is why they represent a major challenge in terms of disinformation[288].

This is an aspect of artificial intelligence that may need more regulation in the short to medium term. Despite efforts, generative technologies surpass those of detection both in the case of text[294,310-312] and in deepfakes[288,299]. The techniques are developing rapidly, which makes it difficult to implement early warning or formulate preventive public policies[285]. To mitigate the problem, materials have been developed with the aim of making it easier for the public to detect them[313] and to guide the efforts of policymakers[286,292,341-317] beyond the generic framework of artificial intelligence[318,319], worldwide[320,321] and in Europe[285,318].

Large language models such as those used by Chat GPT or deepfakes are some of the technological developments with the greatest potential to take disinformation to a previously unknown level.

Key point 2. Technological development as a tool to generate false content.

Large language models, foremost among which are GPT (generative pre-trained transformer)[289], enable analysis and generation of text: summarising content, translation, answering complex questions, etc. Experts point out that machine-generated text could take disinformation to previously unknown levels[290-293] particularly in aspects like health or the climate question[292-295]:

They allow false content to be disguised as real by imitating the styles and formats of trustworthy sources, thus avoiding detection[31,288,296].

They amplify when integrated into fake websites, bots, or influence algorithmic filtering by cross referencing artificially generated content[13,228].

A lack of exhaustive control of the sources used to train models so they can provide answers facilitates the generation of erroneous, imprecise, false or biased content[288,297].

The term deepfake[298] is defined as manipulated or synthetic audiovisual media, produced using artificial intelligence techniques, which seem real and show people apparently saying or doing something that they would never have done or said[285]. The possibilities of sowing doubt about anything that can be seen or heard are almost infinite and general access to it is constantly increasing[288,298-302], which calls into question the concept of *seeing is believing*[234]. As well as enabling more effective a la carte disinformation, there are currently many examples of harmful uses of this technology[303,304]. Among these are their use to generate non-consensual pornographic content[305], including cases with minors[306]. This is an issue that particularly affects women and, apart from defamation of character, involves other forms of violence, like intimidation or blackmail[285,307].

In the field of research into content generation and detection, the most active area is the manipulation of faces in images or videos[13,288,299,308].

Other areas of the digital world that may be important are the use of virtual and augmented reality, including the metaverse[13,309]. Some noteworthy elements here are intelligent assistants or devices, decentralized autonomous organizations (see Cybersecurity[2]), multimedia games and what is known as transmedia storytelling[13].

Impact

There is a need for developments that give us better insight into the impact of disinformation, particularly on society. To date, available evidence relates its impact with major effects on individuals, which can trigger democratic disorders.

There is broad consensus among experts regarding the potential impact and danger that disinformation represents, justifying the significant number of studies devoted to this area. A major challenge persists in strengthening the empirical evidence that allows us to establish causality[7,9,24,29,39,101,115,185,322,323] .

Most causal evidence focuses on the short to medium-term effects on the individual, including effects on online behaviour and emotional response[93,324-329] and the consequences on health in the case of the infodemic[15,84,86,324]. In terms of the individual and social decision-making, there are studies attributing negative effects to false and misleading information and, above all, to the use of social networks[322]. However, it is not always possible to establish causality. The negative effects include developing dangerous conduct or hostile attitudes, committing hate crimes[5,203,330], reduced trust in institutions[44,208,331,332], increased polarization[331,333,334] and changes in vote attributable to the use of misleading or false information[331,334,335].

· Transmedia storytelling: This refers to telling stories whose narrative universe develops across multiple different media and communication platforms such as social networks, websites, video games, podcasts, television programmes, books etc. Each medium or platform contributes a unique part of the story and is used to enrich the global experience of the viewer or reader.

· Causality: Seeks to establish one variable or factor as the cause that explains the behaviour of another, which can be called 'effect'. In research, the reference method to establish this type of relationship are randomized controlled trials, which are very common in the clinical setting. They involve the random assignment of participants to a control group (who do not experience any actions) and others to an experimental group; this makes their application to the social dimension of disinformation difficult.

Several factors impede progress in understanding the impact of disinformation and misinformation. On the one hand, it is methodologically complex to establish a connection between attitudes and behaviours determined by the exposure to false or misleading information and to demonstrate how this translates into changed or reinforced attitudes or decision-making at the individual. Moreover, it is even more challenging to determine how these individual effects translate into broader societal impacts[24,101,115,175,185,324,336]. On the other hand, the consequences of disinformation and misinformation are cumulative over time, which means that they may be modified by factors such as algorithmic or other types of change. However, the lack of longitudinal data and transparent information hinders our ability to address this issue. All of which make it difficult to assess the long-term effects. Alongside other factors, this is related with the existing fragmentation of knowledge about impact, which prevents a systemic view due to considerations that hinder its conceptual definition[24]:

- The scope of impact and its causes occur at many levels, at all stages of the chain, from the individual to society as a whole[324].
- It is possible to distinguish different areas of impact, which are often difficult to connect: the psychological, financial and social[285].
- There is low representativeness and comparability between the studies, which derive from the multiple channels and individual[324-326,337,338] geographic or linguistic factors that influence them[31,257,327].
- Impact studies often lack representation from mainstream journalism, including press, television, and radio, as well as other offline media sources[24]
- Gaps in knowledge about whether social networks are a driver or a threat for democratic systems are related to the difficulties and disagreements in understanding the impact of disinformation[5,71,94,101,322].

To date, there is very limited solid evidence about what causal influence disinformation and other dynamics have on offline attitudes and behaviours. This does not mean, however, that an influence does not exist or that it does not require a response, as existing evidence indicates to be the case[101,245,339]. To progress in this field, it is important to reinforce observational and relational evidence in the disinformation ecosystem and be able to validate the relationships detected with controlled-condition experiments that enable attribution of causality[327,340,341]. Experts highlight the need for a large-scale, multi-disciplinary approach[101], which includes the development and use of harmonized comparable indicators[24]. To achieve this, there should be strengthened collaboration, transparency, access to the data of social networks, digital platforms and mainstream journalism, which may require public policies and incentives[24,101,340-343].

Combatting disinformation: agents and mitigating strategies

The design of strategies to fight disinformation requires a combination of many instruments that can be inspired by the democratic principles of equality, representation and participation.

Experts note the importance of coordinating multiple interventions in response strategies against disinformation[5]. This response should also combine short-term responses aimed at the immediate effects of false information, and structural, long-term ones that seek to improve the resilience of the public, democratic systems and their institutions[10,11]. One type of response should not substitute another as they act in a complementary way, although negative synergies also exist[99]. As a whole, measures should act simultaneously in prevention, mitigation and at systems level. They also need to include the varied group of public and private actors who can apply them, and here the role of digital platforms in moderating is of utmost importance. The set of measures should also consider the public. UNESCO[99] groups strategies in accordance with the role of each actor, but there are other proposals, such as that of the EU's Joint Research Centre, which classify them depending on the desired effect using three democratic principles[5].

- **Equality**: Covers strategies aimed at reducing the asymmetries derived from a lack or accumulation of knowledge, information or data, power or assignment of responsibility of some actors compared to others[5,6]. They are mainly based on strengthening guarantor instruments and institutions by means of identification and neutralization, fostering social and individual resilience, and reinforcing regulations.

- **Representation**: This covers measures that focus on protecting electoral processes and the public's data, privacy and free choice.
- **Participation**: This aims to consolidate the role of the public, research and public or private actors, based on a forward-looking vision.

Guarantees, detection and neutralization

Guarantor institutions: responsibility as the first line of prevention

Strengthening trust in democratic institutions and reinforcing journalist mediation by promoting their capacities, resources, independence, transparency and plurality are foundational measures necessary to mitigate the effects of disinformation.

Democracies and their institutions can respond to and accept responsibility for the public's disaffection and mechanisms, such as disinformation, which can undermine democratic systems and increase the population's vulnerability[6,7,33,344,345]. Institutions face the structural challenge of generating a new dialogue with the public, adapted to the context of digital disintermediation, to manage associated uncertainty and build new trust relationships. On the one hand, institutional communication and the development of public policies designed with an understanding of the mechanisms of disinformation can help institutions mitigate the effects of false and misleading information[7,59,345–347]. In this sense strategic communication plays the very important role of enabling the generation of narratives that neutralize disinformation[347,348]. This requires monitoring and pre-emptive actions[349]. On the other hand, political actors or the media can also combat the effects of disinformation by avoiding exploitation of the vulnerabilities associated with social fragmentation, polarization or the erosion of trust in democracy[33,44,70,98,125,127,175].

There is also a broad consensus among experts and European organizations on the central role that journalists and information professionals can play as a structural check on disinformation[11,29,99,350]. However, there are few projects or scientific studies on the media as a firewall against disinformation[24,248,351] except for fact-checking agencies[99,352–354]. In fact, although there is a wealth of information about the behaviour and information diet on social networks and digital platforms, there has been very little research into offline information, mainstream media, or private spheres such as instant messaging.

While there is no clear consensus among the various proposals to strengthen the journalistic sector[39,352,355–357], the debate revolves around enhancing its capabilities and resources, independence, transparency, and plurality, as well as defining responsibility regarding disinformation and the role of advertising and the relationship with social networks[33,99,352–354]. Other proposals suggest new forms of communication that foster a connection with the public and promote a constructive, proactive view in the face of global difficulties[358–361]. All of which should occur without encroaching on freedom of speech. These strategic objectives require progress in the development of institutions and regulatory frameworks, like the European Digital Media Observatory (EDMO)[362,363] or the European Media Freedom Act[354]. This act aims to protect plurality and independence with a series of regulations that address areas ranging from stable funding of the media and transparency to mechanisms that protect against political or editorial interference and the imbalances derived from concentration of media ownership[364]. Experts note its importance in the medium to long term, although some aspects have caused dissent between the agents involved[365].

Monitoring and fact checking

Fact-checking agencies play an important, positive role in fighting false and misleading information. This role may be extensible to other actors and is not without its challenges .

Facts are also vulnerable since the truth is not the same as objectivity or accuracy[6]. Fact-checking agencies assess the accuracy of information to detect and rectify false information. The way to correct it is mainly based on exposing the facts surrounding inaccuracies[356,366–369] and providing accurate information when there is evidence[370–372] (known as debunking), on one hand, and providing context or available information when the content is not verifiable, on the other. Neutralizing false information can also consist of evaluating and showing the plausibility of sources and their credibility[115]. Agencies also amplify their range by means of collaboration with social networks and the media, which can aid ·progress along the road

· Strategic communication: This is a specialist focus on distributing and receiving information. It consists of communicating the right message through the right channels to the right people at the right time, using feedback from this process to remain focused on the established objectives.

to media literacy, the pre-emptive debunking (prebunking) of disinformation and even the promotion of public policies[368,373-375].

Scientific evidence indicates that contradicting false information has positive social effects[5], is effective to fight against disinformation and misinformation and, in most cases, is preferable to doing nothing[184,376-383]. However, it is not an infallible mechanism and conditions may exist that influence its effectiveness, among which are those derived from:

- Acceptance and scope of the refutation: These are two major challenges related to *why* and *how*[379]. Given information overdose and the constant increase in deepfakes, it is not always possible to fact check all information. So, a critical aspect is the use of clear, transparent methodological criteria that explain what content is prioritized for fact-checking[19,377,379,384]. Experts highlight that such criteria should consider the reliability of sources, the potential of the information to go viral, and potential harm it could cause[379]. On the other hand, rectification should meet standards of quality that boost its acceptance and reach while minimizing rejection[184,210,379,385-388].
- Personal beliefs and scepticism about agencies[184,376-378]: The standards and policies that guarantee trust, such as neutrality and methodological, financial and political transparency, etc. should be internationally defined and certified[389-391]. The activities of fact-checking agencies could also be helped and amplified with mechanisms aimed at inclusion and social listening[392], reinforcing trust and user participation[377,393-395], as well as making use of artificial intelligence[396-399].

Monitoring and refutation also form part of the strategic dimension in the fight against false and misleading information. They are not an action exclusively limited to fact-checkers; other agents from the field of journalism, experts and scientific institutions, etc. can participate to jointly warn the public about hostile operations to influence information[11]. On the other hand, monitoring is the cornerstone that enables the degree of pre-emption necessary to perform strategic communication or other types of structural measure connected with strengthening the public's resilience[11].

Automation: artificial intelligence as an ally

Artificial intelligence has the potential to amplify our capacity to detect disinformation and its agents, connecting it to general narratives and amplifying the reach of refutation.

Although artificial intelligence can foster disinformation (**Key point 2**) some studies have assessed the use of machine learning techniques to fight against it[31,288]. Scientific evidence shows its potential to detect and distinguish both false and misleading information and its agents on social networks, in accordance with the characteristics of the message itself or its context[31,288,400,401]. As well as text, this includes analysis of the images or videos that accompany it, the way in which they are shared or the emotions they provoke, among other aspects[31,257,400]. In addition to currently unveiling and distinguishing bots, trolls or disinformation itself, artificial intelligence can be the key to making the work of fact-checkers easier and identifying interconnections between information and large-scale narratives at international level[396-399]. It can also help amplify the reach of refutations[257].

In general terms, although applications exist, the techniques require further development before they can be used by the general public, public institutions or fact-checkers[288,399,402,403]. Among other challenges are fragmentation of the available data, possible bias, a lack of transparency and algorithmic explainability, the possible impact on privacy and ethical concerns[257,404].

Resilience and social capacity building

Since the battle against false information is being fought in the minds of people, the public needs to have the skills and mechanisms that reinforce their capacity to identify and reduce its influence.

Half the Spanish population does not trust their ability to identify false information[17]. There is a general consensus among experts on the central role that public awareness and capacity building play in reducing the impact of false and misleading information[4,21,405,406]. Education is the foundation of critical thought[19]. On the other hand, the subtlety of the disinformation phenomenon hampers a proper perception of the risk and, in doing so, to the awareness that would enable public policies and measures in this area[6].

·· Algorithmic explainability: This refers to the importance of being able to understand and explain decisions, with the support of artificial intelligence, that have an impact on the lives of people.

There are a wide range of actions that can be classified according to when they are applied:

- Before exposure:
    - In the long term or structurally, for instance with digital and media literacy, techno-cognitive actions or preparation of an ethical framework
    - Aimed at developing specific skills or fight specific types of false information, commonly known as ***prebunking***
- After exposure: These seek to modify the mistaken belief through refutation, and are called ***debunking***

Digital and media literacy

A better understanding of the many forms and dimensions of disinformation could prepare present and future generations against it.

This can combat disinformation at a structural level, as it covers the technical, cognitive, social, civil, ethical and creative skills and abilities the public can use to navigate the world of today's media in a more critical way when producing content, communicating and understanding the information they receive[19,407-411]. It fosters a critical perception of the different areas and challenges of information in the digital age. Although many different actions have been suggested, this is a complex task[19,407,408,410] and there are very few studies evaluating their effectiveness and applicability in different contexts and with different demographic groups[408,410,412-416]. Some recent evidence shows that learning to recognize specific indicators of false and misleading information[271,410] or recurrently highlighting the importance of attention and the accuracy of information[417,418] are effective strategies. Actions that generate an increase in generalized scepticism to all types of information should be avoided[419,420].

From a systemic perspective, experts in Spain advocate media education at all educational levels, which should also be aimed at teaching staff, communications professionals and vulnerable groups[19,121]. The current inclusion of these skills in today's educational syllabus may, however, be insufficient[19,421-425]. Teachers do not always have the information or resources to address this issue and, for instance, only one third of secondary school students correctly distinguish opinion from information[425]. There are many proposals and guidelines aimed at the formal education system[19,406,409,410,13,426-428] in addition to campaigns, guidelines and informal learning approaches[19,413,429-434]. Consolidating a coordinated systems approach with long-term objectives focused on the contexts of networks and issues that young people come into contact with could reinforce both personal and collective resilience. The aim should be to establish a favourable social framework for the analysis and debate of this type of challenge. Currently, the General Law on Audiovisual Communication 13/2022[435] only makes a passing reference to media literacy from the point of view of some experts who highlight that there is room for improvement in actions, plans and public policies in this field[70].

Ethical response and social norms

This consists of fostering an ethical framework to guide the behaviour of people, institutions and agents with a focus on a a

These seek to shape a shared social framework based on an understanding of the risks this threat entails, which appeals to a set of ethical norms and behaviours that, in the long term, carry more weight than a reactive response[12,33,99]. Information and media literacy can contribute to this. Another contribution would be actions from institutions and their associated agents to strengthen democratic debate and avoid its fragmentation. Such actions are in line with international standards related to human rights and ethical codes on behaviour when confronted with disinformation on social networks[99]. This is not only about identifying false or misleading information. A social and public attitude that reports and rejects disinformation can take mitigation to a collective level[33]. For individuals, there is evidence about how reputational harm can act as a check when it comes to sharing information and amplifying attention and accuracy regarding content[115].

Techno-cognitive actions and trust mechanisms

The flow of false information could be checked by redesigning the architectures of social networks and digital platforms, and with the development and application of mechanisms enabling evaluation of the truthfulness of information.

Based on our knowledge of behavioural economics, communication and computer science, there is evidence about how to redesign the way that social platforms are used to reduce the tendency to share or interact with false information, a concept known as nudging[5,173,379,436,437]. Implementation of mechanisms that promote thought and make it more difficult to automatically forward a message are an example[379,438]. Trust mechanisms are tools and methods that help users to better understand the information they see and its sources, enabling them to adjust their level of trust accordingly[99]. This can include labelling, for instance '*forwarded many times*' on some private messaging networks, access to trust indicators about content, its quality, its source etc., or information that contextualizes content. Another option is the use of applications, websites and platforms to scan the origin of information and its versions, which could include the use of artificial intelligence[99,438-440].

Digital platforms are not likely to accept this type of intervention unless there is governmental or public pressure to do so[379].

Refutation and behavioural interventions

Prebunking

Mechanisms to pre-emptively prepare the public for the arrival of false information so they can refute it when they come into contact with it.

