CCN-CERT

BP/13

Disinformation in Cyberspace

BEST PRACTICE REPORT

AUGUST **2021**

ccn-cert centro criptológico nacional

CCN centro criptológico nacional

Edit:

Date of Publication: August 2021

LIMITATION OF LIABILITY

This document is provided in accordance with the terms set forth herein, expressly disclaiming any implied warranties of any kind that may be found to be related. In no event shall the National Cryptologic Centre be held responsible for direct, indirect, incidental or extraordinary damage derived from the use of the information and software indicated, even when such possibility is warned of.

LEGAL NOTICE

Index

1. **About CCN-CERT, National Governmental Cert** 4

2. **Introduction** 5

3. **Objective of the guide** 10

4. **What is the risk in Spain?** 11

5. **What are the consequences of a disinformation attack?** 15
   - 5.1 Loss of confidence in traditional media 16
   - 5.2 Loss of confidence in public institutions 20
   - 5.3 Loss of confidence in the citizen's sovereignty 21
   - 5.4 Social polarization 23

6. **The methodology of disinformation campaigns** 25

7. **The 10 key elements of a disinformation campaign** 28
   - 7.1 Fake news 29
     - 7.1.1 Deep fake news 33
   - 7.2 The approach 33
   - 7.3 New media 35
   - 7.4 Social forums 37
   - 7.5 Malicious digital profiles 37
   - 7.6 Automated accounts of non-human behaviour 39
   - 7.7 Digital coverages or hybrid accounts 41
   - 7.8 The stars invited 44
   - 7.9 Algorithms, resonance chambers and trusted networks 45
   - 7.10 Paid advertisements 46

8. **Decalogue of recomendations** 48

1. About CCN-CERT

The CCN-CERT is the Computer Emergency Response Team of the National Cryptologic Centre, CCN, within the National Intelligence Centre, CNI. This service was created in 2006 as a **Spanish National Governmental CERT** and its functions are included in Law 11/2002 regulating the CNI, RD 421/2004 regulating the CCN and RD 3/2010, dated 8th January, regulating the National Security Scheme (ENS), modified by RD 951/2015 of 23rd October.

Its mission therefore is to contribute to the improvement of Spanish cybersecurity, being the national alert and response centre that cooperates and helps to respond quickly and efficiently to cyberattacks and to actively confront cyber threats, including the coordination at the national public level of the different Incident Response Teams or existing Security Operations Centres.

Its ultimate aim is to make cyberspace more secure and reliable, preserving classified information (as stated in Article 4.F of Law 11/2002) and sensitive information, defending Spanish Technological Heritage, training expert personnel, applying security policies and procedures and using and developing the most appropriate technologies for this purpose.

In accordance with these regulations and Law 40/2015 on the Legal Regulation for the Public Sector, the CCN-CERT is responsible for the management of cyber incidents affecting any public body or company. In the case of critical operators in the public sector, the management of cyber incidents will be carried out by the CCN-CERT in coordination with the CNPIC.

The CCN-CERT is the Computer Emergency Response Team of the National Cryptologic Centre.

2. Introduction

Disinformation as a cyber-attack is designed to erode and weaken the internal cohesion of a state.

Spain suffers three (3) critical or very high danger cyberattacks against the public sector and strategic companies on a daily basis[1]. Some of these offensive actions have their origin mainly in other states whose purposes include weakening and compromising Spain's economic, technological and political capacity in an increasingly complex, competitive and globalised world.

The consequences of these attacks can cause millionaire losses in private companies and compromise for minutes, hours or days the normal functioning of essential public services for Spanish citizens.

Sometimes the damage caused by digital offensive actions is limited to economic and material losses which in the medium and long term can be remedied. In recent years, both large companies and strategic state institutions have developed plans and protocols that allow them to effectively avoid, manage and minimize the consequences of possible attacks against their systems and infrastructures; part of these plans is dedicated to recovering operations as quickly and efficiently as possible: this is known as **cyber resilience.**

Nevertheless, there are increasingly being more attacks against a country's interests through the cyberspace which do not consist of modifying the computer systems of companies and institutions but are aimed at altering the functioning of one of the main elements of the development of a liberal democracy and a modern nation-state: **the public opinion.**

1. CCN-CERT. National Cryptologic Cent Incident Response Capacity (CCN-CERT)

2. Introduction

Thinkers and philosophers, such as Jürgen Habermas, argue that a public sphere based on rational deliberation is the cornerstone of democracy[2]. Hence the consequences of a cyberattack aimed at eroding the public opinion of a country cannot go unnoticed and, if successfully carried out, its damages would not be limited to economic or material losses but could erode the nature and raison d'être of a system of government based on a liberal democracy, affecting the factors that provide integrity to a nation-state.

Communication used as a weapon of war is anything but new. There have been references to the use of communicative wars in military contexts for more than 2,500 years. In fact, the Chinese general Sun Tzu, born in 544 BC, wrote that "the art of war is deceit"[3].

The figure of Joseph Goebbles, Minister for Public Enlightenment and Propaganda of the Third Reich between 1933 and 1945, is already being studied in the faculties of Information Sciences around the world as the greatest representative of the use of propaganda as a weapon of war. The postulates of Hitler and his minister set the guidelines for communication and disinformation, even before the armed conflict: *"It is indispensable to demoralize the enemy nation, to prepare it to capitulate, to constrain it morally to passivity, even before planning any military action... We will not hesitate to promote revolutions in enemy land".*[4]

With the passage of time, the propaganda and disinformation techniques have been improved on all fields and countries. Nonetheless, the technological revolution which has taken place at a global scale over the last ten years has led to an exponential increase in these actions, both in magnitude and in frequency and effectiveness. And it has accomplished this not only through traditional methods of disinformation, but also by using low-cost automatic dissemination tools with complex traceability which considerably increases their consequences and impact.

Those responsible for these attacks are usually governments and subnational groups whose objective is to **erode and weaken the internal cohesion of a state or a group of states** considered as adversaries and, in this way, redefine their geo-strategic position. In fact, some countries already openly acknowledge that they are systematically carrying out and undertaking this type of action. In this sense, Russia has been one of the countries which most has developed the concept of hybrid war

2. MACNAMARA, Jim. 2016. Organizational Listening. New York. Peter Lang. P. 10
3. SUN TZU (s. V a. C. /2013). The art of war. Ed. Medí. S.f.
4. RAUSCHINING, H, 1940. Hitler told me. Ed. Atlas. Page 11

or, as it is stated by the Russian military doctrine, "undeclared wars" and "non-linear wars"[5].