The simplest actions range from presenting factually accurate information with the aim of preventing misinformation, to generic warnings about false and misleading information before it is spread[115]. According to scientific evidence[376,441-446], the most sophisticated and effective actions can be based on:

Addressing a specific subject for which false information content is explained and refuted before exposure to a real campaign[376]. This practice, commonly referred to as inoculation by experts, aims to build resistance.

Implementing logical reasoning mechanisms applicable to any subject to understand the techniques used to mislead and increase resistance. These mechanisms are based on identifying false experts, logical fallacies[447], impossible expectations, biased selection of evidence, and conspiracy theories[448].

Such actions strengthen our ability to identify misleading or false information during real campaigns. Inoculation requires deep strategic knowledge about the specific disinformation being addressed: how to present it, when to do so, etc.[184,376,449]. However, general escalation to the public and across multiple channels and themes is complex and requires further research[184]. Conversely, practices based on logical reasoning are more versatile, and promising developments, especially using games, exist that can be transferred to different contexts such as education or social networks[376,449-454].

Debunking: psychological resistance to rectification

Although the refutation of false information reduces the level of deception and erroneous beliefs, its effectiveness in terms of society and for all types of misleading content is limited[379]. Nevertheless, experts highlight its usefulness compared to not taking any action[115,379].

Rectification of false information should follow specific criteria to maximize its positive impact and have a long-term influence[115,385,455] thus reducing resistance to refutation or the backfire effect. Experts[379,456,457] propose practices based on repetition, empathy, the use of alternative explanations and generally trusted sources, as well as their selective application at the right time.

- Logical fallacies: Arguments that appear to be valid but are not. These include *ad hominem* arguments, arguments from ignorance and *argumentum ad populum*, argument from authority, the straw man fallacy or argument from anecdote among others, or techniques such as the burden of proof fallacy, the slippery slope argument or the false dilemma fallacy.
- Impossible expectations: This consists of the use of unachievable objectives or expectations to discredit or neutralize information. For instance, 'PCR tests for coronavirus are not 100% accurate, so we should not bother to use them'.

Progress in regulations

Establishing regulations about disinformation are complex as they may come into conflict with fundamental rights, there is no consensus on political mandate and there are difficulties determining the responsibility of social networks and digital platforms or the authorship and attribution of instigators.

Any of the proposed interventions, such as behavioral ones, run the risk of being instrumentalized by the very channels exploited by disinformation. Digital platforms and other actors may transfer the responsibility for detection and spreading disinformation to other external agents of the attention economy, users or other agents rather than taking responsibility themselves[5]. Regulatory steps may reinforce these steps and responsibilities[5,115].

Challenges of regulating

A large part of governmental and intergovernmental measures have focused on national security and institutionalization of the mission to detect, report and act against foreign disinformation campaigns[458]. This is the case for NATO[459], foreign policies and counter-intelligence[349], in the EU, the European Centre of Excellence for Countering Hybrid Threats[460], and actions at state level[18,22]. There is another focus aimed at increasing social resilience and the responsibilities of channels for disinformation[458], with implementation of strategies and regulations[5,169] within Spain and Europe.

However, this is a highly complex area, since conceptual understanding and knowledge of it are additional to other transversal challenges such as:

Conflict with fundamental rights: Freedom of speech[22,99,353,461,462] and information[463] should prevail, as should protection of democracy and its values[4,5,7]. Therefore, the criminalization or classification of information using the pretext that it is false should be avoided because it could undermine democracy and give discretion to the state[461].

Neutrality, transparency and political mandate: States do not have the authority to decide what is truthful or false information, neither are they always neutral[4,461]. Transparency and checks in the development of regulations[4], such as direct collaboration with civil society and the private sector[4,346,362,363], as well as international cooperation can broaden the legitimacy of actions[4]. These are not incompatible with government actions[4,464]. Each of them presents risks and advantages, which means that both approaches should be employed[4]. Moreover, political mandate is not universal, for instance, within the EU[142]. Some countries or institutions may not be equally committed to tackling this problem, which constitutes another challenge[142].

Responsibility: Experts highlight the progress made in both technical and human resources to enable attribution to instigators. On the other hand, they also indicate the need to consider the responsibility of digital platforms regarding the effects of their activities[7,33,169]. Defining what constitutes a means of communication affects its social, ethical and legal responsibilities. Some authors indicate that the classification of platforms as 'technologies' allows them to evade their transnational responsibility as mediators of information or even in terms of electoral moderation[465]. For instance, Google states that in the first half of 2023 it showed 20,441 political adverts in the EU, which generated a profit of 4.5 million euros, while it rejected 141,823 political adverts because they did not pass identity-checking processes[466]. Along these lines, various recent advances in regulation at European level make direct allusions to the role of digital platforms as moderators.

Considering these challenges, experts note the importance of avoiding regulation aimed at content and focusing on mechanisms that counteract the phenomenon since, to a large extent, it can be considered a consequence of freedom itself[4,37,467].

The European framework

European regulations focus on a co-regulatory framework and highlight the importance of detecting and counteracting false information, coordination and joint actions, mobilization of the private sector and social resilience.

The European framework has continually progressed in recent years (**Key point 3**). Even so, the effectiveness of the many instruments that comprise the European model has been called into question[37,468-470], which may be related to their evolution from proposals focused on the self-regulation of digital platforms and social networks, such as the Code of Practice on Disinformation, to co-regulatory measures such as the recent Digital Services Act (**Key point 3**).

Key point 3. Regulatory framework in the European framework.

(2015) the European External Action Service (EEAS) establishes its strategic communication division (Stratcom), which currently has four working groups, to monitor and discredit international disinformation affecting the European Union, coordinate EU response and collaborate with international partners[349].

(2018) Publication of the European approach to tackling online disinformation[28] and the report of the international group of experts that advises the EU[11]. These lead to the European Commission Action Plan against Disinformation[29] and the voluntary Code of Practice on Disinformation for the private sector (digital platforms, social networks and the advertising sector)[471].

(2019) Rapid Alert System (RAS)[473] to exchange information and coordinate an early response.

The Digital Services Act is the most recent legislation on disinformation. It aims to assign responsibility for disinformation, among other aspects, to large digital platforms.

The Digital Services Act is the most recent legislation on disinformation. It aims to assign responsibility for disinformation, among other aspects, to large digital platforms.

(2020) Various measures:

- **Communication**: 'Tackling COVID-19 disinformation - Getting the facts right'[474], which examines steps against the infodemic.
- **European Democracy Action Plan**[20], which addresses recommendations to revise and improve the Code of Practice on Disinformation and strengthen European foreign policies as well as attribution in this field[32].
- After a series of **assessments** indicating an insufficient impact and low application of the Code of Practice on Disinformation in 2020[469] and the strategy as a whole in 2021[468], the Code is strengthened in 2022[68,475].
- Creation of the **European Digital Media Observatory** (EDMO)[362] that brings together fact-checkers, academic researchers, digital platforms and social networks, mainstream journalism and professionals in media literacy. There are specific regional observatories; the one for Spain and Portugal is called IBERIFIER[363]. The joint mission is to improve knowledge on disinformation in Europe and consolidate advances that enable implementation of effective public policies.

(2023) The Spanish **Digital Services Act** aimed at defining the responsibilities of digital platforms and legally consolidating a good part of the Code of Practice on Disinformation[472,476]. The **Digital Markets Act** aims to guarantee a competitive, fair digital sector, enabling innovative digital businesses to grow and ensure the online security of users[477].

Foreign interference in European electoral processes is a subject of great concern for the EU, which has resulted in the Commission making recommendations to Member States[478]. In 2020, the European Parliament created a Special Committee for foreign interference in all democratic processes in the EU (ING 1)[479] that was recently renewed (ING2)[480].

Other steps are the proposal for **Regulation on the transparency and targeting of political advertising**[280] and the **European Media Freedom Act**[344]. The Act aims to support and protect the plurality and independence of mass media in the EU[354]. The proposal for regulations called the **AI Act** also mentions related matters, such as the obligation to identify deepfakes created with artificial intelligence[319].

European actions are grouped around four main lines: building capacities to detect, analyse and expose disinformation, creating the mechanisms for joint coordination and actions, mobilizing the private sector and, finally, social awareness and resilience[405]. Some of the measures connecting the proposals are:

- Demonetizing disinformation[68,469,471]
- Regulating propaganda and techniques such as micro-segmentation[4,7,9,280]
- Algorithmic transparency and tolerable practices[319,472]
- Identification of deepfakes for what they are[285,319]
- Developing tools, regulations and incentives systems to access and strengthen scientific progress and the processes of fact checking in the digital sphere[24,68]
- Media literacy and public participation (mechanisms to report content and its labelling, safer design of the service architecture, etc.)

In Europe, within states, there are examples[4,7] of co-regulatory frameworks as well as classic regulations, including financial sanctions, which make social networks responsible for false content and its elimination[7]. A case in point is France (partially reviewed by the *Conseil Constitutionnel*)[481] or Germany[482], among others. However, some experts suggest that this type of approach can have negative effects on freedom of speech or bring serious overload of the legal system[7,9].

Spain

As with other hybrid threats or cybersecurity, Spain is progressing in the consolidation of a cooperative framework that combines the public and private sectors, and civil society to develop regulations and measures against disinformation.

Experts highlight that the Spanish regulatory framework and current public policies are based on the European framework[70] (**Key point 3**), and this should continue to be the case[4]. Also in line with expert consensus, Spain is attempting to consolidate public-private cooperation and civil society in the fight against disinformation[18,22,54] (**Key point 4**).

Today, putting the Digital Services Act into practice requires identification of a body that would guarantee and monitor compliance. Mechanisms related to attribution of content in mainstream journalism (e.g., Article 30 of the Spanish Criminal Code)[46] may inspire the principle of the tiered subsidiary liability of digital platforms as information mediators.

Key point 4. The Spanish experience in the fight against disinformation.

- (2019) Spain's National Cybersecurity Strategy recognizes the danger posed by disinformation[483], an area that is assigned to the authorities and organizations that form part of the DSN (the Spanish Department of National Security). The Spanish National Intelligence Centre (CNI) monitors the agents connected with disinformation campaigns on a domestic level, and when there is foreign involvement, national law enforcement forces and agencies collaborate, each within their field. The National Cryptologic Centre (CCN) dependent on the CNI forms a disinformation unit.
- (2020) the DSN:
  - Publishes the procedure for action against disinformation[464], with the backing of the European Commission.
  - Creates the group of civil society experts who, alongside the representatives of public administrations, will jointly analyse the threat and possible strategies to fight against it using social, information, technological and regulatory means.
- (2021) Major update of the National Security Strategy of Spain includes the risks derived from disinformation campaigns and the challenges of managing the risks[22]; specific threats are covered in annual national security reports since COVID-19[54,484,485].
- (2022) National Forum against Disinformation Campaigns, where civil society, the private sector and public institutions cooperate in an advisory capacity in nine working groups that address all dimensions of the problem[346].
- (2023) The group of experts, the Forum and public institutions involved coincide on the need for a national strategy against disinformation and to work together on developing it[18]

The private sector, a necessary ally

Digital platforms have taken measures to mitigate disinformation, but not enough. The regulatory framework attempts to define their responsibilities in this area but also sees them as necessary allies in terms of co-regulation.

Large digital platforms and social networks have progressively implemented various types of mechanism and initiatives for moderation[19,24,59,67,99]. For instance, they periodically publish progress reports[465] on their implementation of the Code of Practice on Disinformation (**Key point 3**). Evidence indicates the importance and usefulness of including mechanisms such as fact-checking, rectification, and trust-building in the digital ecosystem. However, some studies also highlight the negative aspects of implementing these mechanisms[115,486].

Despite many initiatives, experts indicate the need for greater transparency and access to data for the research community to advance knowledge about disinformation and enable public policies[5,24,101,342]. The platforms and networks are key actors in the fight against this threat, and their cooperation and collaboration are necessary[18,67,475].

Privacy, security and elections

Many international initiatives exist, indicating the need to refine regulations regarding electoral content online, prioritizing the privacy and voting independence of citizens.

Political advertising during election periods is strictly regulated in the European Union in audiovisual and written media. However, most social networks are not covered by these measures[5]. Experts link the protection of electoral processes, equating them with critical infrastructures, with the implementation of advances in cybersecurity, regulation of online campaigns, data protection and the privacy of the public in an attempt to avoid potential abuse derived from personalized propaganda or other threats[2,5,7,10,228].

Regarding elections themselves, the EU has made recommendations for action (**Key point 3**) that Spain has followed, such as the creation of the Network for Coordination of Security in Electoral Processes[487]. The group of experts formed at the Spanish Department of National Security have made specific proposals of a strategic nature to strengthen democratic resilience against disinformation in the long term[67]. Specifically, experts highlight the need to tackle reform of Organic Law 5/1985 on the General Electoral System in order to avoid interference and guarantee a framework that minimizes the impact of disinformation[10,67]. Countries like Canada[488,489], the USA[490] or New Zealand[491] but also neighbours[492] such as France[493], Ireland[494] or the United Kingdom[495] have included or are working on reforms in their electoral regulations. Among other aspects, these aim to improve transparency regarding the content and distribution of political advertising online and offline or reinforce institutional communication, about which there is also a European proposal[280]. These initiatives can serve to guide Spain in this sphere[67].

A strategic, participative vision of the future

Key aspects to halt the advance of disinformation are developments in multi-disciplinary knowledge about the threat and consolidation of a national strategy in diverse dimensions.

To address the complexity of disinformation, experts highlight the importance of a systemic approach[59,405] that can develop into a National Strategy identifying the weaknesses, principles and objectives of the fight[18,22]. This approach should combine the foreign policies of states with a reinforcement of security and social resilience, connecting the different actors involved and putting them at the centre of defending democratic values[5]. This will require each of the actors to accept individual responsibility and build trust between them, whether they are the media, citizens, public institutions or political agents[5,18,33].

These goals will necessitate further development of knowledge about disinformation and mapping of its agents, which is, by definition, a multi-disciplinary task[18]. It is important to move forward with large-scale studies and combined assessment of the various mitigation actions in order to design effective strategies, and to use prospective research to plan long-term tactics[5,24,101,115]. Another element that could foster integration and representation within institutions and public policies is to professionalize the sector.

Likewise, cooperation must go beyond the private sphere and reach the whole of civil society, in line with the National Forum on disinformation (**Key point 4**), and cover an international context[18,70]. Given the sensitivity of this phenomenon, key considerations should be the transparency of actions undertaken by all actors, whether public or private, in addition to public accountability[5]. Bearing in mind the rapid, continuous developments in this field, we need adaptable, dynamic, up-to-date tools[70]. This could benefit from the development of new platforms for public debate and media literacy, not only to increase resilience to false information but also to strengthen the social framework. This would generate strong foundations, forming a cornerstone for strategies whose mission is to defend democracy, like the fight against disinformation[5].

Key concepts

- The digital age facilitates an unprecedented amplification of disinformation and other information disorders, which constitute a major risk for democracies.
- Handling disinformation is a challenge because it implies protecting and expanding citizen rigjts without restricting others such as freedom of expression.
- The success of a disinformation campaign does not necessarily reside in generating false beliefs, but in the creation of confusion, distrust and division, alongside its amplification of bias and prejudice. To achieve their aims, its instigators usually exploit affective components and replace truthfulness with verisimilitude. The goal is to achieve structural changes in the public's perception rather than specific short-term effects.
- Disinformation in the digital age is favored by an environment where traditional intermediation and information flow blur: anyone can generate content, disseminate it, and share it. This results in an information explosion of varying quality that hinders the identification of truthful content, creating uncertainty
- Disinformation is explained within a socio-political context where the crisis of democratic trust, geopolitical situations, social and psychological factors, and the digital business model itself, supported by opaque and constantly evolving technologies, play a very significant role
- Although consensus exists about the risks and need to start up mechanisms to combat disinformation, the complexity of this phenomenon makes its comprehensive analysis difficult.
- Calls are being made on the responsibility and cooperation of all agents (politicians, the media, business) to prevent the exploitation of uncertainty and false and misleading information.
- Democratic institutions and their guarantors should foster a dialogue with the public that reinforces trust and is tailored to the new informational context.
- Measures to combat disinformation have the ultimate goal of achieving digital and media literacy, as well as making society as a whole more resilient.
- The European framework promotes steps aimed at defending and strengthening democracy against disinformation, and at consolidating mechanisms that fight it in a systemic way. This ranges from attribution of responsibilities or demonetizing content, to extending the freedom and plurality of the media or the moderation of online electoral content.
- New developments in artificial intelligence could represent a turning point for disinformation. Although this technology increases the scope and danger of the threat, it also offers new opportunities to detect and combat false and misleading information.

Bibliography:

1. Nai Fovino, I. *et al.* Cybersecurity, our digital anchor. EUR 30276 EN, Publications Office of the European Union. Luxemburgo. (2020)
www.doi.org/10.2760/352218.

2. Oficina de Ciencia y Tecnología del Congreso de los Diputados (Oficina C). Informe C: Ciberseguridad. (2022)
www.doi.org/10.57952/c8hy-6c31.

3. Marciel, R. Democracia, desinformación y conocimiento político: algunas aclaraciones conceptuales. *Dilemata* 65–82 (2022).

4. Grupo de expertos de la sociedad civil – Departamento de Seguridad Nacional. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil.* Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática
https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf (2022).

5. Lewandowsky, S. *et al.* Technology and Democracy: Understanding the influence of online technologies on political behaviour and decision-making. *JRC Publications Repository*
https://publications.jrc.ec.europa.eu/repository/handle/JRC122023 [30/05/2023]
www.doi.org/10.2760/709177.

6. Innerarity, D. & Colomina, C. La verdad en las democracias algorítmicas - Truth in algorithmic democracies. *Revista CIDOB d'Afers Internacionals* 11–24 (2020).

7. Bayer, J., Bitiukova, N., Bard, P., Szakács, J., Alemanno, A. & Uszkiewicz, E. Disinformation and Propaganda – Impact on the Functioning of the Rule of Law in the EU and its Member States. *SSRN Electronic Journal* (2019)
www.doi.org/10.2139/ssrn.3409279.

8. Panel for the Future of Science and Technology (STOA). *Automated tackling of disinformation: major challenges ahead.*
https://data.europa.eu/doi/10.2861/368879 (2019).

9. Colomina, C., Margalef, H. S. & Youngs, R. *The impact of disinformation on democratic processes and human rights in the world.* (2021).

10. Rubio Núñez, R. La amenaza tecnológica en los procesos electorales: Una respuesta jurídica. *Revista de privacidad y derecho digital* **3**, 109–146 (2018).

11. High Level Expert Group on Fake News and Online Disinformation. *A multi-dimensional approach to disinformation.*
https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation (2018).

12. Wagner, A. Retos filosóficos de las sociedades digitales: esbozo de un enfoque sistémico. *Dilemata* **38**, 13–29 (2022)
https://www.dilemata.net/revista/index.php/dilemata/article/view/412000497.

13. van Boheemen, P., Munnichs, G. & Dujso, E. *Digital threats to democracy. On new technology and disinformation.*
https://www.rathenau.nl/en/digitalisering/digital-threats-democracy (2020).

14. Tennøe, T. & Barland, M. *Elections, technology and political influencing*. Norwegian Board of Technology (NBT) (2019).

15. Bruns, H., Dessart, F. J. & Pantazi, M. *Covid-19 misinformation: Preparing for future crises.* Joint Research Center
https://publications.jrc.ec.europa.eu/repository/handle/JRC130111 (2022)
www.doi.org/10.2760/41905.

16. Del-Fresno-García, M. Desórdenes informativos: sobreexpuestos e infrainformados en la era de la posverdad. *Profesional de la información / Information Professional* **28**, (2019)
www.doi.org/10.3145/epi.2019.may.02.

17. Eurobarometer. *Fake news and disinformation online.*
https://europa.eu/eurobarometer/surveys/detail/2183 (2018).

18. Grupo de expertos de la sociedad civil – Departamento de Seguridad Nacional. Principios para una estrategia contra la desinformación. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática.*
https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf (2022).

19. Grupo de expertos de la sociedad civil – Departamento de Seguridad Nacional. La alfabetización mediática, herramienta clave en la lucha contra la desinformación. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática.*
https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf (2022).

20. *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS On the European democracy action plan*. (2020).

21. Centro Criptológico Nacional. Ministerio de Defensa. Gobierno de España. *Buenas Prácticas en la Desinformación en el Ciberespacio.* CCN-CERT BP/13
https://www.ccn-cert.cni.es/es/informes/informes-de-buenas-practicas-bp/3549-ccn-cert-bp-13-desinformacion-en-el-ciberespacio/file?format=html (2021).

22. Departamento de Seguridad Nacional. Estrategia de Seguridad Nacional 2021.
https://www.dsn.gob.es/es/documento/estrategia-seguridad-nacional-2021 [18/04/2022].

23. Chong, M. & Choy, M. An Empirically Supported Taxonomy of Misinformation: *Advances in Media, Entertainment, and the Arts* (eds. Dalkir, K. & Katz, R.) 117–138 (IGI Global, 2020). ISBN: 978-1-79982-543-2.

24. Badillo-Matos, A. *et al. Analysis of the Impact of Disinformation on Political, Economic, Social and Security Issues, Governance Models and Good Practices: The cases of Spain and Portugal.* IBERIFIER Report. DOI: www.doi.org/10.15581/026.002
https://iberifier.eu/2023/06/21/report-analysis-impact-disinformation-june-2023/ (2023).

25. Wardle, C. & Derakhshan, H. *Information Disorder. Toward an interdisciplinary framework for research and policymaking.* Council of Europe (2017).

26. Bennett, W. L. & Livingston, S. The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication* **33**, 122–139 (2018)
www.doi.org/10.1177/0267323118760317.

27. House of Commons Digital, Culture, Media and Sport Committee. UK Paliament. Disinformation and 'fake news'. **Eighth Report of Session 2017–19**, (2019).

28. *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Tackling online disinformation: a European Approach.* (2018).

29. European Commission. Action Plan against Disinformation.
https://digital-strategy.ec.europa.eu/en/library/action-plan-against-disinformation [05/06/2023].

30. Borges do Nascimento, I. J., Pizarro, A. B., Almeida, J. M., Azzopardi-Muscat, N., Gonçalves, M. A., Björklund, M. & Novillo-Ortiz, D. Infodemics and health misinformation: a systematic review of reviews. *Bulletin of the World Health Organization* **100**, 544–561 (2022)
www.doi.org/10.2471/BLT.21.287654.

31. Ruffo, G., Semeraro, A., Giachanou, A. & Rosso, P. Studying fake news spreading, polarisation dynamics, and manipulation by bots: A tale of networks and language. *Computer Science Review* **47**, (2023)
www.doi.org/10.1016/j.cosrev.2022.100531.

32. Comisión Europea. Plan de Acción para la Democracia Europea.
https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan_es [23/06/2023].

33. Wagner, A. Deliberación, polarización y posverdad. Repensar la responsabilidad en la sociedad digital. *Quaderns de Filosofia* **10**, 51–67 (2023)
www.doi.org/10.7203/qfia.10.2.26616.

34. Salaverría-Aliaga, R. *Informe del GTM: Entender y combatir la desinformación sobre ciencia y salud.*
http://hdl.handle.net/10261/239480 (2021).

35. Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M. & Procter, R. Detection and Resolution of Rumours in Social Media: A Survey. *ACM Computing Surveys* **51**, 32:1-32:36 (2018)
www.doi.org/10.1145/3161603.

36. Zubiaga, A., Liakata, M., Procter, R., Hoi, G. W. S. & Tolmie, P. Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. *PLOS ONE* **11**, e0150989 (2016)
www.doi.org/10.1371/journal.pone.0150989.

37. Corredoira y Alfonso, L. European Regulatory Responses to Disinformation. Special Attention to Election Campaigns. *Derecom* 5 (2020).

38. *Willemo J., (2019). Trends and Developments in the Malicious Use of Social Media. Riga: NATO Strategic Communications Centre of Excellence.*
https://stratcomcoe.org/pdfjs/?file=/publications/download/nato_report_-_trends_and_developments.pdf?zoom=page-fit .

39. IBERIFIER Reports. *The Impact of Disinformation on the Media Industry in Spain and Portugal.*
https://iberifier.eu/2023/02/15/iberifier-reports-the-impact-of-disinformation-on-the-media-industry-in-spain-and-portugal/ (2023).

40. Rinehart, A. Fake news. It's complicated. *First Draft*
https://firstdraftnews.org/articles/fake-news-complicated/ [22/06/2023].

41. Starbird, K. Disinformation's spread: bots, trolls and all of us. *Nature* **571**, 449–450 (2019).

42. Matos, Á. B. La sociedad de la desinformación: propaganda, «fake news» y la nueva geopolítica de la información. *Real Institututo Elcano*
https://www.realinstitutoelcano.org/documento-de-trabajo/la-sociedad-de-la-desinformacion-propaganda-fake-news-y-la-nueva-geopolitica-de-la-informacion/ [30/08/2023].

43. Martens, B., Aguiar, L., GGmez, E. & Mueller-Langer, F. The Digital Transformation of News Media and the Rise of Disinformation and Fake News. *SSRN Electronic Journal* (2018)
www.doi.org/10.2139/ssrn.3164170.

44. Ognyanova, K., Lazer, D., Robertson, R. E. & Wilson, C. Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. *Harvard Kennedy School Misinformation Review* (2020)
www.doi.org/10.37016/mr-2020-024.

45. Bobadilla, Á. M. & Isidoro, B. del C. M. *Fake News, desinformación y otros desórdenes informativos.* (Editorial Fragua, 2022). ISBN: 978-84-7074-963-6.

46. Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal.

47. Fiscalía General del Estado. *Tratamiento penal de las 'fake news'*. (2020).

48. Secretaría de Estado de Seguridad. Ministerio del Interior. *Informe sobre la evolución de los delitos de odio en España 2021.*

https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/publicaciones-periodicas/informe-sobre-la-evolucion-de-los-delitos-de-odio-en-Espana/Informe_evolucion_delitos_odio_Espana_2021_126200207.pdf (2021).

49. Salaverría, R. & León, B. Misinformation Beyond the Media: 'Fake News' in the Big Data Ecosystem. *Total Journalism* (eds. Vázquez-Herrero, J., Silva-Rodríguez, A., Negreira-Rey, M.-C., Toural-Bran, C. & López-García, X.) vol. 97 109–121 (Springer International Publishing, 2022). ISBN: 978-3-030-88027-9.

50. Ricard, J. & Medeiros, J. Using misinformation as a political weapon: COVID-19 and Bolsonaro in Brazil. *Harvard Kennedy School Misinformation Review* **1**, (2020) www.doi.org/10.37016/mr-2020-013.

51. Pantti, I. K., Mervi. Fake News: The narrative battle over the Ukrainian conflict. *The Future of Journalism: Risks, Threats and Opportunities* (Routledge, 2019). ISBN: 978-0-429-46203-0.

52. Levinger, M. Master Narratives of Disinformation Campaigns. *Journal of International Affairs* **71**, 125–134 (2018).

53. EUvsDisinfo. Actualización del informe especial del SEAE: evaluación de las narrativas y la desinformación en torno a la pandemia de COVID-19 (actualizada del 23 de abril al 18 de mayo). *EU vs Disinfo* https://euvsdisinfo.eu/es/actualizacion-del-informe-especial-del-seae-breve-evaluacion-de-las-narrativas-y-la-desinformacion-en-torno-a-la-pandemia-de-covid-19-actualizada-del-23-de-abril-al-18-de-mayo/ [19/09/2023].

54. Presidencia del Gobierno. *Informe Anual de Seguridad Nacional 2022*. https://www.dsn.gob.es/es/documento/informe-anual-seguridad-nacional-2022 (2023).

55. Tandoc, E. C., Lim, Z. W. & Ling, R. Defining "Fake News". *Digital Journalism* **6**, 137–153 (2018) www.doi.org/10.1080/21670811.2017.1360143.

56. McCright, A. M. & Dunlap, R. E. Combatting misinformation requires recognizing its types and the factors that facilitate its spread and resonance. *Journal of Applied Research in Memory and Cognition* **6**, 389–396 (2017) www.doi.org/10.1016/j.jarmac.2017.09.005.

57. Giachanou, A., Ghanem, B. & Rosso, P. Detection of conspiracy propagators using psycho-linguistic characteristics. *Journal of Information Science* **49**, 3–17 (2023) www.doi.org/10.1177/0165551520985486.

58. *Michlin-Shapir, V. (2021). The Long Decade of Disinformation. Defence Strategic Communications, 9, 17-33. doi: 10.30966/2018.RIGA.9.5.* https://stratcomcoe.org/pdfjs/?file=/publications/download/web_shapir_dsc_vol9-1.pdf?zoom=page-fit .

59. Grupo de expertos de la sociedad civil - Departamento de Seguridad Nacional. La desinformación: una amenaza a la democracia. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf* (2022).

60. *Giannopoulos, G., Smith, H., Theocharidou, M. The landscape of hybrid threats: a conceptual model. The Landscape of Hybrid Threats: A conceptual model.* EUR 30585; JRC123305. ISBN 978-92-76-29819-9, doi:10.2760/44985, https://data.europa.eu/doi/10.2760/44985 (2021).

61. Jungwirth, R. *et al. Hybrid Threats: A Comprehensive Resilience Ecosystem*. Joint Research Centre. ISBN 978-92-76-53293-4, doi:10.2760/867072, JRC129019. https://publications.jrc.ec.europa.eu/repository/handle/JRC129019 (2023) www.doi.org/10.2760/37899.

62. Pamment, J. *The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework*. Carnegie Endowment for International Peace. https://carnegieendowment.org/2020/09/24/eu-s-role-in-fighting-disinformation-crafting-disinformation-framework-pub-82720 (2020).

63. Johns Hopkins University & Imperial College London. Countering cognitive warfare: awareness and resilience. *NATO Review* https://www.nato.int/docu/review/articles/2021/05/20/countering-cognitive-warfare-awareness-and-resilience/index.html [28/09/2023].

64. Bradshaw, S. & Howard, P. The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation. *Copyright, Fair Use, Scholarly Communication, etc.* (2019).

65. Benedicto Solsona, M. Á. La UE frente a la desinformación de China y Rusia durante la COVID-19. La necesidad de una mayor proactividad narrativa europea a nivel internacional. *Janus.net* **11**, 84–98 www.doi.org/10.26619/1647-7251.DT21.6.

66. Suau, J. & Puertas-Graell, D. Disinformation narratives in Spain: reach, impact and spreading patterns. *Profesional de la información* **32**, (2023) www.doi.org/10.3145/epi.2023.sep.08.

67. Grupo de expertos de la sociedad civil - Departamento de Seguridad Nacional. Propuesta para combatir las campañas de desinformación en proceso electoral. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf.*

68. European Commission. *Strengthened Code of Practice on Disinformation.* https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation (2022).

69. Hameleers, M. Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination. *Communication Theory* **33**, (2023) www.doi.org/10.1093/ct/qtac021.

70. IBERIFIER Reports. *Political and Legal Aspects of Disinformation in Portugal and Spain.* https://iberifier.eu/2023/10/20/iberifier-reports-legal-and-political-aspects-of-disinformation-in-portugal-and-spain-october-2023/ (2023).

71. Jeangène vILMER, J. B., Escorcia, A., Guillaume, M. & Herrera, J. *Information manipulation: A challenge for our democracies. By the Policy Planning Staff (CAPS) of the Ministry for Europe and Foreign Affairs and the Institute for Strategic Research (IRSEM) of the Ministry for the Armed Forces, Paris, August 2018.* (2018).

72. Posetti, J. & Matthews, A. *A short guide to the history of 'fake news' and disinformation.* International Center for Journalists https://www.icfj.org/sites/default/files/2018-07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ%20Final.pdf (2018).

73. Jamieson, K. H. *Cyberwar: How Russian Hackers and Trolls Helped Elect a President: What We Don't, Can't, and Do Know.* (Oxford University Press, 2020). ISBN: 978-0-19-752895-2.

74. Lemke, T. & Habegger, M. W. Foreign Interference and Social Media Networks: A Relational Approach to Studying Contemporary Russian Disinformation. *Journal of Global Security Studies* **7**, ogac004 (2022) www.doi.org/10.1093/jogss/ogac004.

75. Bastos, M. T. & Mercea, D. The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review* **37**, 38–54 (2019) www.doi.org/10.1177/0894439317734157.

76. Greene, C. M., Nash, R. A. & Murphy, G. Misremembering Brexit: partisan bias and individual predictors of false memories for fake news stories among Brexit voters. *Memory* **29**, 587–604 (2021) www.doi.org/10.1080/09658211.2021.1923754.

77. Gerken, T. EU gives Meta and TikTok formal Hamas disinformation deadline. *BBC News* https://www.bbc.com/news/technology-67157733 [21/10/2023].

78. Informe EDMO fact-checking network #29 | Iberifier. https://iberifier.eu/2023/11/22/informe-edmo-numero-29/ [23/11/2023].

79. García-Saisó, S. *et al.* Infodemia en tiempos de COVID-19. *Revista Panamericana de Salud Pública* **45**, e89 (2021) www.doi.org/10.26633/RPSP.2021.89.

80. Siwakoti, S., Yadav, K., Thange, I., Bariletto, N., Zanotti, L., Ghoneim, A. & Shapiro, J. N. *Localized Misinformation in a Global Pandemic: Report on COVID-19 Narratives around the World.* Empirical Study of Conflict, Princeton University, pages 1-68, https://esoc.princeton.edu/publications/localized-misinformation-global-pandemic-report-covid-19-narratives-around-world (2021).

81. Zou, S. Mistranslation as disinformation: COVID-19, global imaginaries, and self-serving cosmopolitanism. *The Cultural Politics of COVID-19* (Routledge, 2022). ISBN: 978-1-00-331041-9.

82. Clemente-Suárez, V. J. *et al.* Mis–Dis Information in COVID-19 Health Crisis: A Narrative Review. *International Journal of Environmental Research and Public Health* **19**, 5321 (2022) www.doi.org/10.3390/ijerph19095321.

83. Lee, S. K., Sun, J., Jang, S. & Connelly, S. Misinformation of COVID-19 vaccines and vaccine hesitancy. *Scientific Reports* **12**, 13681 (2022) www.doi.org/10.1038/s41598-022-17430-6.

84. León, B., Martínez-Costa, M.-P., Salaverría, R. & López-Goñi, I. Health and science-related disinformation on COVID-19: A content analysis of hoaxes identified by fact-checkers in Spain. *PLOS ONE* **17**, e0265995 (2022) www.doi.org/10.1371/journal.pone.0265995.

85. Theocharis, Y. *et al.* Does the platform matter? Social media and COVID-19 conspiracy theory beliefs in 17 countries. *New Media and Society* (2021) www.doi.org/10.1177/14614448211045666.

86. IBERIFIER Reports & Fundación Española para la Ciencia y la Tecnología. *Desinformación científica en España.* https://www.fecyt.es/es/publicacion/desinformacion-cientifica-en-espana (2022).

87. Cano Orón, L., Calvo, D., López García, G. & Baviera, T. Disinformation in Facebook Ads in the 2019 Spanish General Election Campaigns. (2021).

88. Paniagua Rojano, F., Seoane Pérez, F. & Magallón Rosa, R. Anatomía del bulo electoral: la desinformación política durante la campaña del 28-A en España. *Revista CIDOB d'Afers Internacionals* **124**, 123–145 (2020) www.doi.org/doi.org/10.24241/rcai.2020.124.1.123.

89. IBERIFIER Reports. *Spain & Portugal fact-checking brief.* Q1 2023 https://iberifier.eu/2023/05/26/fact-checking-brief-q1-2023/ (2023).

90. Aparici, R., García-Marín, D. & Rincón-Manzano, L. Noticias falsas, bulos y trending topics. Anatomía y estrategias de la desinformación en el conflicto catalán. *Profesional de la información* **28**, (2019) www.doi.org/10.3145/epi.2019.may.13.