According to General Valery Gerasimov, head of the General Staff of the Armed Forces of the Russian Federation, "a perfectly thriving state can, in a matter of months and even days, be transformed into an arena of fierce armed conflict, become a victim of foreign intervention, and sink into a web of chaos, humanitarian catastrophe, and civil war".

In the words of this Russian General, promotor of the so-called "Gerasimov Doctrine", which is based on the study of the Arab Springs, considers that: "The role of non-military means of achieving political and strategic goals has grown, and, in many cases, they have exceeded the power of force of weapons in their effectiveness. The focus of applied methods of conflict has altered in the direction of the broad use of political, economic, informational, humanitarian, and other non-military measures — applied in coordination with the protest potential of the population. All this is supplemented by military means of a concealed character, including carrying out actions of informational conflict and the actions of special-operations forces[6]."

There are at least six (6) factors that contribute to the increase of the ever more recurrent use of hostile actions based on the distribution of disinformation:

1. **High level of effectiveness.** The technological revolution has allowed the democratization of accessing the media and the technology for producing informative messages. Currently, it is relatively inexpensive and easy to produce multimedia messages of high technical quality and disseminate them directly and effectively to the appropriate audiences.

   In a matter of a few days, it is possible to create websites and multimedia communication platforms with the same appearance and professional quality as other media with centenary trajectories, or to manipulate photographs with editing programs that are easy to access and use.

   Similarly, with limited resources, even with a mobile device, it is possible to massively transmit live content or generate videos with artificially manipulated images, which will be disseminated through the Internet and social networks.

5. See the paper "Hybrid Threats: New Tools for Old Aspirations". Carlos Galán, PhD. Real Instituto Elcano (2018).
6. COALSON, Robert. "Top Russian General Lays Bare Putin´s Plan for Ukraine" [online], in *Huffington Post*. S.f. Available at: https://www.huffingtonpost.com/robert-coalson/valery-gerasimov-putin-ukraine_b_5748480.html.

2

Difficulty in establishing direct attribution. One of the main characteristics of disinformation campaigns is to generate confusion, both through messages and sources.

Digital platforms and the very nature of the Internet favour the emergence of anonymous actors who maliciously influence the shaping of public opinion. Anonymous digital profiles, programs to automate the distribution of messages, IP address concealment software, technology to create new media which imitate the appearance of consolidated communication companies...

Any Internet user today has at his fingertips the technology needed to create potentially influential communication networks, which make the traceability of information and its source of origin very complex.

Therefore, it is a l**aborious and difficult task to substantiate in evidence a direct accusation** or to bring legal charges or take coercive action against a country or subnational group accused of initiating a communication war.

3

Complex regulation. Unlike other offensive actions, such as open warfare on a battlefield, terrorist actions or digital hacking, **disinformation and manipulation actions aiming to affect the public opinion are not easy to combat** from the legal perspective of liberal democracies.

Freedom of expression and opinion are fundamental principles in a democratic nation-state and it is often unfeasible, by democratic principle, to limit these rights to both national and foreign citizens.

It is not illegal to create a means of communication which disseminates non-contrasted information, just as it is not a crime to manage several anonymous accounts on social networks, or to create subsidiaries or media associated with groups linked to foreign governments. The actors, both state and non-state, that currently carry out this type of operations are aware of and take advantage of the limitations and contradictions posed by the aforementioned regulations.

4

Limitation for establishing a causal relationship. Current technical methodologies make it possible to detect disinformation attempts and attribute them, with a greater or lesser degree of certainty, to certain national or sub-national agents with the intention of maliciously conditioning the public debate in a state.

However, it is still **very difficult to prove a causal link** between attempts to alter public opinion and changes in citizens' behaviour.

5

Exploitation of existing social vulnerabilities. The agents responsible for committing disinformation actions against a state do not start from scratch. First, they detect real and spontaneous social and political vulnerabilities which are occurring in a state's public debate and then focus on **increasing and polarizing that debate.**

In this way, it is difficult to accuse these agents of provoking political or social crises, since their actual role is to distort, making pre-existing conflicts to rise or fall in intensity or introducing new factors as to modify their course.

6

Infiltration of illegitimate disinformation into the methods of legitimate social and political communication. The proliferation of illegitimate disinformation actions by actors interested in influencing citizen audiences in countries occurs within the framework of the legitimate use that political and social actors make of the newest technological platforms for mass dissemination of information to distribute their own messages and content.

In this scenario of conversations with thousands of actors in social networks and cross conversations on socially or politically controversial topics, the challenge to evaluate and react appropriately to disinformation campaigns, nullifying or counteracting them, is to **"separate the wheat from the chaff"**: to discern which opinions and information from those massively distributed on digital platforms are part of the legitimate influence attempt by social, economic or political actors; and which others use influence techniques and new social networks' features for malicious interference purposes.

For example, the *Computational Propaganda* project of the Internet Institute at Oxford University[7] in the United Kingdom studies "how *bots*, algorithms and other forms of automation are used by political actors in countries around the world", on the basis that automation technologies are already part of political conversations in democracies as well and, therefore, that disinformation for malicious interference purposes will be interspersed in that reality.

7. https://www.oii.ox.ac.uk/research/projects/computational-propaganda/

3. Objective of the guide

Currently, **the most effective way to develop effective resilience to disinformation actions is to protect the main target of these attacks: citizens.** It is necessary that the inhabitants of a state have the resources and develop the **necessary skills to identify products and communication platforms that are part of the disinformation tools.**

These disinformation strategies will be successful to the extent that their messages become hegemonic, assumed and shared by end users who, in most cases, do not know the true origin and motivation of the information sources they consume and share.

The purpose of this guide of good practices is precisely to explain the main characteristics and methodology of the current disinformation actions, with the aim that citizens and end users of digital media have the tools that allow them to consume and share information critically and avoid being involuntary accomplices of offensive actions against the interests of the state.

It is necessary for the inhabitants of a state to have the resources and develop the skills to identify communication products and platforms that are characteristic of disinformation tools.

4. What is the risk in Spain?

More than 20 million Spanish citizens, at risk of being victims of disinformation.

Countries such as Spain have already officially recognised the security threat posed by this new type of action. **The National Security Strategy**[8] developed by the Government of Spain in 2017, explicitly includes the threat of hybrid actions as one of the main security challenges facing the country[9].