91. Vicente, A. R. *Disinformation landscape in Spain.* EU DesinfoLab https://www.disinfo.eu/publications/disinformation-landscape-in-spain/ (2023).

92. Rhodes, S. C. Filter Bubbles, Echo Chambers, and Fake News: How Social Media Conditions Individuals to Be Less Critical of Political Misinformation. *Political Communication* **39**, 1–22 (2022) www.doi.org/10.1080/10584609.2021.1910887.

93. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018) www.doi.org/10.1126/science.aap9559.

94. Rubio Núñez, R. Los efectos de la posverdad en la democracia. *Revista de derecho político* 191–228 (2018) www.doi.org/10.5944/rdp.103.2018.23201.

95. Newman, N., Fletcher, R., Eddy, K., Robertson, C. T. & Nielsen, R. K. *Reuters Institute Digital News Report 2023.*

https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf (2023).

96. Amoedo-Casais, A., Moreno-Moreno, E., Negredo-Bruna, S., Kaufmann-Argueta, J. & Vara-Miguel, A. Digital News Report España 2023. (2023) www.doi.org/10.15581/019.2023.

97. Bradshaw, S., Bailey, H. & Howard, P. N. *Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation.* Oxford University, UK: Programme on Democracy&Technology https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/02/CyberTroop-Report20-Draft9.pdf .

98. Lasser, J., Aroyehun, S. T., Carrella, F., Simchon, A., Garcia, D. & Lewandowsky, S. From alternative conceptions of honesty to alternative facts in communications by US politicians. *Nature Human Behaviour* 1–12 (2023) www.doi.org/10.1038/s41562-023-01691-w.

99. UNESCO. *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression.* https://en.unesco.org/publications/balanceact (2020).

100. IBERIFIER. Iberian Digital Media Map. https://map.iberifier.eu/ [28/09/2023].

101. González-Bailón, S. & Lelkes, Y. Do social media undermine social cohesion? A critical review. *Social Issues and Policy Review* **17**, 155–180 (2023) www.doi.org/10.1111/sipr.12091.

102. Bode, L. & Vraga, E. K. In Related News, That was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media. *Journal of Communication* **65**, 619–638 (2015) www.doi.org/10.1111/jcom.12166.

103. *Mair D., Smillie L., La Placa G., Schwendinger F., Raykovska M., Pasztor Z., van Bavel R., Understanding our political nature: How to put knowledge and reason at the heart of political decision-making, EUR 29783 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-08621-5, doi:10.2760/374191, JRC117161.* https://policycommons.net/artifacts/2162869/understanding-our-political-nature/2918428/ (2019).

104. Innerarity, D. & Colomina, C. Introducción: desinformación y poder, la crisis de los intermediarios. *Revista CIDOB d'afers internacionals* 7–10 (2020) www.doi.org/doi.org/10.24241/rcai.2020.124.1.7.

105. Törnberg, P. Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLOS ONE* **13**, e0203958 (2018) www.doi.org/10.1371/journal.pone.0203958.

106. Torreblanca, J.-ıgnacıo. Social Networks and Democracy: Problems and Dilemmas of Regulating the Digital Ecosystem. *Siyasal: Journal of Political Sciences* **32**, 15–33 (2023) www.doi.org/10.26650/siyasal.2023.32.1252061.

107. Xaudiera, S. & Cardenal, A. S. Ibuprofen narratives in five European countries during the COVID-19 pandemic. *Harvard Kennedy School Misinformation Review* **1**, (2020) www.doi.org/10.37016/mr-2020-029.

108. Benkler, Y., Faris, R. & Roberts, H. The Propaganda Feedback Loop. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (eds. Benkler, Y., Faris, R. & Roberts, H.) 0 (Oxford University Press, 2018). ISBN: 978-0-19-092362-4.

109. Benkler, Y., Faris, R. & Roberts, H. Mainstream Media Failure Modes and Self-Healing in a Propaganda-Rich Environment. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (eds. Benkler, Y., Faris, R. & Roberts, H.) 0 (Oxford University Press, 2018). ISBN: 978-0-19-092362-4.

110. Heiberger, R., Majó-Vázquez, S., Castro Herrero, L., Nielsen, R. K. & Esser, F. Do Not Blame the Media! The Role of Politicians and Parties in Fragmenting Online Political Debate. *International Journal of Press/Politics* **27**, 910–941 (2022) www.doi.org/10.1177/19401612211015122.

111. Rodríguez-Virgili, J., Serrano-Puche, J. & Fernández, C. B. Digital Disinformation and Preventive Actions: Perceptions of Users from Argentina, Chile, and Spain. *Media and Communication* **9**, 323–337 (2021) www.doi.org/10.17645/mac.v9i1.3521.

112. Rodríguez-Pérez, C., Seibt, T., Magallón-Rosa, R., Paniagua-Rojano, F. J. & Chacón-Peinado, S. Purposes, Principles, and Difficulties of Fact-checking in Ibero-America: Journalists' Perceptions. *Journalism Practice* **0**, 1–19 (2022) www.doi.org/10.1080/17512786.2022.2124434.

113. Abonizio, H. Q., de Morais, J. I., Tavares, G. M. & Barbon Junior, S. Language-Independent Fake News Detection: English, Portuguese, and Spanish Mutual Features. *Future Internet* **12**, 87 (2020) www.doi.org/10.3390/fi12050087.

114. Unkelbach, C., Koch, A., Silva, R. R. & Garcia-Marques, T. Truth by Repetition: Explanations and Implications. *Current Directions in Psychological Science* **28**, 247–253 (2019) www.doi.org/10.1177/0963721419827854.

115. Ecker, U. K. H. *et al.* The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology* **1**, 13–29 (2022) www.doi.org/10.1038/s44159-021-00006-y.

116. Fazio, L. K. Repetition Increases Perceived Truth Even for Known Falsehoods. *Collabra: Psychology* **6**, 38 (2020) www.doi.org/10.1525/collabra.347.

117. Almansa-Martínez, A., Fernández-Torres, M. J. & Rodríguez-Fernández, L. Desinformación en España un año después de la COVID-19. Análisis de las verificaciones de Newtral y Maldita. *Revista Latina de Comunicación Social* 183–200 (2022) www.doi.org/10.4185/RLCS-2022-1538.

118. Magallón-Rosa, R. The Agenda Below the Radar. Disinformation and Fact-Checking on (IM) Migration. *Migraciones* **52**, 59–87 (2021) www.doi.org/10.14422/mig.i52.y2021.003.

119. Elsevier. 'Fake news', bulos y contenidos en salud: una tendencia con muchos riesgos. *Elsevier Connect* https://www.elsevier.com/es-es/connect/actualidad-sanitaria/fake-news-bulos-y-contenidos-en-salud-una-tendencia-con-muchos-riesgos [23/10/2023].

120. European Commison. €1.2 million for new project to deepen understanding of disinformation on war, elections and gender | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/news/eu12-million-new-project-deepen-understanding-disinformation-war-elections-and-gender [23/11/2023].

121. Moore, R. C. & Hancock, J. T. A digital media literacy intervention for older adults improves resilience to fake news. *Scientific Reports* **12**, 6008 (2022) www.doi.org/10.1038/s41598-022-08437-0.

122. Blanco-Herrero, D., Amores, J. J. & Sánchez-Holgado, P. Citizen Perceptions of Fake News in Spain: Socioeconomic, Demographic, and Ideological Differences. *Publications* **9**, 35 (2021) www.doi.org/10.3390/publications9030035.

123. Serrano Maíllo, M. I. ¿Qué hace que seamos tan vulnerables a la desinformación?: ¿estamos perdidos o aún podemos hacer algo? *Fake News, desinformación y otros desórdenes informativos, 2022, ISBN 9788470749636, págs. 135-149* 135–149 (Fragua, 2022).

124. Repucci, S. & Slipowitz, A. *Democracy under Siege.* Freedom House https://freedomhouse.org/report/freedom-world/2021/democracy-under-siege .

125. Gerschewski, J. Erosion or decay? Conceptualizing causes and mechanisms of democratic regression. *Democratization* **28**, 43–62 (2021) www.doi.org/10.1080/13510347.2020.1826935.

126. A new low for global democracy. *The Economist* https://www.economist.com/graphic-detail/2022/02/09/a-new-low-for-global-democracy [29/09/2023].

127. Humprecht, E. The Role of Trust and Attitudes toward Democracy in the Dissemination of Disinformation—a Comparative Analysis of Six Democracies. *Digital Journalism* **0**, 1–18 (2023) www.doi.org/10.1080/21670811.2023.2200196.

128. The deep roots of polarisation, or on the need to recover the lost narrative. *CaixaBank Research* https://www.caixabankresearch.com/en/economics-markets/public-sector/deep-roots-polarisation-or-need-recover-lost-narrative [23/11/2023].

129. OECD. The territorial impact of COVID-19: Managing the crisis across levels of government. *OECD* https://www.oecd.org/coronavirus/policy-responses/the-territorial-impact-of-covid-19-managing-the-crisis-across-levels-of-government-d3e314e1/ [25/10/2023].

130. Eurofound. *Maintaining trust during the COVID-19 pandemic.* https://data.europa.eu/doi/10.2806/707872 (2022).

131. Roccato, M., Colloca, P., Cavazza, N. & Russo, S. Coping with the COVID-19 pandemic through institutional trust: Rally effects, compensatory control, and emotions. *Social Science Quarterly* **102**, 2360–2367 (2021) www.doi.org/10.1111/ssqu.13002.

132. University of Oxford. *Social media manipulation by political actors now an industrial scale problem prevalent in over 80 countries – annual Oxford report.* https://www.oii.ox.ac.uk/news-events/news/social-media-manipulation-by-political-actors-now-an-industrial-scale-problem-prevalent-in-over-80-countries-annual-oxford-report (2021).

133. Comisión Europea. *Comunicación conjunta al Parlamento Europeo, al Consejo Europeo y al Consejo. Aumentar la resiliencia y desarrollar las capacidades para hacer frente a las amenazas híbridas.* https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=CELEX:52018JC0016 (2018).

134. Comisión Europea. *Comunicación conjunta al Parlamento Europeo y al Consejo. Comunicación conjunta sobre la lucha contra las amenazas híbridas Una respuesta de la Unión Europea.* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016JC0018 (2016).

135. EUvsDisinfo. EEAS SPECIAL REPORT UPDATE: Short Assessment of Narratives and Disinformation around the COVID-19/Coronavirus Pandemic (Updated 2 – 22 April). *EUvsDisinfo* https://euvsdisinfo.eu/eeas-special-report-update-2-22-april/ [26/06/2023].

136. Council of the EU. 11th package of sanctions on Russia's war of aggression against Ukraine: additional 71 individuals and 33 entities included in the EU's sanctions list and new tools to counter circumvention and information warfare. https://www.consilium.europa.eu/en/press/press-releases/2023/06/23/11th-package-of-sanctions-on-russia-s-war-of-aggression-against-ukraine-additional-71-individuals-and-33-entities-included-in-the-eu-s-sanctions-list-and-new-tools-to-counter-circumvention-and-information-warfare/ [08/09/2023].

137. EUvsDisinfo. UKRAINE. https://euvsdisinfo.eu/ukraine/ [08/09/2023].

138. European Digital Media Observatory. EDMO Task Force on Disinformation on the War in Ukraine. https://edmo.eu/edmo-task-force-on-disinformation-on-the-war-in-ukraine/ [08/09/2023].

139. EUvsDisinfo. 'To Challenge Russia's Ongoing Disinformation Campaigns': Eight Years of EUvsDisinfo. *EUvsDisinfo* https://euvsdisinfo.eu/to-challenge-russias-ongoing-disinformation-campaigns-eight-years-of-euvsdisinfo/ [23/11/2023].

140. Council of the EU. The fight against pro-Kremlin disinformation. https://www.consilium.europa.eu/en/documents-publications/library/library-blog/posts/the-fight-against-pro-kremlin-disinformation/ [23/11/2023].

141. European Commission. *Digital Services Act: application of the risk management framework to Russian disinformation campaigns.* https://data.europa.eu/doi/10.2759/764631 (2023).

142. Pamment, J. The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework. *Carnegie Endowment for International Peace*

https://carnegieendowment.org/2020/09/24/eu-s-role-in-fighting-disinformation-crafting-disinformation-framework-pub-82720 [26/06/2023].
143. European Partnership for Democracy. *Louder than words? Connecting the dots of European democracy support*.
https://epd.eu/news-publications/louder-than-words-connecting-the-dots-of-european-democracy-support/
https://epd.eu/news-publications/louder-than-words-connecting-the-dots-of-european-democracy-support/ (2019).
144. Bel Mallén, I. La realidad es un proceso pernicioso. *Fake news, desinformación y otros desordenes informativos. Directoras Ángela Moreno Bobadilla y Beatriz del Carmen Martínez Isidoro.* (Fragua, 2022). ISBN: 978-84-7074-963-6.
145. La independencia de los medios españoles ante los grupos de presión, bajo sospecha | Digital News Report España 2023 (DNR): informe de noticias digitales en español.
https://www.digitalnewsreport.es/2022/la-independencia-de-los-medios-espanoles-ante-los-grupos-de-presion-bajo-sospecha/ [18/09/2023].
146. Mont'Alverne, C., Badrinathan, S., Ross Arguedas, A., Toff, B., Fletcher, R. & Kleis Nielsen, R. *The trust gap: how and why news on digital platforms is viewed more sceptically versus news in general.*
https://reutersinstitute.politics.ox.ac.uk/trust-gap-how-and-why-news-digital-platforms-viewed-more-sceptically-versus-news-general (2022).
147. Burgess, J. & Hurcombe, E. Digital Journalism as Symptom, Response, and Agent of Change in the Platformed Media Environment. *Digital Journalism* **7**, 359–367 (2019)
www.doi.org/10.1080/21670811.2018.1556313.
148. García-Gordillo, M., Palau-Sampio, D. & Rivas-de-Roca, R. Capítulo 3. Pero ¿qué me cuentas? Una revisión del concepto de verdad en el S. XXI. *Espejo de Monografías de Comunicación Social* 61–81 (2023)
www.doi.org/10.52495/c3.emcs.11.p98.
149. Gómez Mompart, J. L., Gutiérrez Lozano, J. F. & Palau-Sampio, D. La calidad periodística en España según la percepción de los periodistas. *Estudios sobre el mensaje periodístico* 13–30 (2015).
150. Sampio, D. P. & Carratalá, A. Injecting disinformation into public space: pseudo-media and reality-altering narratives. *Profesional de la información* **31**, (2022)
www.doi.org/10.3145/epi.2022.may.12.
151. Palomares, P. J. & Silva, F. G.-F. e. Visibilidad de la información en redes sociales: los algoritmos de Facebook y su influencia en el clickbait. *Caleidoscopio - Revista Semestral de Ciencias Sociales y Humanidades* 173–211 (2019)
www.doi.org/10.33064/41crscsh1772.
152. Bravo Araujo, A., Serrano Puche, J. & Novoa Jaso, M. F. Uso del clickbait en los medios nativos digitales españoles. Un análisis de El Confidencial, El Español, Eldiario.es y Ok Diario. *Dígitos: Revista de Comunicación Digital, 7: 185-* (2021)
www.doi.org/10.7203/rd.v1i7.184.
153. McGregor, S. C. Social media as public opinion: How journalists use social media to represent public opinion. *Journalism* **20**, 1070–1086 (2019)
www.doi.org/10.1177/1464884919845458.
154. Denisova, A. 'Viral journalism', is it a thing? Adapting quality reporting to shifting social media algorithms and wavering audiences. (eds. Morrison, J., Birks, J. & Berry, M.) 271–278 (Routledge, 2021). ISBN: 978-0-367-24822-2.
155. Petre, C. *All the News That's Fit to Click: How Metrics Are Transforming the Work of Journalists*. (Princeton University Press, 2021). ISBN: 978-0-691-17764-9.
156. Corredoira, L. Anonimato, transparencia e identificación de fuentes, informantes y robots en la era del algoritmo. *Derecho Público de la Inteligencia Artificial* (2023).
157. Mancini, P. Comparing Media Systems and the Digital Age. *International Journal of Communication* **14**, 14 (2020).
158. Busquet Durán, J. Sistemas mediáticos comparados: tres modelos de relación entre los medios de comunicación y la política. *REIS: Revista Española de Investigaciones Sociológicas* 165–172 (2010).
159. Newman, N., Fletcher, R., Robertson, C. T., Eddy, K. & Nielsen, R. K. *Reuters Institute Digital News Report 2022.*
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf (2022).
160. Vara-Miguel, A., Amoedo-Casais, A., Moreno-Moreno, E., Negredo-Bruna, S. & Kaufmann-Argueta, J. *Digital News Report España 2022*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra. DOI: www.doi.org/10.15581/019.2022 (2022).
161. FAPE. Código Deontológico.
https://fape.es/home/codigo-deontologico/ [02/11/2023].
162. Democratic Efficacy and the Varieties of Populism in Europe. Partisan Media Erodes Trust in Politics, New Study Claims.
https://demos-h2020.eu/en/partisan-media-erodes-trust-in-politics-new-study-claims [18/09/2023].
163. Majó-Vázquez, S. & González-Bailón, S. Polarización en las audiencias de los medios en España. *Center for Economic Policy - EsadeEcPol*
https://www.esade.edu/ecpol/es/publicaciones/polarizacion-medios-espana/ [30/09/2023].
164. Rosa, R. M. La transparencia en la publicidad institucional como garantía de pluralismo informativo. *The Conversation* http://theconversation.com/la-transparencia-en-la-publicidad-institucional-como-garantia-de-pluralismo-informativo-149716 [30/09/2023].
165. Rosa, R. M. La publicidad institucional en España. Evolución legislativa, tecnológica y social. (2020) http://dx.doi.org/10.5209/arab.67255.
166. España. *Ley 29/2005, de 29 de diciembre, de Publicidad y Comunicación Institucional. Jefatura del Estado BOE núm. 312, de 30 de diciembre de 2005 Referencia: BOE-A-2005-21524.*
167. Waisbord, S. Truth is What Happens to News. *Journalism Studies* **19**, 1866–1878 (2018)
www.doi.org/10.1080/1461670X.2018.1492881.
168. Krasni, J. How to hijack a discourse? Reflections on the concepts of post-truth and fake news. *Humanities and Social Sciences Communications* **7**, 1–10 (2020)
www.doi.org/10.1057/s41599-020-0527-z.
169. De Blasio, E. & Selva, D. Who Is Responsible for Disinformation? European Approaches to Social Platforms' Accountability in the Post-Truth Era. *American Behavioral Scientist* **65**, 825–846 (2021)
www.doi.org/10.1177/0002764221989784.
170. European Commission. Identifying conspiracy theories.
https://commission.europa.eu/strategy-and-policy/coronavirus-response/fighting-disinformation/identifying-conspiracy-theories_en [12/09/2023].
171. Lewandowsky, S. & Cook, J. *The Conspiracy Theory Handbook*.
https://digitalcommons.unl.edu/scholcom/246 (2020).
172. Orriols, L. La polarización afectiva en España: bloques ideológicos enfrentados. *ESADE*
https://dobetter.esade.edu/es/polarizacion-afectiva [28/09/2023].
173. Lewandowsky, S., Ecker, U. K. H. & Cook, J. Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era. *Journal of Applied Research in Memory and Cognition* **6**, 353–369 (2017)
www.doi.org/10.1016/j.jarmac.2017.07.008.
174. Lewandowsky, S. Hannah Arendt and the contemporary social construction of conspiracy theorists. *PsyArXiv* (2020)
www.doi.org/10.31234/osf.io/fm8yg.
175. Pantazi, M., Hale, S. & Klein, O. Social and Cognitive Aspects of the Vulnerability to Political Misinformation. *Political Psychology* **42**, 267–304 (2021)
www.doi.org/10.1111/pops.12797.
176. Fiedler, K. Metacognitive Myopia: Gullibility as a Major Obstacle in the Way of Rational Behavior. *The Social Psychology of Gullibility* (Routledge, 2019). ISBN: 978-0-429-20378-7.
177. Cazorla, A., Montabes, J. & López-López, P. C. Medios de comunicación, información política y emociones hacia partidos políticos en España. *Revista Española de Ciencia Política* **58**, 83–109 (2022)
www.doi.org/10.21308/recp.58.03.
178. Ross Arguedas, A., Robertson, C., Fletcher, R. & Nielsen, R. *Echo chambers, filter bubbles, and polarisation: a literature review.*
https://ora.ox.ac.uk/objects/uuid:6e357e97-7b16-450a-a827-a92c93729a08 (2022).
179. Pérez-Escolar, M. & Noguera-Vivo, J. M. *Hate Speech and Polarization in Participatory Society. Routledge Studies in Media, Communication, and Politics*. (2022). ISBN: 978-0-367-62601-3.
180. Torcal, M. *De votantes a hooligans. La polarización política en España - Catarata.* (Catarata). ISBN: 978-84-13-52614-0.
181. Crockett, M. J. Moral outrage in the digital age. *Nature Human Behaviour* **1**, 769–771 (2017)
www.doi.org/10.1038/s41562-017-0213-3.
182. Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A. & Van Bavel, J. J. Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* **114**, 7313–7318 (2017)
www.doi.org/10.1073/pnas.1618923114.
183. Petersen, M. B., Osmundsen, M. & Arceneaux, K. The "Need for Chaos" and Motivations to Share Hostile Political Rumors. *American Political Science Review* **117**, 1486–1505 (2023)
www.doi.org/10.1017/S0003055422001447.
184. van der Linden, S. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine* **28**, 460–467 (2022)
www.doi.org/10.1038/s41591-022-01713-6.
185. Pennycook, G. & Rand, D. G. The Psychology of Fake News. *Trends in Cognitive Sciences* **25**, 388–402 (2021)
www.doi.org/10.1016/j.tics.2021.02.007.
186. Wolters, H., Stricklin, K., Carey, N. & McBride, M. K. *The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms*.
https://www.cna.org/reports/2021/10/psychology-of-disinformation-key-psychological-mechanisms (2021).
187. García, P. A. P., Danieli, N. E. & Freire, I. E. M. Cognitive processing of political fake news. Review of experimental studies. *Dixit* **37**, 44–60
www.doi.org/10.22235/d.v37i1.3112.
188. Vidgen, B., Taylor, H., Pantazi, M., Anastasiou, Z., Inkster, B. & Margetts, H. *Understanding vulnerability to online misinformation*. The Alan Turing Institute
https://www.turing.ac.uk/news/publications/understanding-vulnerability-online-misinformation (2021).
189. Howard, P. N., Neudert, L.-M., Prakash, N. & Vosloo, S. *Digital misinformation / disinformation and children*. UNICEF
https://www.unicef.org/globalinsight/reports/digital-misinformation-disinformation-and-children (2021).
190. Pantazi, M., Kissine, M. & Klein, O. The Power of the Truth Bias: False Information Affects Memory and Judgment Even in the Absence of Distraction. *Social Cognition* **36**, (2018)
www.doi.org/10.1521/soco.2018.36.2.167.
191. Prike, T., Arnold, M. M. & Williamson, P. The relationship between anomalistic belief, misperception of chance and the base rate fallacy. *Thinking & Reasoning* **26**, 447–477 (2020)
www.doi.org/10.1080/13546783.2019.1653371.
192. Brashier, N. M. & Marsh, E. J. Judging Truth. *Annual Review of Psychology* **71**, 499–515 (2020)
www.doi.org/10.1146/annurev-psych-010419-050807.
193. Begg, I. M., Anas, A. & Farinacci, S. Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General* **121**, 446–458 (1992)

www.doi.org/10.1037/0096-3445.121.4.446.

194. Unkelbach, C. & Rom, S. C. A referential theory of the repetition-induced truth effect. *Cognition* **160**, 110–126 (2017) www.doi.org/10.1016/j.cognition.2016.12.016.

195. Mitchell, K. J. & Johnson, M. K. Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin* **135**, 638–677 (2009) www.doi.org/10.1037/a0015849.

196. Nadarevic, L., Reber, R., Helmecke, A. J. & Köse, D. Perceived truth of statements and simulated social media postings: an experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research: Principles and Implications* **5**, 56 (2020) www.doi.org/10.1186/s41235-020-00251-4.

197. Briñol, P. & Petty, R. E. Source factors in persuasion: A self-validation approach. *European Review of Social Psychology* **20**, 49–96 (2009) www.doi.org/10.1080/10463280802643640.

198. Altay, S., de Araujo, E. & Mercier, H. "If This account is True, It is Most Enormously Wonderful": Interestingness-If-True and the Sharing of True and False News. *Digital Journalism* **10**, 373–394 (2022) www.doi.org/10.1080/21670811.2021.1941163.

199. Ceylan, G., Anderson, I. A. & Wood, W. Sharing of misinformation is habitual, not just lazy or biased. *Proceedings of the National Academy of Sciences* **120**, e2216614120 (2023) www.doi.org/10.1073/pnas.2216614120.

200. Pennycook, G. & Rand, D. G. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* **188**, 39–50 (2019) www.doi.org/10.1016/j.cognition.2018.06.011.

201. Brady, W. J., Crockett, M. J. & Van Bavel, J. J. The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online. *Perspectives on Psychological Science* **15**, 978–1010 (2020) www.doi.org/10.1177/1745691620917336.

202. Islam, A. K. M. N., Laato, S., Talukder, S. & Sutinen, E. Misinformation sharing and social media fatigue during COVID-19: An affordance and cognitive load perspective. *Technological Forecasting and Social Change* **159**, 120201 (2020) www.doi.org/10.1016/j.techfore.2020.120201.

203. Flore, M. Understanding Citizens' Vulnerabilities (II): From Disinformation to Hostile Narratives. *JRC Publications Repository* https://publications.jrc.ec.europa.eu/repository/handle/JRC118914 [30/05/2023] www.doi.org/10.2760/271224.

204. Brady, W. J., Gantman, A. P. & Van Bavel, J. J. Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General* **149**, 746–756 (2020) www.doi.org/10.1037/xge0000673.

205. Forgas, J. P. Happy Believers and Sad Skeptics? Affective Influences on Gullibility. *Current Directions in Psychological Science* **28**, 306–313 (2019) www.doi.org/10.1177/0963721419834543.

206. Morosoli, S., Van Aelst, P., Humprecht, E., Staender, A. & Esser, F. Identifying the Drivers Behind the Dissemination of Online Misinformation: A Study on Political Attitudes and Individual Characteristics in the Context of Engaging With Misinformation on Social Media. *American Behavioral Scientist* 00027642221118300 (2022) www.doi.org/10.1177/00027642221118300.

207. González-Bailón, S. *et al.* Asymmetric ideological segregation in exposure to political news on Facebook. *Science* **381**, 392–398 (2023) www.doi.org/10.1126/science.ade7138.

208. Guess, A. M., Barberá, P., Munzert, S. & Yang, J. The consequences of online partisan media. *Proceedings of the National Academy of Sciences* **118**, e2013464118 (2021) www.doi.org/10.1073/pnas.2013464118.

209. McIlhiney, P., Gignac, G. E., Weinborn, M. & Ecker, U. K. H. Sensitivity to misinformation retractions in the continued influence paradigm: Evidence for stability. *Quarterly Journal of Experimental Psychology* **75**, 1259–1271 (2022) www.doi.org/10.1177/17470218211048986.

210. Prike, T., Blackley, P., Swire-Thompson, B. & Ecker, U. K. H. Examining the replicability of backfire effects after standalone corrections. *Cognitive Research: Principles and Implications* **8**, (2023) www.doi.org/10.1186/s41235-023-00492-z.

211. Swire-Thompson, B., Miklaucic, N., Wihbey, J. P., Lazer, D. & DeGutis, J. The backfire effect after correcting misinformation is strongly associated with reliability. *Journal of Experimental Psychology. General* **151**, 1655–1665 (2022) www.doi.org/10.1037/xge0001131.

212. Stoycheff, E., Burgess, G. S. & Martucci, M. C. Online censorship and digital surveillance: the relationship between suppression technologies and democratization across countries. *Information, Communication & Society* **23**, 474–490 (2020) www.doi.org/10.1080/1369118X.2018.1518472.

213. Ververis, V., Marguel, S. & Fabian, B. Cross-Country Comparison of Internet Censorship: A Literature Review. *Policy & Internet* **12**, 450–473 (2020) www.doi.org/10.1002/poi3.228.

214. Andrés, M. B. Modelos de negocio basados en datos, publicidad programática, inteligencia artificial y regulación: algunas reflexiones. *IDP. Revista de Internet, Derecho y Política* 1–13 (2022) www.doi.org/10.7238/idp.v0i36.401947.

215. Ryan, C. D., Schaul, A. J., Butner, R. & Swarthout, J. T. Monetizing disinformation in the attention economy: The case of genetically modified organisms (GMOs). *European Management Journal* **38**, 7–18 (2020) www.doi.org/10.1016/j.emj.2019.11.002.

216. Bakir, V. & McStay, A. Fake News and The Economy of Emotions. *Digital Journalism* **6**, 154–175 (2018) www.doi.org/10.1080/21670811.2017.1345645.

217. Mourão, R. R. & Robertson, C. T. Fake News as Discursive Integration: An Analysis of Sites That Publish False, Misleading, Hyperpartisan and Sensational Information. *Journalism Studies* **20**, 2077–2095 (2019) www.doi.org/10.1080/1461670X.2019.1566871.

218. Staender, A., Humprecht, E., Esser, F., Morosoli, S. & Van Aelst, P. Is Sensationalist Disinformation More Effective? Three Facilitating Factors at the National, Individual, and Situational Level. *Digital Journalism* **10**, 976–996 (2022) www.doi.org/10.1080/21670811.2021.1966315.

219. Braun, J. A. & Eklund, J. L. Fake News, Real Money: Ad Tech Platforms, Profit-Driven Hoaxes, and the Business of Journalism. *Digital Journalism* **7**, 1–21 (2019) www.doi.org/10.1080/21670811.2018.1556314.

220. Global disinformation Index. *The Quarter Billion Dollar Question: How is Disinformation Gaming Ad Tech?* https://www.disinformationindex.org/files/gdi_ad-tech_report_screen_aw16.pdf (2019).

221. Global disinformation Index. *Research Brief: Ad Tech Fuels Disinformation Sites in Europe – The Numbers and Players*. https://www.disinformationindex.org/research/2020-3-1-research-brief-ad-tech-fuels-disinformation-sites-in-europe-the-numbers-and-players/ (2020).

222. European Commison. A European Strategy for data | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/policies/strategy-data [10/06/2022].

223. Garcia, D. Privacy beyond the individual. *Nature Human Behaviour* **3**, 112–113 (2019) www.doi.org/10.1038/s41562-018-0513-2.

224. Bagrow, J. P., Liu, X. & Mitchell, L. Information flow reveals prediction limits in online social activity. *Nature Human Behaviour* **3**, 122–128 (2019) www.doi.org/10.1038/s41562-018-0510-5.

225. Dommett, K. Data-driven political campaigns in practice: understanding and regulating diverse data-driven campaigns. *Internet Policy Review* **8**, (2019) www.doi.org/10.14763/2019.4.1432.

226. Zuiderveen Borgesius, F. J. *et al.* Online Political Microtargeting: Promises and Threats for Democracy. *Utrecht Law Review* **14**, 82 (2018) www.doi.org/10.18352/ulr.420.

227. Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L. & Fei-Fei, L. Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proceedings of the National Academy of Sciences* **114**, 13108–13113 (2017) www.doi.org/10.1073/pnas.1700035114.

228. García Sanz, R. M. Questioning the protection of European General Data Regulation: identifying key problems. *Derecom* 3 (2019).

229. Center for Humane Technology. Ledger of Harms. https://ledger.humanetech.com/ [11/09/2023].

230. Metzler, H. & Garcia, D. Social Drivers and Algorithmic Mechanisms on Digital Media. *Perspectives on Psychological Science* (2023) www.doi.org/10.1177/17456916231185057.

231. Crain, M. & Nadler, A. Political Manipulation and Internet Advertising Infrastructure. *Journal of Information Policy* **9**, 370–410 (2019) www.doi.org/10.5325/jinfopoli.9.2019.0370.

232. Kreps, S. *The role of technology in online misinformation*. Brookings https://www.brookings.edu/articles/the-role-of-technology-in-online-misinformation/ (2020).

233. Cardenal, A. S., Aguilar-Paredes, C., Galais, C. & Pérez-Montoro, M. Digital Technologies and Selective Exposure: How Choice and Filter Bubbles Shape News Media Exposure. *International Journal of Press/Politics* **24**, 465–486 (2019) www.doi.org/10.1177/1940161219862988.

234. Innerarity, D. Justicia algorítmica y autodeterminación deliberativa. *Isegoría* e23–e23 (2023) www.doi.org/10.3989/isegoria.2023.68.23.

235. Shin, D. *et al.* Countering Algorithmic Bias and Disinformation and Effectively Harnessing the Power of AI in Media. *Journalism & Mass Communication Quarterly* **99**, 887–907 (2022) www.doi.org/10.1177/10776990221129245.

236. Mohseni, S. & Ragan, E. Combating fake news with interpretable news feed algorithms. *arXiv* (2018) www.doi.org/10.48550/arXiv.1811.12349.

237. Saurwein, F. & Spencer-Smith, C. Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe. *Digital Journalism* **8**, 820–841 (2020) www.doi.org/10.1080/21670811.2020.1765401.

238. Saurwein, F. Emerging structures of control for algorithms on the Internet: Distributed agency - distributed accountability. 196–211 (2019). ISBN: 978-1-351-11578-0.

239. Shaw, A. Social media, extremism, and radicalization. *Science Advances* **9**, eadk2031 (2023) www.doi.org/10.1126/sciadv.adk2031.

240. Hosseinmardi, H., Ghasemian, A., Clauset, A., Mobius, M., Rothschild, D. M. & Watts, D. J. Examining the consumption of radical content on YouTube. *Proceedings of the National Academy of Sciences* **118**, e2101967118 (2021) www.doi.org/10.1073/pnas.2101967118.

241. Wolfowicz, M., Weisburd, D. & Hasisi, B. Examining the interactive effects of the filter bubble and the echo chamber on radicalization. *Journal of Experimental Criminology* **19**, 119–141 (2023) www.doi.org/10.1007/s11292-021-09471-0.

242. O'Callaghan, D., Greene, D., Conway, M., Carthy, J. & Cunningham, P. Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems. *Social Science Computer Review* **33**, 459–478 (2015) www.doi.org/10.1177/0894439314555329.

243. Colleoni, E., Rozza, A. & Arvidsson, A. Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. *Journal of Communication* **64**, 317–332 (2014) www.doi.org/10.1111/jcom.12084.

244. Pariser, E. *The Filter Bubble: What The Internet Is Hiding From You*. (Penguin UK, 2011). ISBN: 978-0-14-196992-3.

245. Garcia, D. Influence of Facebook algorithms on political polarization tested. *Nature* **620**, 39–41 (2023) www.doi.org/10.1038/d41586-023-02325-x.

246. Nyhan, B. *et al.* Like-minded sources on Facebook are prevalent but not polarizing. *Nature* **620**, 137–144 (2023) www.doi.org/10.1038/s41586-023-06297-w.

247. Guess, A., Lyons, B., Nyhan, B. & Reifler, J. *Avoiding the echo chamber about echo chambers: Why selective exposure to like-minded political news is less prevalent than you think*. (2018).

248. Masip, P., Suau, J. & Ruiz-Caballero, C. Incidental exposure to non-like-minded news through social media: opposing voices in echo-chambers' news feeds. *Media and Communication* **8**, 53–62 (2020) www.doi.org/10.17645/mac.v8i4.3146.