Spain's National Security Strategy places the origin of these new threats in the complex socio-political and economic context and also recognises that this type of threat can come from both "state" and "non-state" agents, and that they combine the use of "military means with cyberattacks, elements of economic pressure or influence campaigns through social networks"[10].

There are currently 27.6 million Internet users in Spain, of which 25.5 million use social networks on a daily basis. The sources consulted indicate that 92 per cent of the Spanish population aged between 16 and 65 use the Internet daily to be informed and that 85 per cent does so through social networks, according to data from the National Observatory for Telecommunications and the Information Society (ONTSI) in 2017[11].

8. GOVERNMENT OF SPAIN. Department of Homeland Security. http://www.dsn.gob.es/sites/dsn/files/Estrategia_de_Seguridad_Nacional_ESN%20Final.pdf
9. GOVERNMENT OF SPAIN. Op. cit. , p. 16.
10. GOVERNMENT OF SPAIN. Op. cit. , p. 32.
11. NATIONAL OBSERVATORY OF TELECOMMUNICATIONS AND YES. 2018. Sociodemographic profile of Internet users. Análisis de datos INE 2017.https://www.ontsi.red.es/ontsi/sites/ontsi/files/Perfil%20sociodemogr%C3%A1fico%20de%20los%20internautas%2 02017.pdf

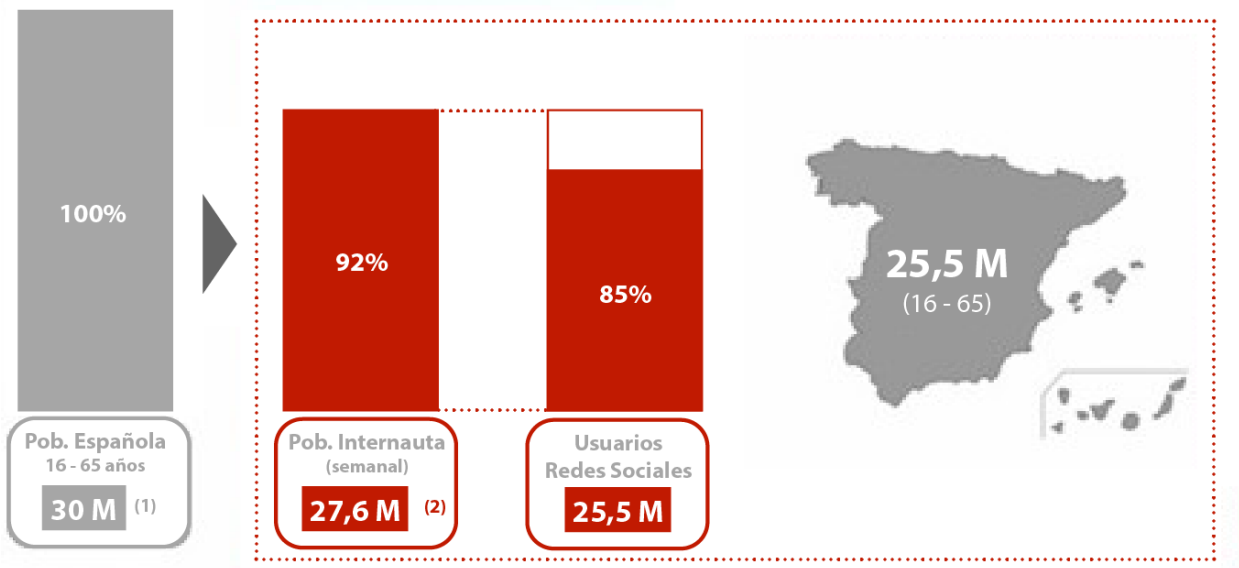


Illustration 1.- Chart elaborated by the Interactive Advertising Bureau (IAB Spain). 2018 Annual Social Networks Study.

According to data from the Association for Media Research (AIMC), reading current news is the main use Spanish citizens make of the Internet, as expressed by 84.6 per cent of people surveyed between October and December 2017[12].

These studies on Internet use and digital information consumption habits suggest that nearly **90 per cent of the Spanish population between 16 and 65 years of age may potentially be victims of a disinformation attack.** Other recent researches have shown that, despite the widespread use of the Internet and social networks among Spanish citizens, there is still a significant percentage of users who do not know how news distribution works on digital platforms.

12. ASSOCIATION FOR MEDIA RESEARCH (AIMC). 2018. 20th Edition Navigators on the Net. Survey AIMC to Internet users. (October-December 2017). https://www.aimc.es/otros-estudios-trabajos/navegantes-la-red/infografia- resumen-20o-navegantes-la-red/