249. Guess, A. M. *et al.* How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science* **381**, 398–404 (2023) www.doi.org/10.1126/science.abp9364.

250. Chen, A. Y., Nyhan, B., Reifler, J., Robertson, R. E. & Wilson, C. Subscriptions and external links help drive resentful users to alternative and extremist YouTube channels. *Science Advances* **9**, eadd8080 (2023) www.doi.org/10.1126/sciadv.add8080.

251. Kolomeets, M. & Chechulin, A. Analysis of the Malicious Bots Market. *2021 29th Conference of Open Innovations Association (FRUCT)* 199–205 (2021). www.doi.org/10.23919/FRUCT52173.2021.9435421.

252. Keller, T. R. & Klinger, U. Social Bots in Election Campaigns: Theoretical, Empirical, and Methodological Implications. *Political Communication* **36**, 171–189 (2019) www.doi.org/10.1080/10584609.2018.1526238.

253. Luceri, L., Deb, A., Giordano, S. & Ferrara, E. Evolution of bot and human behavior during elections. *First Monday* (2019) www.doi.org/10.5210/fm.v24i9.10213.

254. Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A. & Menczer, F. The spread of low-credibility content by social bots. *Nature Communications* **9**, 4787 (2018) www.doi.org/10.1038/s41467-018-06930-7.

255. González-Bailón, S. & De Domenico, M. Bots are less central than verified accounts during contentious political events. *Proceedings of the National Academy of Sciences* **118**, e2013443118 (2021) www.doi.org/10.1073/pnas.2013443118.

256. Hussain, M. N., Tokdemir, S., Agarwal, N. & Al-Khateeb, S. Analyzing Disinformation and Crowd Manipulation Tactics on YouTube. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* 1092–1095 (2018). www.doi.org/10.1109/ASONAM.2018.8508766.

257. Giachanou, A., Zhang, X., Barrón-Cedeño, A., Koltsova, O. & Rosso, P. Online information disorder: fake news, bots and trolls. *International Journal of Data Science and Analytics* **13**, 265–269 (2022) www.doi.org/10.1007/s41060-022-00325-0.

258. García-Orosa, B. Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy. *Profesional de la información* **30**, (2021) www.doi.org/10.3145/epi.2021.nov.03.

259. Jachim, P., Sharevski, F. & Pieroni, E. TrollHunter2020: Real-time Detection of Trolling Narratives on Twitter During the 2020 U.S. Elections. *Proceedings of the 2021 ACM Workshop on Security and Privacy Analytics* 55–65 (Association for Computing Machinery, 2021). www.doi.org/10.1145/3445970.3451158.

260. Santini, R. M., Salles, D., Tucci, G., Ferreira, F. & Grael, F. Making up Audience: Media Bots and the Falsification of the Public Sphere. *Communication Studies* **71**, 466–487 (2020) www.doi.org/10.1080/10510974.2020.1735466.

261. Zannettou, S., Caulfield, T., De Cristofaro, E., Sirivianos, M., Stringhini, G. & Blackburn, J. Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web. *Companion Proceedings of The 2019 World Wide Web Conference* 218–226 (Association for Computing Machinery, 2019). www.doi.org/10.1145/3308560.3316495.

262. Keller, F. B., Schoch, D., Stier, S. & Yang, J. Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign. *Political Communication* **37**, 256–280 (2020) www.doi.org/10.1080/10584609.2019.1661888.

263. Lukito, J. *et al.* The Wolves in Sheep's Clothing: How Russia's Internet Research Agency Tweets Appeared in U.S. News as Vox Populi. *The International Journal of Press/Politics* **25**, 196–216 (2020) www.doi.org/10.1177/1940161219895215.

264. Schoch, D., Keller, F. B., Stier, S. & Yang, J. Coordination patterns reveal online political astroturfing across the world. *Scientific Reports* **12**, 4572 (2022) www.doi.org/10.1038/s41598-022-08404-9.

265. Chan, J. Online astroturfing: A problem beyond disinformation. *Philosophy & Social Criticism* 01914537221108467 (2022) www.doi.org/10.1177/01914537221108467.

266. Bail, C. A. *et al.* Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 9216–9221 (2018) www.doi.org/10.1073/pnas.1804840115.

267. Van Bavel, J. J., Robertson, C., del Rosario, K., Rasmussen, J. & Rathje, S. Social Media and Morality. *Annual Review of Psychology* www.doi.org/10.31234/osf.io/ywevq.

268. Gonzalez-Cabañas, J., Cuevas, A., Cuevas, R., Lopez-Fernandez, J. & Garcia, D. Unique on Facebook: Formulation and evidence of (nano)targeting individual users with non-PII data. 464–479 (2021). www.doi.org/10.1145/3487552.3487861.

269. Kim, Y. M. *et al.* The Stealth Media? Groups and Targets behind Divisive Issue Campaigns on Facebook. *Political Communication* **35**, 515–541 (2018) www.doi.org/10.1080/10584609.2018.1476425.

270. Heawood, J. Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal. *Information Polity* **23**, 429–434 (2018) www.doi.org/10.3233/IP-180009.

271. Lorenz-Spreen, P., Lewandowsky, S., Sunstein, C. R. & Hertwig, R. How behavioural sciences can promote truth, autonomy and democratic discourse online. *Nature Human Behaviour* **4**, 1102–1109 (2020) www.doi.org/10.1038/s41562-020-0889-7.

272. Rothwell, J. Perspectives: Find your audience on digital and storytell with data. *Think with Google* https://www.thinkwithgoogle.com/intl/en-apac/marketing-strategies/video/perspectives-find-your-audience-digital-and-storytell-data/ [05/07/2023].

273. Martin, K. *Ethics of Data and Analytics: Concepts and Cases*. (CRC Press, 2022). ISBN: 978-1-00-056626-0.

274. Blasi Casagran, C. & Vermeulen, M. Reflections on the murky legal practices of political micro-targeting from a GDPR perspective. *International Data Privacy Law* **11**, 348–359 (2021) www.doi.org/10.1093/idpl/ipab018.

275. Dobber, T., Ó Fathaigh, R. & Zuiderveen Borgesius, F. J. The regulation of online political micro-targeting in Europe. *Internet Policy Review* **8**, (2019) www.doi.org/10.14763/2019.4.1440.

276. López-López, P. C., Barredo-Ibáñez, D. & Jaráiz-Gulías, E. Research on Digital Political Communication: Electoral Campaigns, Disinformation, and Artificial Intelligence. *Societies* **13**, 126 (2023) www.doi.org/10.3390/soc13050126.

277. Ryabtsev, K. Political Micro-Targeting in Europe: A Panacea for the Citizens' Political Misinformation or the New Evil for Voting Rights. *Groningen Journal of International Law* **8**, 69–89 (2020) www.doi.org/10.21827/GroJIL.8.1.69-89.

278. Gibney, E. The scant science behind Cambridge Analytica's controversial marketing techniques. *Nature* (2018) www.doi.org/10.1038/d41586-018-03880-4.

279. Hinds, J., Williams, E. J. & Joinson, A. N. "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies* **143**, 102498 (2020) www.doi.org/10.1016/j.ijhcs.2020.102498.

280. Comisión Europea. *Propuesta de reglamento del Parlamento Europeo y del Consejo sobre la transparencia y la segmentación de la publicidad política*. (2021).

281. Ozawa, J. V. S., Woolley, S. C., Straubhaar, J., Riedl, M. J., Joseff, K. & Gursky, J. How Disinformation on WhatsApp Went From Campaign Weapon to Governmental Propaganda in Brazil. *Social Media + Society* **9**, 20563051231160630 (2023) www.doi.org/10.1177/20563051231160632.

282. Paris, B. & Pasquetto, I. Hidden Virality and the Everyday Burden of Correcting WhatsApp Mis- and Disinformation. *Governing Everyday Misinformation* (Cambridge University Press, 2024).

283. Chauchard, S. & Garimella, K. What Circulates on Partisan WhatsApp in India? Insights from an Unusual Dataset. *Journal of Quantitative Description: Digital Media* **2**, (2022) www.doi.org/10.51685/jqd.2022.006.

284. European Parliament. Artificial intelligence: threats and opportunities. https://www.europarl.europa.eu/news/en/headlines/society/20200918STO87404/artificial-intelligence-threats-and-opportunities [29/09/2023].

285. European Parliament. Directorate General for Parliamentary Research Services. *Tackling deepfakes in European policy*. (Publications Office, 2021).

286. Westerlund, M. The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review* **9**, 39–52 (2019) www.doi.org/10.22215/timreview/1282.

287. Guo, D., Chen, H., Wu, R. & Wang, Y. AIGC Challenges and Opportunities Related to Public Safety: A Case Study of ChatGPT. *Journal of Safety Science and Resilience* (2023) www.doi.org/10.1016/j.jnlssr.2023.08.001.

288. Montoro-Montarroso, A. *et al.* Fighting disinformation with artificial intelligence: fundamentals, advances and challenges. *Profesional de la información* **32**, (2023) www.doi.org/10.3145/epi.2023.may.22.

289. Brown, T. B. *et al.* Language Models are Few-Shot Learners. *arXiv:2005.14165 [cs]* (2020) www.doi.org/10.48550/arXiv.2005.14165.

290. Spitale, G., Biller-Andorno, N. & Germani, F. AI model GPT-3 (dis)informs us better than humans. *Science Advances* **9**, eadh1850 (2023) www.doi.org/10.1126/sciadv.adh1850.

291. Fütterer, T., Fischer, C., Alekseeva, A., Chen, X., Tate, T., Warschauer, M. & Gerjets, P. ChatGPT in education: global reactions to AI innovations. *Scientific Reports* **13**, 15310 (2023) www.doi.org/10.1038/s41598-023-42227-6.

292. Meskó, B. & Topol, E. J. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *npj Digital Medicine* **6**, 1–6 (2023) www.doi.org/10.1038/s41746-023-00873-0.

293. De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E. & Rizzo, C. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. *Frontiers in Public Health* **11**, (2023).

294. Sun, Y., He, J., Lei, S., Cui, L. & Lu, C. T. Med-MMHL: A Multi-Modal Dataset for Detecting Human-and LLM-Generated Misinformation in the Medical Domain. *arXiv:2306.08871 [cs.SI]* (2023)

www.doi.org/10.48550/arXiv.2306.08871.

295. Galaz, V., Metzler, H., Daume, S., Olsson, A., Lindström, B. & Marklund, A. *AI could create a perfect storm of climate misinformation*. Stockholm Resilience Centre (Stockholm University) and the Beijer Institute of Ecological Economics (Royal Swedish Academy of Sciences) http://arxiv.org/abs/2306.12807 (2023).

296. Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y. & Pearson, A. T. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *npj Digital Medicine* **6**, 1–5 (2023)
www.doi.org/10.1038/s41746-023-00819-6.

297. Wang, B. *et al.* Wang, B., Chen, W., Pei, H., Xie, C., Kang, M., Zhang, C., ... & Li, B. (2023). DecodingTrust: A Comprehensive Assessment of Trustworthiness in GPT Models. arXiv preprint arXiv:2306.11698. *arXiv:2306.11698 [cs.CL]* (2023)
www.doi.org/10.48550/arXiv.2306.11698.

298. Boucher, P. *What if deepfakes made us doubt everything we see and hear?* European Parliamentary Research Service
https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2021)690046 (2021).

299. Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A. & Malik, H. Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward. *Applied Intelligence* **53**, 3974–4026 (2023)
www.doi.org/10.1007/s10489-022-03766-z.

300. Rafique, R., Gantassi, R., Amin, R., Frnda, J., Mustapha, A. & Alshehri, A. H. Deep fake detection and classification using error-level analysis and deep learning. *Scientific Reports* **13**, 7422 (2023)
www.doi.org/10.1038/s41598-023-34629-3.

301. Greengard, S. Will deepfakes do deep damage? *Communications of the ACM* **63**, 17–19 (2019)
www.doi.org/10.1145/3371409.

302. UC Berkeley School of Information. Hany Farid Explains How AI Voice Cloning Fuels Misinformation.
https://www.ischool.berkeley.edu/news/2023/hany-farid-explains-how-ai-voice-cloning-fuels-misinformation [16/09/2023].

303. Rini, R. & Cohen, L. Deepfakes, Deep Harms. *Journal of Ethics and Social Philosophy* **22**, 143 (2022).

304. Home Security Heroes. 2023 State Of Deepfakes: Realities, Threats, And Impact.
https://www.homesecurityheroes.com/state-of-deepfakes/#appendix [03/11/2023].

305. Mania, K. Legal Protection of Revenge and Deepfake Porn Victims in the European Union: Findings From a Comparative Legal Study. *Trauma, Violence, & Abuse* **25**, 117–129 (2024)
www.doi.org/10.1177/15248380221143772.

306. Ocaña, J. Caso de Almendralejo: ¿por qué los menores crean contenidos falsos de carácter sexual con IA? *EFE Verifica*
https://verifica.efe.com/almendralejo-inteligencia-artiificial-machismo-menores/ [23/11/2023].

307. Universitat Oberta de Catalunya. El 99% del deepfake dirigido contra las mujeres.
https://www.uoc.edu/portal/es/news/actualitat/2023/265-deepfakes-pornograficos-cuando-IA-desnuda-tu-intimidad-vulnera-tus-derechos.html [23/11/2023].

308. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A. & Ortega-Garcia, J. Deepfakes and beyond: A Survey of face manipulation and fake detection. *Information Fusion* **64**, 131–148 (2020)
www.doi.org/10.1016/j.inffus.2020.06.014.

309. Dwivedi, Y. K. *et al.* Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management* **66**, 102542 (2022)
www.doi.org/10.1016/j.ijinfomgt.2022.102542.

310. Mitchell, E., Lee, Y., Khazatsky, A., Manning, C. D. & Finn, C. Detectgpt: Zero-shot machine-generated text detection using probability curvature. *arXiv:2301.11305 [cs.CL]* (2023)
www.doi.org/10.48550/arXiv.2301.11305.

311. Sarvazyan, A. M., González, J. Á., Franco-Salvador, M., Rangel, F., Chulvi, B. & Rosso, P. Overview of AuTexTification at IberLEF 2023: Detection and Attribution of Machine-Generated Text in Multiple Domains. *Procesamiento del Lenguaje Natural* **71**, 275–288 (2023).

312. Sarvazyan, A. M., González, J. Á., Rosso, P. & Franco-Salvador, M. Supervised Machine-Generated Text Detectors: Family and Scale Matters. *Experimental IR Meets Multilinguality, Multimodality, and Interaction* (eds. Arampatzis, A. et al.) 121–132 (Springer Nature Switzerland, 2023).
www.doi.org/10.1007/978-3-031-42448-9_11.

313. U.S. GAO. *Science & Tech Spotlight: Deepfakes*.
https://www.gao.gov/products/gao-20-379sp (2020).

314. van der Sloot, B. & Wagensveld, Y. Deepfakes: regulatory challenges for the synthetic society. *Computer Law & Security Review* **46**, 105716 (2022)
www.doi.org/10.1016/j.clsr.2022.105716.

315. Hine, E. & Floridi, L. New deepfake regulations in China are a tool for social stability, but at what cost? *Nature Machine Intelligence* **4**, 608–610 (2022)
www.doi.org/10.1038/s42256-022-00513-4.

316. Khowaja, S. A., Khowaja, P. & Dev, K. ChatGPT Needs SPADE (Sustainability, PrivAcy, Digital divide, and Ethics) Evaluation: A Review. *arXiv:2305.03123 [cs.CY]* (2023)
www.doi.org/10.48550/arXiv.2305.03123.

317. Hacker, P., Engel, A. & Mauer, M. Regulating ChatGPT and other Large Generative AI Models. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* 1112–1123 (Association for Computing Machinery, 2023).
www.doi.org/10.1145/3593013.3594067.

318. European Parliament. EU AI Act: first regulation on artificial intelligence.
https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence [29/09/2023].

319. Comisión Europea. *Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO POR EL QUE SE ESTABLECEN NORMAS ARMONIZADAS EN MATERIA DE INTELIGENCIA ARTIFICIAL (LEY DE INTELIGENCIA ARTIFICIAL) Y SE MODIFICAN DETERMINADOS ACTOS LEGISLATIVOS DE LA UNIÓN*. (2021).

320. Lawson, A. A Look at Global Deepfake Regulation Approaches. *RAI Institute*
https://www.responsible.ai/post/a-look-at-global-deepfake-regulation-approaches [17/09/2023].

321. Pavis, M. Rebalancing our regulatory response to Deepfakes with performers' rights. *Convergence* **27**, 974–998 (2021)
www.doi.org/10.1177/13548565211033418.

322. Lorenz-Spreen, P., Oswald, L., Lewandowsky, S. & Hertwig, R. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour* **7**, 74–101 (2023)
www.doi.org/10.1038/s41562-022-01460-1.

323. Lazer, D. M. J. *et al.* The science of fake news. *Science* **359**, 1094–1096 (2018)
www.doi.org/10.1126/science.aao2998.

324. Colley, T. P., Granelli, F. & Althuis, J. Disinformation's Societal Impact: Britain, COVID and Beyond. *Defence Strategic Communications* **8**, 89–140 (2020)
www.doi.org/10.30966/2018.RIGA.8.3.

325. Bastick, Z. Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation. *Computers in Human Behavior* **116**, 106633 (2021)
www.doi.org/10.1016/j.chb.2020.106633.

326. Bovet, A. & Makse, H. A. Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications* **10**, (2019)
www.doi.org/10.1038/s41467-018-07761-2.

327. Arcos, R., Gertrudix, M., Arribas, C. & Cardarilli, M. Responses to digital disinformation as part of hybrid threats: a systematic review on the effects of disinformation and the effectiveness of fact-checking/debunking. *Open Research Europe* (2022)
www.doi.org/10.12688/openreseurope.14088.1.

328. Ghanem, B., Rosso, P. & Rangel, F. An Emotional Analysis of False Information in Social Media and News Articles. *ACM Transactions on Internet Technology* **20**, 19:1-19:18 (2020)
www.doi.org/10.1145/3381750.