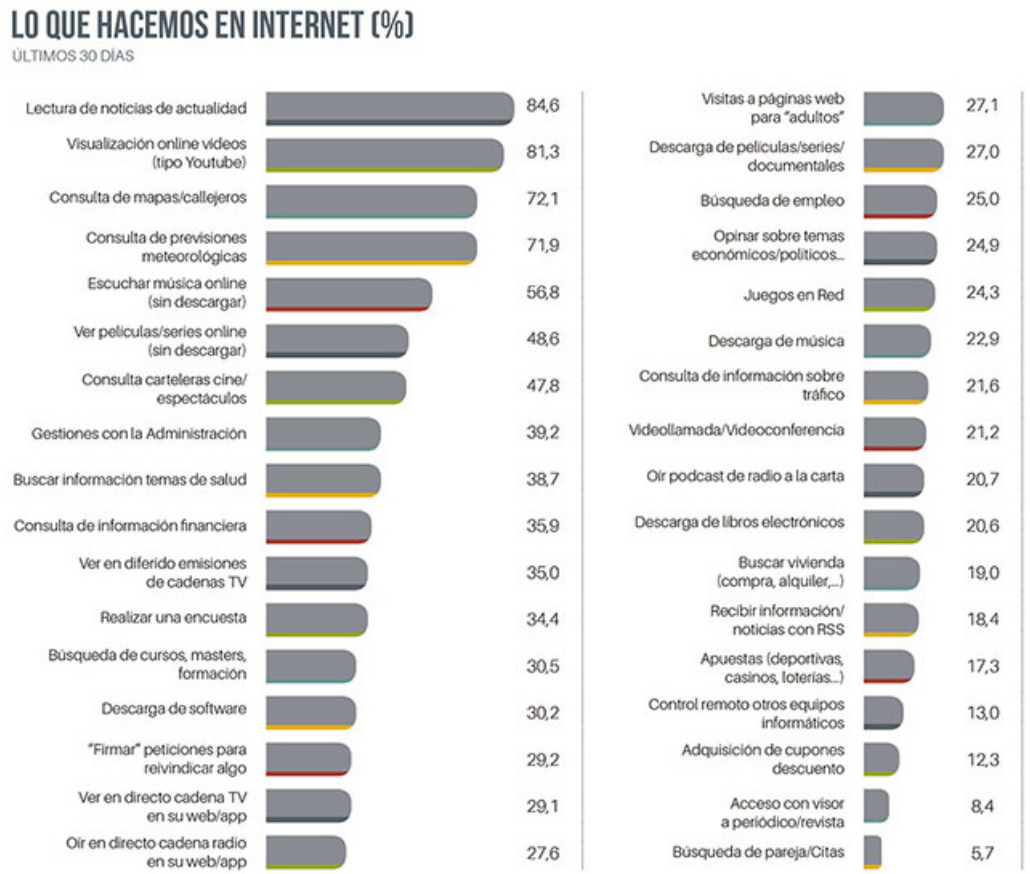


Illustration 2.- Chart elaborated by the Association for Media Research (AIMC13).

The "Digital News Report", done in 2018 by the University of Oxford and the University of Navarra, acknowledges for instance that only three out of ten digital users in Spain are aware that the news they read on Facebook depends on an algorithm[14]. **This lack of knowledge about the digital information environment constitutes a vulnerability of the Spanish public opinion.**

13. Available at https://www.aimc.es/otros-estudios-trabajos/navegantes-la-red/infografia-resumen-20o-navegantes-la-red/

14. UNIVERSIDAD DE NAVARRA/OXFORD UNIVERSITY. 2018. Digital News Report.ES 2018. Una audiencia diversa y preocupada por la desinformación. Coordinado por Avelino Amoedo, Alfonso Vara-Miguel y Samuel Negredo. Centre for Internet Studies and Digital Life School of Communication and Reuters Institute for the Study of Journalism. https://drive.google.com/file/d/1_MqxpbPvMQM1IpvjsGm4QOKxIMC8IZ_D/view

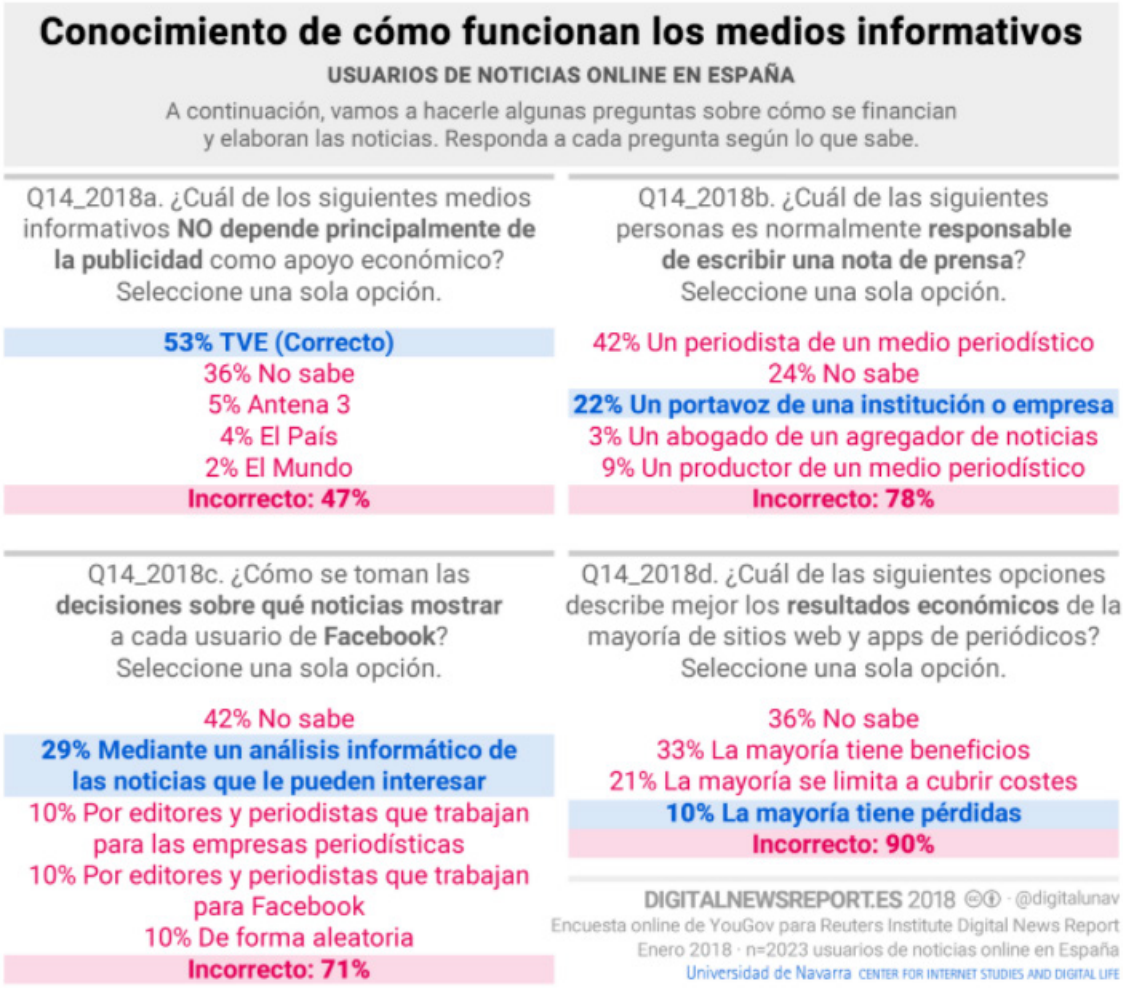


Illustration 3.- Digital News Report. ES 2018. University of Navarra. Oxford University[15].

15. Available at: https://drive.google.com/file/d/1_MqxpbPvMQM1lpvjsGm4QOKxlMC8IZ_D/view

5. What are the consequences of a disinformation attack?

The consequences of a systematic and malicious campaign of disinformation among the public opinion can lead to dangerous consequences for a liberal democracy.