329. Rosso, P., Ghanem, B. & Giachanou, A. On the Impact of Emotions on the Detection of False Information. *2021 International Symposium on Electrical, Electronics and Information Engineering* 277–282 (Association for Computing Machinery, 2021).
www.doi.org/10.1145/3459104.3459150.

330. Flore, M., Balahur-Dobrescu, A., Podavini, A. & Verile, M. Understanding Citizens' Vulnerabilities to Disinformation and Data-Driven Propaganda. *JRC Publications Repository*
https://publications.jrc.ec.europa.eu/repository/handle/JRC116009 [30/05/2023]
www.doi.org/10.2760/919835.

331. Zimmermann, F. & Kohring, M. Mistrust, Disinforming News, and Vote Choice: A Panel Survey on the Origins and Consequences of Believing Disinformation in the 2017 German Parliamentary Election. *Political Communication* **37**, 215–237 (2020)
www.doi.org/10.1080/10584609.2019.1686095.

332. Sabatini, F. & Sarracino, F. Online Social Networks and Trust. *Social Indicators Research* **142**, 229–260 (2019)
www.doi.org/10.1007/s11205-018-1887-2.

333. Au, C. H., Ho, K. K. W. & Chiu, D. K. W. The Role of Online Misinformation and Fake News in Ideological Polarization: Barriers, Catalysts, and Implications. *Information Systems Frontiers* **24**, 1331–1354 (2022)
www.doi.org/10.1007/s10796-021-10133-9.

334. Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. & Westwood, S. J. The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science* **22**, 129–146 (2019)
www.doi.org/10.1146/annurev-polisci-051117-073034.

335. Schaub, M. & Morisi, D. Voter mobilisation in the echo chamber: Broadband internet and the rise of populism in Europe. *European Journal of Political Research* **59**, 752–773 (2020)
www.doi.org/10.1111/1475-6765.12373.

336. Tucker, J. *et al.* Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. *SSRN Electronic Journal* (2018)
www.doi.org/10.2139/ssrn.3144139.

337. Bronstein, M. V., Pennycook, G., Buonomano, L. & Cannon, T. D. Belief in fake news, responsiveness to cognitive conflict, and analytic reasoning engagement. *Thinking & Reasoning* **27**, 510–535 (2021)
www.doi.org/10.1080/13546783.2020.1847190.

338. Dabbous, A., Aoun Barakat, K. & de Quero Navarro, B. Fake news detection and social media trust: a cross-cultural perspective. *Behaviour & Information Technology* **41**, 2953–2972 (2022)
www.doi.org/10.1080/0144929X.2021.1963475.

339. Bak-Coleman, J. B. *et al.* Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences* **118**, e2025764118 (2021)
www.doi.org/10.1073/pnas.2025764118.

340. Lazer, D. M. J. *et al.* Computational social science: Obstacles and opportunities. *Science* **369**, 1060–1062 (2020)
www.doi.org/10.1126/science.aaz8170.

341. Lazer, D., Hargittai, E., Freelon, D., Gonzalez-Bailon, S., Munger, K., Ognyanova, K. & Radford, J. Meaningful measures of human society in the twenty-first century. *Nature* **595**, 189–196 (2021)
www.doi.org/10.1038/s41586-021-03660-7.

342. Green, Y., Gully, A., Roth, Y., Abhishek, R., Tucker, J. A. & Wanless, A. *Evidence-Based Misinformation Interventions: Challenges and Opportunities for Measurement and Collaboration*.
https://carnegieendowment.org/2023/01/09/evidence-based-misinformation-interventions-challenges-and-opportunities-for-measurement-and-collaboration-pub-88661 (2023).

343. Allcott, H., Braghieri, L., Eichmeyer, S. & Gentzkow, M. The Welfare Effects of Social Media. *American Economic Review* **110**, 629–676 (2020) www.doi.org/10.1257/aer.20190658.

344. European Commission. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing a common framework for media services in the internal market (European Media Freedom Act) and amending Directive 2010/13/EU*. (2022).

345. Aba-Catoira, A. Los desórdenes informativos en un sistema de comunicación democrático. *Revista Derecho Político* **1**, 119–151 (2020) www.doi.org/10.5944/rdp.109.2020.29056.

346. Departamento de Seguridad Nacional. Constitución del Foro contra las Campañas de Desinformación en el ámbito de la Seguridad Nacional. https://www.dsn.gob.es/es/actualidad/sala-prensa/constituci%C3%B3n-del-foro-contra-campa%C3%B1as-desinformaci%C3%B3n-%C3%A1mbito-seguridad [25/09/2023].

347. European Commission. *RAN C&N Working Group meeting - How to respond to disinformation in public communications from the perspective of frontline practitioners, 27-29 March 2023*. Migration and Home Affairs https://home-affairs.ec.europa.eu/whats-new/publications/ran-cn-working-group-meeting-how-respond-disinformation-public-communications-perspective-frontline_en (2023).

348. Bolt, N. *Strategic communications and disinformation in the early 21st century*. https://cadmus.eui.eu/handle/1814/74494 (2021).

349. European External Action Service. Tackling disinformation: Information on the work of the EEAS Strategic Communication division and its task forces (SG.STRAT.2). https://www.eeas.europa.eu/countering-disinformation/tackling-disinformation-information-work-eeas-strategic-communication_en [08/09/2023].

350. European Digital Media Observatory. EDMO at a Glance. https://edmo.eu/edmo-at-a-glance/ [22/09/2023].

351. Farkas, J. Fake News in Metajournalistic Discourse. *Journalism Studies* **24**, 423–441 (2023) www.doi.org/10.1080/1461670X.2023.2167106.

352. Gov UK. *The Cairncross Review: A sustainable future for journalism*. https://www.gov.uk/government/publications/the-cairncross-review-a-sustainable-future-for-journalism (2020).

353. Judit, B., Irini, K., Olga, B., Bernd, H., Sarah, H. & Katarzyna, L. *The fight against disinformation and the right to freedom of expression*. European Parliament's Committee on Civil Liberties, Justice and Home Affairs. PE 695.445 https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2021)695445 (2021).

354. European Media Freedom Act. *European Commission* https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5504 [22/09/2023].

355. Palau-Sampio, D., Carratalá, A., Tarullo, R. & Crisóstomo Flores, P. Reconocimiento de la calidad como prescriptor contra la desinformación. *Comunicar: Revista Científica de Comunicación y Educación* 59–70 (2022) www.doi.org/10.3916/C72-2022-05.

356. Cavaliere, P. From journalistic ethics to fact-checking practices: defining the standards of content governance in the fight against disinformation. *Journal of Media Law* **12**, 133–165 (2020) www.doi.org/10.1080/17577632.2020.1869486.

357. Stanton, L. Trusting News. *Trusting News* https://trustingnews.org/ [11/12/2023].

358. Lough, K. & McIntyre, K. A systematic review of constructive and solutions journalism research. *Journalism* **24**, 1069–1088 (2023) www.doi.org/10.1177/14648849211044559.

359. Aitamurto, T. & Varma, A. The Constructive Role of Journalism. *Journalism Practice* **12**, 695–713 (2018) www.doi.org/10.1080/17512786.2018.1473041.

360. McIntyre, K. Solutions Journalism. *Journalism Practice* **13**, 16–34 (2019) www.doi.org/10.1080/17512786.2017.1409647.

361. Li, Y. Assessing the Role Performance of Solutions Journalism in a Global Pandemic. *Journalism Practice* **17**, 1445–1464 (2023) www.doi.org/10.1080/17512786.2021.1990787.

362. EDMO – United against disinformation. https://edmo.eu/ [07/07/2023].

363. Iberifier | Iberian Digital Media Research and Fact-Checking Hub. https://iberifier.eu/ [25/09/2023].

364. Comisión Europea. Ley Europea de Libertad de los Medios de Comunicación. *European Commission - European Commission* https://ec.europa.eu/commission/presscorner/detail/es/ip_22_5504 [20/11/2023].

365. Cabrera Blázquez, F. J. *The proposal for a European Media Freedom Act*. European Audiovisual Observatory (2022).

366. EFE Verifica. ¿Qué es EFE Verifica? https://verifica.efe.com/que-es-efe-verifica/ [18/10/2023].

367. Maldita.es. Metodología de Maldito Bulo. *Maldita.es — Periodismo para que no te la cuelen* https://maldita.es/metodologia-de-maldito-bulo/ [18/10/2023].

368. Newtral. Metodología y transparencia. *Newtral* https://www.newtral.es/metodologia-transparencia/ [18/10/2023].

369. RTVE. Herramientas de Verificación de Bulos por internet de RTVE. *RTVE.es* https://www.rtve.es/noticias/verificartve/herramientas-de-verificacion/index.shtml [07/07/2023].

370. Barrera, O., Guriev, S., Henry, E. & Zhuravskaya, E. Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of Public Economics* **182**, 104123 (2020) www.doi.org/10.1016/j.jpubeco.2019.104123.

371. Ecker, U. K. H., O'Reilly, Z., Reid, J. S. & Chang, E. P. The effectiveness of short-format refutational fact-checks. *British Journal of Psychology* **111**, 36–54 (2020) www.doi.org/10.1111/bjop.12383.

372. Brashier, N. M., Pennycook, G., Berinsky, A. J. & Rand, D. G. Timing matters when correcting fake news. *Proceedings of the National Academy of Sciences* **118**, e2020043118 (2021) www.doi.org/10.1073/pnas.2020043118.

373. Maldita.es. Qué es pre-bunking y cómo se lucha contra la desinformación antes del desmentido. *Maldita.es — Periodismo para que no te la cuelen* https://maldita.es/nosotros/20230323/prebunking-que-es-antes-desmentido/ [18/09/2023].

374. Rojas Caja, F. El fact checking. Las agencias de verificación de noticias en España. *bie3: Boletín IEEE* 1492–1505 (2020).

375. Maldita.es. Políticas Públicas. https://maldita.es/politicas-publicas-desarrollo-institucional/ [19/09/2023].

376. Roozenbeek, J. & Van der Linden, S. *Inoculation Theory and Misinformation*. NATO Strategic Communications Centre of Excellence https://stratcomcoe.org/publications/inoculation-theory-and-misinformation/217 (2021).

377. Markowitz, D. M., Levine, T. R., Serota, K. B. & Moore, A. D. Cross-checking journalistic fact-checkers: The role of sampling and scaling in interpreting false and misleading statements. *PLOS ONE* **18**, e0289004 (2023) www.doi.org/10.1371/journal.pone.0289004.

378. Bavel, J. J. V. & Pereira, A. The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences* **22**, 213–224 (2018) www.doi.org/10.1016/j.tics.2018.01.004.

379. Vraga, E. K., Ecker, U. K. H., Žeželj, I., Lazić, A. & Azlan, A. A. To Debunk or Not to Debunk? Correcting (Mis)Information. *Managing Infodemics in the 21st Century: Addressing New Public Health Challenges in the Information Ecosystem* 85–98 (2023). ISBN: 978-3-031-27789-4.

380. Porter, E. & Wood, T. J. The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. *Proceedings of the National Academy of Sciences* **118**, e2104235118 (2021) www.doi.org/10.1073/pnas.2104235118.

381. Kyriakidou, M., Cushion, S., Hughes, C. & Morani, M. Questioning Fact-Checking in the Fight Against Disinformation: An Audience Perspective. *Journalism Practice* **0**, 1–17 (2022) www.doi.org/10.1080/17512786.2022.2097118.

382. Walter, N., Cohen, J., Holbert, R. L. & Morag, Y. Fact-Checking: A Meta-Analysis of What Works and for Whom. *Political Communication* **37**, 350–375 (2020) www.doi.org/10.1080/10584609.2019.1668894.

383. Porter, E. & Wood, T. J. Political Misinformation and Factual Corrections on the Facebook News Feed: Experimental Evidence. *The Journal of Politics* **84**, 1812–1817 (2022) www.doi.org/10.1086/719271.

384. Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I. & Erviti, M.-C. Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19. *Profesional de la información* **29**, (2020) www.doi.org/10.3145/epi.2020.may.15.

385. Carey, J. M., Guess, A. M., Loewen, P. J., Merkley, E., Nyhan, B., Phillips, J. B. & Reifler, J. The ephemeral effects of fact-checks on COVID-19 misperceptions in the United States, Great Britain and Canada. *Nature Human Behaviour* **6**, 236–243 (2022) www.doi.org/10.1038/s41562-021-01278-3.

386. European Media and Information Fund. Boosting Fact-Checking Activities in Europe. https://gulbenkian.pt/emifund/bolsas-lista/boosting-fact-checking-activities-in-europe/ [22/09/2023].

387. McIlhiney, P., Gignac, G. E., Ecker, U. K. H., Kennedy, B. L. & Weinborn, M. Executive function and the continued influence of misinformation: A latent-variable analysis. *PLoS One* **18**, (2023) www.doi.org/10.1371/journal.pone.0283951.

388. Calabrese, C. & Albarracín, D. Bypassing misinformation without confrontation improves policy support as much as correcting it. *Scientific Reports* **13**, 6005 (2023) www.doi.org/10.1038/s41598-023-33299-5.

389. The European Fact-Checking Standards Network. https://eufactcheckingproject.com/ [30/09/2023].

390. IFCN Code of Principles. https://ifcncodeofprinciples.poynter.org/ [07/07/2023].

391. European Digital Media Observatory. Fact-checking EDMO. https://edmo.eu/fact-checking/ [07/07/2023].

392. Allen, J., Arechar, A. A., Pennycook, G. & Rand, D. G. Scaling up fact-checking using the wisdom of crowds. *Science Advances* **7**, eabf4393 (2021) www.doi.org/10.1126/sciadv.abf4393.

393. Li, J. & Chang, X. Combating Misinformation by Sharing the Truth: a Study on the Spread of Fact-Checks on Social Media. *Information Systems Frontiers* 1–15 (2022) www.doi.org/10.1007/s10796-022-10296-z.

394. Vraga, E. K., Bode, L. & Tully, M. Creating News Literacy Messages to Enhance Expert Corrections of Misinformation on Twitter. *Communication Research* **49**, 245–267 (2022) www.doi.org/10.1177/0093650219898094.

395. Nissen, I. A., Walter, J. G., Charquero-Ballester, M., Bechmann, A. & DATALAB Aarhus University. *A method for auditing fact-checking tools and databases*. NORDIS https://www.tjekdet.dk/files/2022-05/AAReport%20task%201.1%20-%20A%20method%20for%20auditing%20fact-checking%20tools%20and%20databases.pdf (2022).

396. Zeng, X., Abumansour, A. S. & Zubiaga, A. Automated fact-checking: A survey. *Language and Linguistics Compass* **15**, e12438 (2021) www.doi.org/10.1111/lnc3.12438.

397. Das, A., Liu, H., Kovatchev, V. & Lease, M. The state of human-centered NLP technology for fact-checking. *Information Processing & Management* **60**, 103219 (2023) www.doi.org/10.1016/j.ipm.2022.103219.

398. Lazarski, E., Al-Khassaweneh, M. & Howard, C. Using NLP for Fact Checking: A Survey. *Designs* **5**, 42 (2021) www.doi.org/10.3390/designs5030042.

399. Ortega, J. Periodismo e inteligencia artificial: los avances de Newtral. https://www.newtral.es/periodismo-inteligencia-artificial-avances-newtral/20220624/ [18/10/2023].

400. Giachanou, A., Zhang, G. & Rosso, P. Multimodal fake news detection with textual, visual and semantic information. *Lecture Notes in Computer Science* **12284 LNAI**, 30–38 (2020) www.doi.org/10.1007/978-3-030-58323-1_3.

401. Zhou, X. & Zafarani, R. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Computing Surveys* **53**, 109:1-109:40 (2020) www.doi.org/10.1145/3395046.

402. IBERIFIER & Iberian Media Research and Fact-Checking. *Is the 'AI toolbox for disinformation' ready?* https://iberifier.eu/app/uploads/2023/06/202303-IBERIFIER-Report-Is-the-%E2%80%98AI-toolbox-for-disinformation-ready.pdf (2023).

403. Linden, C. G., Dang-Nguyen, D. T., Salas-Gulliksen, C., Khan, S. A., Amelie, M. & Dierickx, L. State of the art in fact-checking technology. *NORDIS - NORdic observatory for digital media and information DISorders* (2022).

404. Oficina de Ciencia y Tecnología del Congreso de los Diputados (Oficina C). Informe C: Inteligencia artificial y salud. (2022) www.doi.org/10.57952/tcsx-b678.

405. *JOINT COMMUNICATION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Action Plan against Disinformation*. (2018).

406. European Commission. Media Literacy Guidelines | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/library/media-literacy-guidelines [24/09/2023].

407. Buckingham, D. Epilogue: Rethinking digital literacy: Media education in the age of digital capitalism. *Digital Education Review* 230–239 (2020).

408. Buckingham, D. Teaching media in a 'post-truth' age: fake news, media bias and the challenge for media/digital literacy education. *Culture and Education* **31**, 213–231 (2019) www.doi.org/10.1080/11356405.2019.1603814.

409. Council of Europe. Supporting Quality Journalism through Media and Information Literacy. *Freedom of Expression* https://www.coe.int/en/web/freedom-expression/-/supporting-quality-journalism-through-media-and-information-literacy [24/09/2023].

410. Goodman, E. *Media literacy in Europe and the role of EDMO.* https://edmo.eu/wp-content/uploads/2022/02/Media-literacy-in-Europe-and-the-role-of-EDMO-Report-2021.pdf (2021).

411. NAMLE. Media Literacy Defined. https://namle.net/resources/media-literacy-defined/ [30/10/2023].

412. McDougall, J., Zezulkova, M. & van Driel, B. Teaching media literacy in Europe: evidence of effective school practices in primary and secondary education. *NESET* https://nesetweb.eu/en/resources/library/teaching-media-literacy-in-europe-evidence-of-effective-school-practices-in-primary-and-secondary-education/ [24/09/2023].

413. Dumitru, E.-A., Ivan, L. & Loos, E. A Generational Approach to Fight Fake News: In Search of Effective Media Literacy Training and Interventions. *Human Aspects of IT for the Aged Population. Design, Interaction and Technology Acceptance* (eds. Gao, Q. & Zhou, J.) 291–310 (Springer International Publishing, 2022). www.doi.org/10.1007/978-3-031-05581-2_22.

414. Lee, N. M. Fake news, phishing, and fraud: a call for research on digital media literacy education beyond the classroom. *Communication Education* **67**, 460–466 (2018) www.doi.org/10.1080/03634523.2018.1503313.

415. Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J. & Sircar, N. A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences* **117**, 15536–15545 (2020) www.doi.org/10.1073/pnas.1920498117.

416. Jones-Jang, S. M., Mortensen, T. & Liu, J. Does Media Literacy Help Identification of Fake News? Information Literacy Helps, but Other Literacies Don't. *American Behavioral Scientist* **65**, 371–388 (2021) www.doi.org/10.1177/0002764219869406.

417. Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D. & Rand, D. G. Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595 (2021) www.doi.org/10.1038/s41586-021-03344-2.

418. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science* **31**, 770–780 (2020) www.doi.org/10.1177/0956797620939054.

419. Literat, I., Abdelbagi, A., Law, N. Y., Cheung, M. Y.-Y. & Tang, R. Research note: Likes, sarcasm and politics: Youth responses to a platform-initiated media literacy campaign on social media. *Harvard Kennedy School Misinformation Review* (2021) www.doi.org/10.37016/mr-2020-67.

420. Vraga, E., Tully, M. & Bode, L. Assessing the relative merits of news literacy and corrections in responding to misinformation on Twitter. *New Media & Society* **24**, 2354–2371 (2022) www.doi.org/10.1177/1461444821998691.

421. Herrero-Curiel, E. & La-Rosa, L. Los estudiantes de secundaria y la alfabetización mediática en la era de la desinformación. *Comunicar: Revista Científica de Comunicación y Educación* **73**, 95–106 (2022) www.doi.org/10.3916/C73-2022-08.

422. Herrero Curiel, E. & La- Rosa Barrolleta, L. Cultura, economía y educación: nuevos desafíos en la sociedad digital. *La alfabetización mediática en secundaria: transversalidad y voluntad* 76–95 (Dykinson S.L., 2021). ISBN: 978-84-13-77585-2.

423. Bernabeu Morón, N., Esteban Ruiz, N., Gallego Hernández, L. & Rosales Páez, A. *Alfabetización mediática y competencias básicas. Proyecto Mediascopio Prensa La lectura de la prensa escrita en el aula.* Ministerio de Educación. Instituto de Formación del Profesorado, Investigación e Innovación Educativa (IFIIE) https://sede.educacion.gob.es/publiventa/alfabetizacion-mediatica-y-competencias-basicas-proyecto-mediascopio-prensa-la-lectura-de-la-prensa-escrita-en-el-aula/ensenanza-tecnologias-de-la-informacion-prensa/14484 (2011).

424. Ministerio de Educación, Formación Profesional y Deportes. Desinformación y alfabetización mediática e informacional. https://www.educacionyfp.gob.es/gl/mc/sgctie/comunicacion/blog/2020/octubre2020/alfabetizacion-mediatica.html [30/10/2023].

425. Herrero Curiel, E. & La-Rosa, L. *Estudio de alfabetización mediática en centros de Educación Secundaria Obligatoria*. Ministerio de Educación y Formación Profesional https://sede.educacion.gob.es/publiventa/d/26772/19/1 (2023).

426. European Commison. Reporting on Media Literacy in Europe | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/library/reporting-media-literacy-europe [24/09/2023].

427. Tsourapas, E. Media, Information and Digital Literacy Organisations in Europe. *EAVI* https://eavi.eu/media-information-digital-literacy-organisations-in-europe/ [24/09/2023].

428. Council of Europe. Dealing with propaganda, misinformation and fake news. *Democratic Schools for All* https://www.coe.int/en/web/campaign-free-to-speak-safe-to-learn/dealing-with-propaganda-misinformation-and-fake-news [25/10/2023].

429. European Regulators Group for Audiovisual Media Services. *Improving Media Literacy campaigns on disinformation (ERGA Report).* https://erga-online.eu/wp-content/uploads/2021/01/ERGA-SG2-Report-2020-Improving-Media-Literacy-campaigns-on-disinformation.pdf .

430. Orosz, G., Paskuj, B., Faragó, L. & Krekó, P. A prosocial fake news intervention with durable effects. *Scientific Reports* **13**, 3958 (2023) www.doi.org/10.1038/s41598-023-30867-7.

431. EU-Citizen.Science. The News Evaluator. https://eu-citizen.science/project/3 [18/10/2023].

432. Maldita.es. Educación: la caja de herramientas de verificación para que no te la cuelen. *Maldita.es — Periodismo para que no te la cuelen* https://maldita.es/malditobulo/20181128/educacion-la-caja-de-herramientas-de-verificacion-para-que-no-te-la-cuelen/ [18/09/2023].

433. CAPCIT. Consell Assessordel Parlament sobre Ciència i Tecnologia. *Desinformació a les xarxes socials: Què és i com identificar-la.* https://www.parlament.cat/document/intrade/267188726 (2022).

434. Arroyo, D. & Degli Esposti, S. *Como protegerme de la desinformación.* Ministerio de Educación y Formación Profesional (2022).

435. Jefatura del Estado. *Ley 13/2022, de 7 de julio, General de Comunicación Audiovisual*. vol. BOE-A-2022-11311 96114–96220 (2022).

436. Konstantinou, L., Caraban, A. & Karapanos, E. Combating Misinformation Through Nudging. *Human-Computer Interaction – INTERACT 2019* (eds. Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M. & Zaphiris, P.) 630–634 (Springer International Publishing, 2019). www.doi.org/10.1007/978-3-030-29390-1_51.

437. Sartori, R., Tommasi, F., Ceschi, A., Falser, M., Genero, S. & Belotto, S. Enhancing critical thinking skills and media literacy in initial vocational education and training via self-nudging: The contribution of NERDVET project. *Frontiers in Psychology* **13**, (2022) www.doi.org/10.3389/fpsyg.2022.935673.

438. de Freitas Melo, P., Vieira, C. C., Garimella, K., de Melo, P. O. S. V. & Benevenuto, F. Can WhatsApp Counter Misinformation by Limiting Message Forwarding? *Complex Networks and Their Applications VIII* (eds. Cherifi, H., Gaito, S., Mendes, J. F., Moro, E. & Rocha, L. M.) 372–384 (Springer International Publishing, 2020). www.doi.org/10.1007/978-3-030-36687-2_31.

439. Maldita.es. Desinformación en WhatsApp: el chatbot de Maldita.es y el atributo 'Reenviado Frecuentemente' . https://maldita.es/nosotros/20210603/desinformacion-whatsapp-chatbot-frequently-forwarded-reenviado-frecuentemente/ [21/10/2023].

440. Arroyo, D., Degli-Esposti, S., Gómez-Espés, A., Palmero-Muñoz, S. & Pérez-Miguel, L. On the Design of a Misinformation Widget (MsW) Against Cloaked Science. *Network and System Security* (eds. Li, S., Manulis, M. & Miyaji, A.) 385–396 (Springer Nature Switzerland, 2023). www.doi.org/10.1007/978-3-031-39828-5_21.

441. Banas, J. A. & Rains, S. A. A Meta-Analysis of Research on Inoculation Theory. *Communication Monographs* **77**, 281–311 (2010) www.doi.org/10.1080/03637751003758193.

442. Lewandowsky, S. & van der Linden, S. Countering Misinformation and Fake News Through Inoculation and Prebunking. *European Review of Social Psychology* **32**, 348–384 (2021) www.doi.org/10.1080/10463283.2021.1876983.

443. van der Linden, S., Leiserowitz, A., Rosenthal, S. & Maibach, E. Inoculating the Public against Misinformation about Climate Change. *Global Challenges* **1**, 1600008 (2017) www.doi.org/10.1002/gch2.201600008.

444. van der Linden, S., Maibach, E., Cook, J., Leiserowitz, A. & Lewandowsky, S. Inoculating against misinformation. *Science* **358**, 1141–1142 (2017) www.doi.org/10.1126/science.aar4533.

445. Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P. & Linden, S. van der. Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data & Society* **8**, 20539517211013868 (2021) www.doi.org/10.1177/20539517211013868.

446. Schmid, P. & Betsch, C. Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour* **3**, 931–939 (2019) www.doi.org/10.1038/s41562-019-0632-4.

447. Maldita.es. Falacias lógicas que pueden hacer que te la cuelen: aprende a identificarlas. *Maldita.es — Periodismo para que no te la cuelen* https://maldita.es/malditateexplica/20211213/falacias-tipos-discusion-argumentos/ [19/10/2023].

448. Cook, J. Deconstructing climate science denial. *Research Handbook on Communicating Climate Change* 62–78 (Edward Elgar Publishing, 2020). ISBN: 978-1-78990-040-8.

449. McPhedran, R., Ratajczak, M., Mawby, M., King, E., Yang, Y. & Gold, N. Psychological inoculation protects against the social media infodemic. *Scientific Reports* **13**, 5780 (2023) www.doi.org/10.1038/s41598-023-32962-1.

450. Cook, J. *et al.* The cranky uncle game—combining humor and gamification to build student resilience against climate misinformation. *Environmental Education Research* **29**, 607–623 (2023) www.doi.org/10.1080/13504622.2022.2085671.

451. Roozenbeek, J. & van der Linden, S. Fake news game confers psychological resistance against online misinformation. *Palgrave Communications* **5**, 1–10 (2019) www.doi.org/10.1057/s41599-019-0279-9.

452. Butler, L. H. *et al.* The (Mis)Information Game: A social media simulator. *Behavior Research Methods* (2023) www.doi.org/10.3758/s13428-023-02153-x.

453. Roozenbeek, J., Traberg, C. S. & van der Linden, S. Technique-based inoculation against real-world misinformation. *Royal Society Open Science* **9**, 211719 (2022) www.doi.org/10.1098/rsos.211719.

454. Lewsey, F. How to 'inoculate' millions against misinformation on social media. *University of Cambridge* https://www.cam.ac.uk/stories/inoculateexperiment [24/09/2023].

455. Swire-Thompson, B., Cook, J., Butler, L. H., Sanderson, J. A., Lewandowsky, S. & Ecker, U. K. H. Correction format has a limited role when debunking misinformation. *Cognitive Research: Principles and Implications* **6**, (2021) www.doi.org/10.1186/s41235-021-00346-6.

456. Lewandowsky, S. *et al. The Debunking Handbook 2020*. (DOI:10.17910/b7.1182, 2020).

457. Paynter, J. *et al.* Evaluation of a template for countering misinformation—Real-world Autism treatment myth debunking. *PLOS ONE* **14**, e0210746 (2019) www.doi.org/10.1371/journal.pone.0210746.

458. Casero-Ripollés, A., Tuñón, J. & Bouza-García, L. The European approach to online disinformation: geopolitical and regulatory dissonance. *Humanities and Social Sciences Communications* **10**, 1–10 (2023) www.doi.org/10.1057/s41599-023-02179-8.

459. StratCom | NATO Strategic Communications Centre of Excellence Riga, Latvia. https://stratcomcoe.org/ [07/03/2022].

460. The European Centre of Excellence for Countering Hybrid Threats. Hybrid CoE. https://www.hybridcoe.fi/ [27/09/2023].

461. Naciones Unidas. A/HRC/47/25: La desinformación y la libertad de opinión y de expresión Informe de la Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, Irene Khan. *OHCHR* https://www.ohchr.org/es/documents/reports/disinformation-and-freedom-opinion-and-expression-report-special-rapporteur [06/07/2023].

462. Diario Oficial de la Unión Europea. *Carta de los Derechos Fundamentales de la Unión Europea*. (2016).

463. Constitución Española. *Título I. De los derechos y deberes fundamentales*. https://app.congreso.es/consti/constitucion/indice/titulos/articulos.jsp?ini=20&tipo=2 (1978).

464. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. *Orden PCM/1030/2020, de 30 de octubre, por la que se publica el Procedimiento de actuación contra la desinformación aprobado por el Consejo de Seguridad Nacional*. vol. BOE-A-2020-13663 96673–96680 (2020).

465. Transparency Centre. Reports Archive. https://disinfocode.eu/reports-archive/ [27/09/2023].

466. European Commison. Code of Practice on Disinformation: new reports available in the Transparency Centre | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/news/code-practice-disinformation-new-reports-available-transparency-centre [27/09/2023].

467. Corredoira y Alfonso, L. Como dijera cicerón, no todo lo molesto es delito: (pese al art. 510 del código penal). *OTROSÍ.: Revista del Colegio de Abogados de Madrid* 28–31 (2021).

468. Tribunal de Cuentas Europeo. *El impacto de la desinformación en la UE: una cuestión abordada, pero no atajada*. https://www.eca.europa.eu/es/publications?did=58682 (2021).

469. *Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement | Shaping Europe's digital future*. https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement (2020).

470. Plasilova, I., Hill, J., Carlberg, M., Goubet, M. W. & Procee, R. Study for the "Assessment of the implementation of the Code of Practice on Disinformation". *Comisión Europea* (2020) www.doi.org/10.2759/188091.

471. European Commison. Code of Practice on Disinformation | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/news/code-practice-disinformation [05/06/2023].

472. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*. (2020).

473. European Union External Action. Factsheet: Rapid Alert System | EEAS. https://www.eeas.europa.eu/node/59644_en [07/07/2023].

474. *COMUNICACIÓN CONJUNTA AL PARLAMENTO EUROPEO, AL CONSEJO EUROPEO, AL CONSEJO, AL COMITÉ ECONÓMICO Y SOCIAL EUROPEO Y AL COMITÉ DE LAS REGIONES La lucha contra la desinformación acerca de la COVID-19: contrastando los datos*. (2020).

475. *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS European Commission Guidance on Strengthening the Code of Practice on Disinformation*. (2021).

476. *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance). OJ L* vol. 277 (2022).

477. European Parliament. EU Digital Markets Act and Digital Services Act explained. https://www.europarl.europa.eu/news/en/headlines/society/20211209STO19124/eu-digital-markets-act-and-digital-services-act-explained [25/09/2023].

478. European Commission. State of the Union 2018: European Commission proposes measures for securing free and fair European elections. https://ec.europa.eu/commission/presscorner/detail/en/IP_18_5681 [26/09/2023].

479. Parlamento Europeo. Textos aprobados - Injerencias extranjeras en todos los procesos democráticos de la Unión Europea. https://www.europarl.europa.eu/doceo/document/TA-9-2022-0064_ES.html [26/09/2023].

480. Parlamento Europeo. Comisión Especial sobre Injerencias Extranjeras en Todos los Procesos Democráticos de la Unión Europea, en particular la Desinformación, y sobre el Refuerzo de la Integridad, la Transparencia y la Rendición de Cuentas en el Parlamento Europeo. ING2. https://www.europarl.europa.eu/committees/es/ing2/home/highlights [26/09/2023].

481. République Française. *LOI organique n° 2018-1201 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information (1).*

482. *Alemania (2017). Network Enforcement Act (Netzdurchsetzunggesetz, NetzDG) de 1 de septiembre (Federal Law Gazette I, p. 3352), https:// germanlawarchive.iuscomp.org/?p=1245.*

483. Departamento de Seguridad Nacional. Gobierno de España. *Estrategia Nacional de Ciberseguridad 2019.* https://www.dsn.gob.es/es/documento/estrategia-nacional-ciberseguridad-2019 (2019).

484. Departamento de Seguridad Nacional. *Informe Anual de Seguridad Nacional 2021.* https://www.dsn.gob.es/es/documento/informe-anual-seguridad-nacional-2021 (2022).

485. Presidencia del Gobierno. *Informe Anual de Seguridad Nacional 2020.* https://www.dsn.gob.es/es/documento/informe-anual-seguridad-nacional-2020 (2021).

486. Saltz, E., Barari, S., Leibowicz, C. & Wardle, C. Misinformation interventions are common, divisive, and poorly understood. *Harvard Kennedy School Misinformation Review* (2021) www.doi.org/10.37016/mr-2020-81.

487. La Moncloa. Interior activa la Red de Coordinación para la Seguridad en Procesos Electorales para el 23J. https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/interior/paginas/2023/040723-red-coordinacion-seguridad-elecciones.aspx [26/09/2023].

488. Government of Canada, P. S. and P. C. *Political communications in the digital age. Discussion paper 1: the regulation of political communications under the Canada Elections Act.* https://publications.gc.ca/site/eng/9.886736/publication.html (2020).

489. Elections Canada. Registry Requirements for Political Ads on Online Platforms. https://www.elections.ca/content.aspx?section=pol&dir=regifaq&document=index&lang=e [26/09/2023].

490. Senate of the United States. S.1989 - Honest Ads Act. 115th Congress (2017-2018). https://www.congress.gov/bill/115th-congress/senate-bill/1989/text [26/09/2023].

491. Electoral Commission in New Zealand. About election advertising. *Elections* https://elections.nz/guidance-and-rules/advertising-and-campaigning/about-election-advertising/ [26/09/2023].

492. Furnémont, J. F. & Deirdre, K. *Regulation of Political advertising - A comparative study with reflections on the situation in South-East Europe -*. Council of Europe https://rm.coe.int/study-on-political-advertising-eng-final/1680a0c6e0 (2020).

493. Assemblée Nationale. *Proposition de loi n°419 - 15e législature visant à lutter contre les contenus haineux sur internet*. (2020).

494. Government of Ireland. Government publishes first Report of the Interdepartmental Group on security of Ireland's Electoral Process and Disinformation. https://www.gov.ie/ga/preasraitis/37e936-government-publishes-first-report-of-the-interdepartmental-group-on-/ [26/09/2023].

495. Gov.UK. Transparency in digital campaigning: technical consultation on digital imprints. https://www.gov.uk/government/consultations/transparency-in-digital-campaigning-technical-consultation-on-digital-imprints [26/09/2023].