The main objective of a disinformation campaign is to disseminate fake news when the public opinion in a country is developing, as well as half-truths, highly subjective information presented as objective **(deliberate confusion between opinion and information)** and information elaborated to have an emotional effect on the receiver, minimizing the likelihood of it being processed by applying critical judgment.

This information is disseminated through platforms and profiles which seem to be credible in appearance but hide its true origin and hinder its traceability. The malicious and systematic distribution of poor-quality information in public debate is intended to break the trust between the citizens of a country and two (2) of the main actors responsible for maintaining social cohesion: institutions and the media.

The modern nation-state is based on a **social contract based on the trust that citizens place in its administration and institutions.** The breakdown of this relationship of trust can compromise the solidity of the democratic fabric of a state. In this sense, **the consequences of a systematic and malicious campaign of disinformation among the public opinion can lead to dangerous consequences for a liberal democracy.**

5.1 Loss of confidence in traditional media

European citizens, and especially Spaniards, trust the media less and less, as reflected in a 2018 survey conducted by the Edelman Trust Barometer, which concluded that only 44 per cent of Spanish citizens still trust them[16].

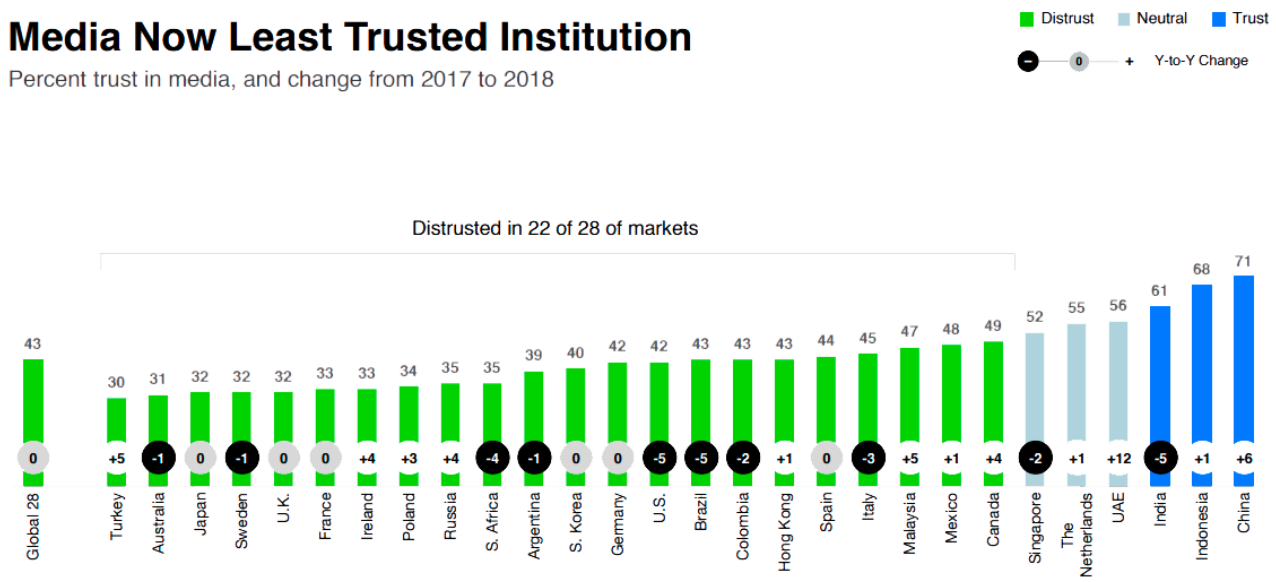


Illustration 4.- Chart prepared by Edelman Trust Barometer[17].

16. EDELMAN. 2018. "2018 Edelman Trust Barometer. Global Report". https://www.edelman.com/sites/g/files/aatuss191/files/201810/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf
17. Available at: https://www.edelman.com/sites/g/files/aatuss191/files/2018-10/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf

5.1 Loss of confidence in traditional media

Until the end of the 20th century, traditional media played the role of mediators in the process of shaping public opinion. They were identified and credible agents who created and transmitted messages that set the agenda and shaped public debate and, consequently, the social cohesion of a state.

This coincides with data from the aforementioned joint study by the University of Navarra and the University of Oxford, which also indicated that only 44 per cent of Internet users in Spain trust the news they read in the media[18].

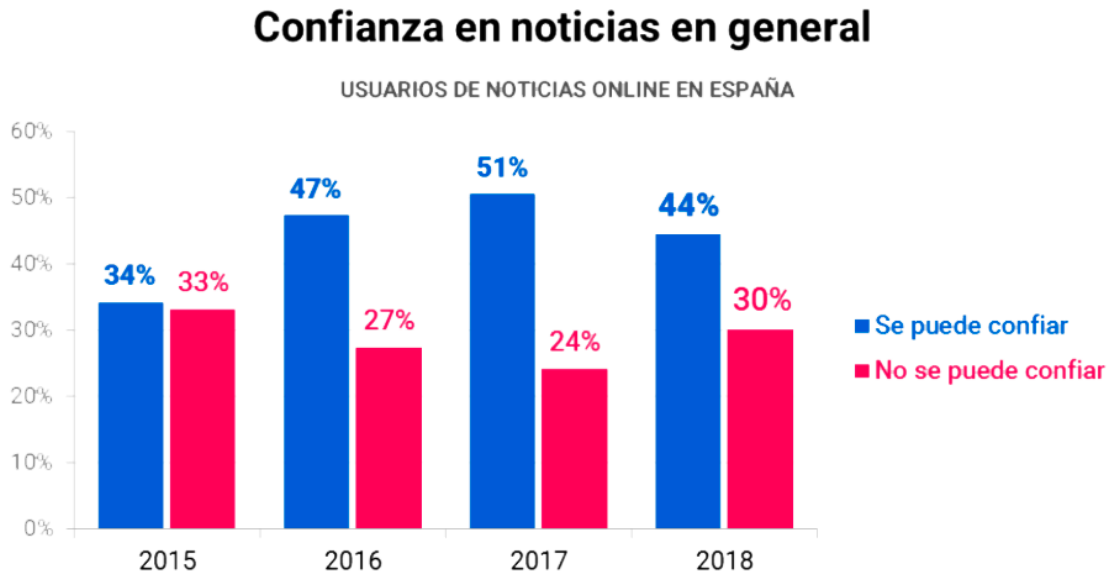


Illustration 5.- Digital News Report.ES 2018. University of Navarra. Oxford University[19].

The decline in public trust in the media is explained by historical causal mechanisms in which numerous structural factors linked to the evolution of societies, politics, technology, the generation and transmission of information as well as the current revolution in journalism comes into play.

18. UNIVERSIDAD DE NAVARRA/OXFORD UNIVERSITY. 2018. Digital News Report.ES 2018. A diverse audience concerned about misinformation. Coordinated by Avelino Amoedo, Alfonso Vara-Miguel and Samuel Negredo. Centre for Internet Studies and Digital Life School of Communication of Journalism. https://drive.google.com/file/d/1_MqxpbPvMQM1lpvjsGm4QOKxlMC8IZ_D/view
19. Available at: https://drive.google.com/file/d/1_MqxpbPvMQM1lpvjsGm4QOKxlMC8IZ_D/view

5.1 Loss of confidence in traditional media

This scenario of **changing public opinion regarding the credibility of the media** is used by offensive disinformation strategies to multiply and generate instability in public opinion.

According to the aforementioned report by the University of Navarra and the University of Oxford, Spain is one of the countries in the world where Internet users are most concerned about being victims of digital disinformation campaigns. Specifically, **69 per cent of Internet users acknowledge their concern about not being able to differentiate between what is true and what is false on the Internet**[20]. The Edelman report also coincides in highlighting Spain as one of the countries in the world with the highest number of Internet users concerned about fake news on the Internet[21].

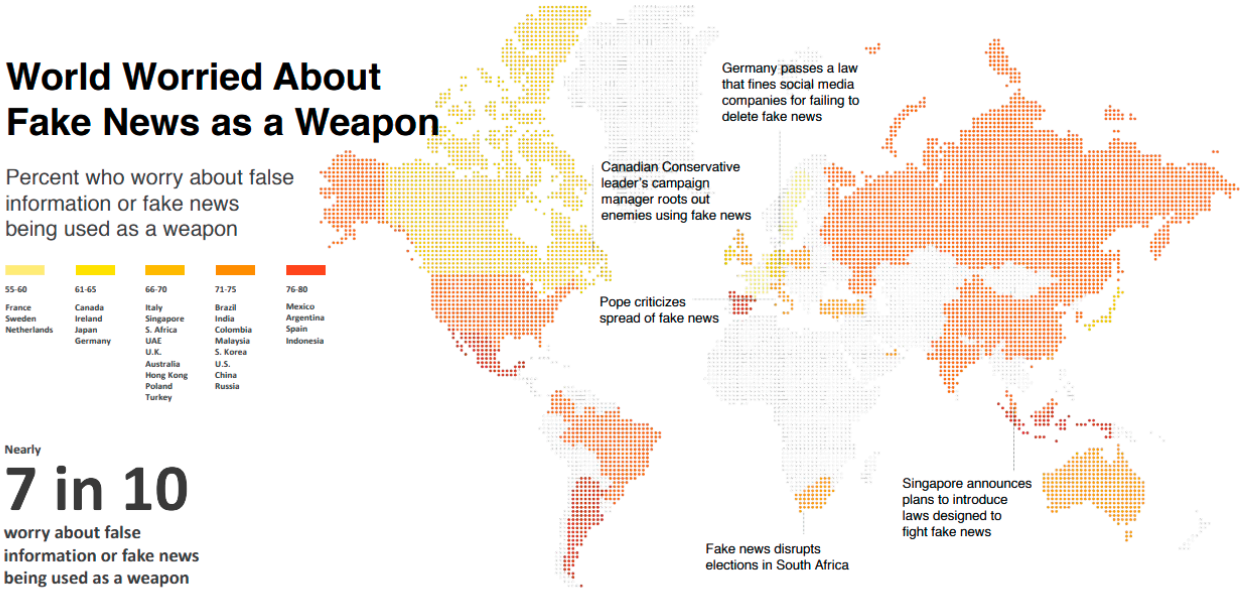


Illustration 6.- Chart prepared by Edelman Trust Barometer[22]

20. Ibid.
21. EDELMAN. 2018. "2018 Edelman Trust Barometer. Global Report". https://www.edelman.com/sites/g/files/aatuss191/files/2018-10/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf
22. Available at: https://www.edelman.com/sites/g/files/aatuss191/files/2018- 10/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf

5.1 Loss of confidence in traditional media

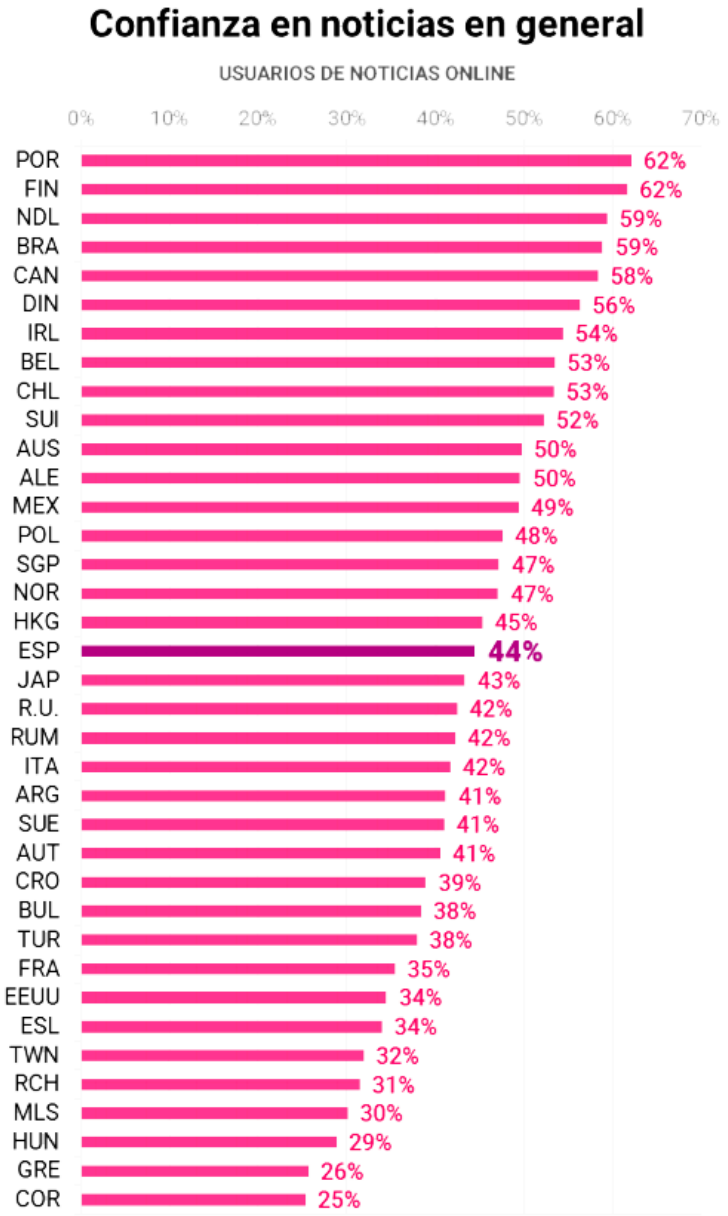


Illustration 7.- Digital News Report. ES2018. University of Navarra. Oxford University[23]

Finally, the European Union Eurobarometer published in December 2018 reported that 83 per cent of European citizens consider false news to be a real threat to democracy, while 73 per cent are concerned about digital disinformation campaigns in pre-election periods[24].

23. Available at: https://drive.google.com/file/d/1_MqxpbPvMQM1lpvjsGm4QOKxlMC8IZ_D/view
24. COMISIÓN EUROPEA. 2018. Action Plan Against Desinformation. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=55907

5.2 Loss of confidence in public institutions

There is evidence to suspect that disinformation campaigns are taking advantage of the social crisis of confidence in the media to be implemented and spread more easily. However, it is even more worrying to see how citizens' confidence in their countries' public institutions is also falling to historic lows.

In the case of Spain, only 24 per cent of citizens trust their government institutions, according to the Edelman report[25]. **One of the main objectives of disinformation campaigns is precisely to detect vulnerabilities in a country's social contract and strengthen them** with the aim of increasing mistrust between citizens and public institutions.

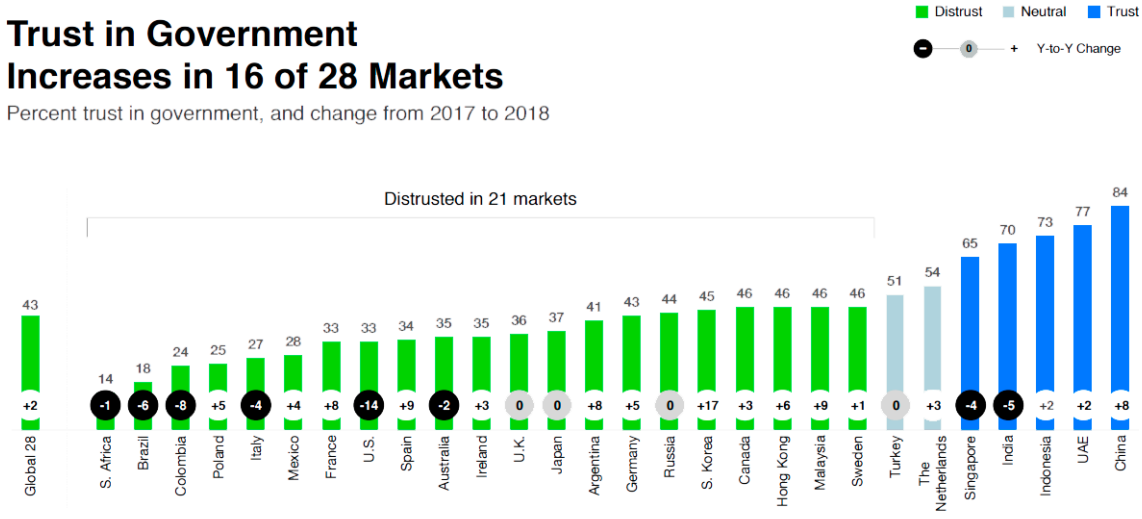


Illustration 8.- Chart elaborated by Edelman Trust Barometer[26]

25. EDELMAN. 2018. "2018 Edelman Trust Barometer. Global Report". https://www.edelman.com/sites/g/files/aatuss191/files/2018-10/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf

26. Available at: https://www.edelman.com/sites/g/files/aatuss191/files/2018-10/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf

5.3 Loss of confidence in the citizen's sovereignty

In relation to both the crisis of confidence in the traditional media and the erosion of confidence in democratic public institutions, some disinformation strategies may seek to destabilise the very basis of both: **the democracies being confident that a well-informed citizen will be able to make the best decisions for their governance.**

This disinformative line of action seeks to **break the confidence of the citizen in the essence of information as an element of democratic decision.** On the one hand, if it is possible to sow and germinate the feeling that traditional press media are not credible and that, on the other hand, only low quality information circulates through social networks (which is also used by foreign countries to influence votes in democratic elections), an attempt will be made to install in the citizen the feeling that there is no way to obtain information of quality in order to establish a value judgment so they can make their voting and governance decisions.

In the analysis of disinformation campaigns, particularly those attributed to foreign countries with geoeconomic or geopolitical interests, it is **good practice to consider the extent to which citizen confidence in the media** (press media, social networks, digital platforms) **is being undermined,** rather than the fact that the intention is to polarize the conversation of social or political issues in spurious directions.

For example, if a state with malicious intentions manages to grab headlines and academic analyses all over the world attributing to it the capacity to manipulate citizens and influence democratic processes in a good part of the electoral scenarios of countries in any continent, one of the effects generated is that this malicious state is attributed with the power and capacity to exert that influence.

From this point forward, the citizen may infer that they have less chance of legitimately informing himself to decide on their democracy, since there is a foreign state that manipulates social networks in a scenario where confidence in the press and institutions has also decreased.

Thus, the challenge to **counteract and annul disinformation campaigns** is at least fivefold:

5.3 Loss of confidence in the citizen's sovereignty

1. To detect and stop illegitimate influence attempts on specific issues of social and political polarization.
2. To adequately analyse these illegitimate attempts to distinguish them from processes of political and social influence by legitimate actors.
3. To adequately contextualize the attributions on intentions and capacities of actors with disinformation objectives, in order to correctly frame the small or large dimension of these actors in the global set of mass communication and transmedia that is already taking place in any democratic process.
4. To arbitrate information campaigns to the citizen about disinformation, its intentions and its scope, as well as about the attempts of some actors to make it seem to the public opinion that they can exert on it a "total manipulation", when they neither have the resources nor such ambitious intentions in most cases.
5. To introduce good practices in academic analysis or through the press, describing attempts at disinformative campaigns in social networks, in order to adequately dimension the scope that these campaigns may have and how these malicious campaigns share space with legitimate information which, in most cases, is the one that is reaching the citizen in the greatest quantity and quality.

5.4 Social polarization

The very nature of digital information platforms, which use algorithms to show the user those news items **considered to be of their taste, contributes to the creation of highly polarized digital conversations.** Disinformation campaigns seek precisely to increase this polarization.

First, by detecting those digital conversations that are the most controversial or cause the most confrontation in public debate and, in a malicious way, by encouraging and broadening these debates in order to confront a country's citizens around certain political or social issues.

A study by MIT's Social Machines Laboratory on the digital conversation (chats) during the U.S. electoral campaign in 2016 detected high rates of polarization around two political communities: one in favour of Donald Trump and the other in favour of Hillary Clinton[27]. The debate between these two communities in the digital environment was practically non-existent and the conversation only contributed to radicalize the positions of the members of each of these communities.

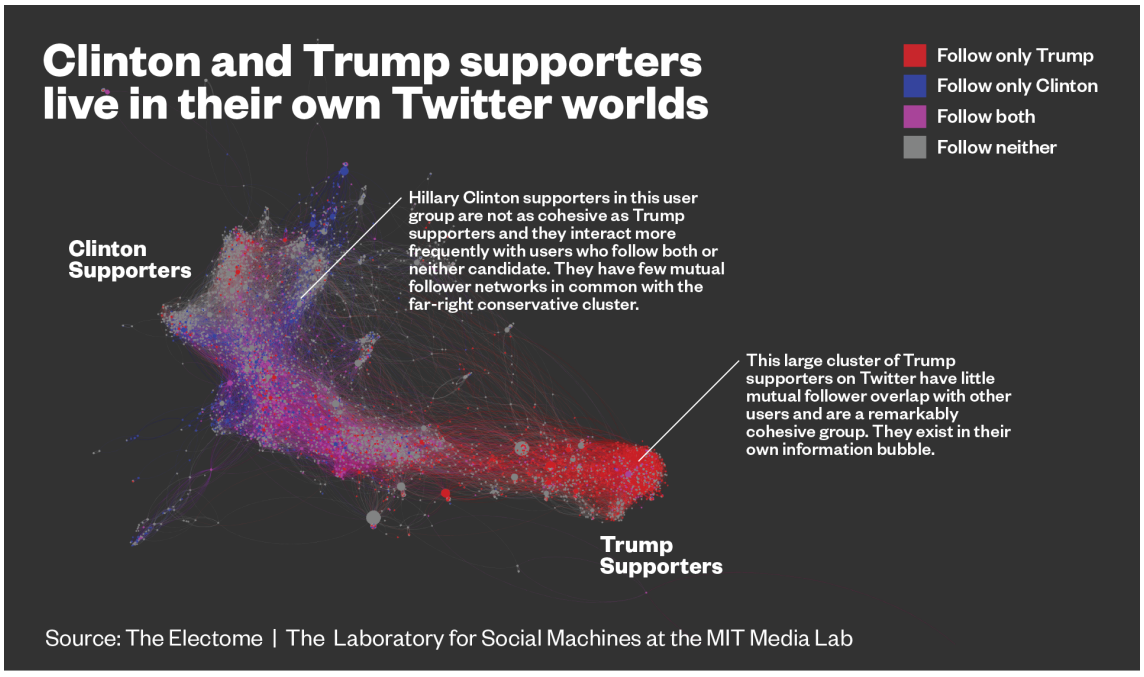


Illustration 9.- Graphic made by Vice News.
8th December 2016 with data from the Social Machines of the MIT Media Lab[27].

27. Available at: https://news.vice.com/en_us/article/d3xamx/journalists-and-trump-voters-live-in-separate-online-bubbles-mit-analysis-shows

5.4 Social polarization

Currently, judicial authorities and the U.S. Congress itself are analysing whether foreign countries used this polarization to increase social tension and maliciously erode the country's internal cohesion.

Similarly, a study carried out by the Spanish company Alto Analytics demonstrated that the digital debate on immigration in the Italian pre-campaign in 2017 also showed that Italian society was highly polarised around two communities and that they barely interrelated.

This same study provided indications that foreign media and other unidentified actors were able to contribute to increasing this division in Italian society through systematic disinformation campaigns[29].

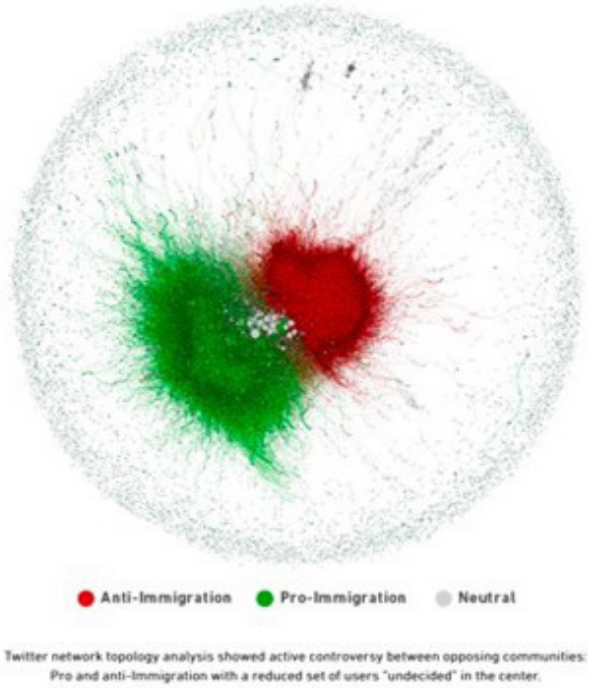


Illustration 10.- Chart elaborated by Alto Analytics[30].

28. VICE NEWS. 2016. Parallel narratives. Clilnton and Trump supporters really don´t listen to each other on Twitter. Alex Thompson. 8 diciembre 2016. https://news.vice.com/en_us/article/d3xamx/journalists-and-trump-voters-live-in-separate-online-bubbles-mit-analysis-shows
29. ALTO ANALYTICS. 2017. The Construction of Anti-immigration electoral messages in Italy. The role of foreign media in the anti-immigration debate one year before the 2018 election. https://www.alto-analytics.com/en_US/the-construction-of-anti-immigration-messages-in-italy/
30. Available at: https://www.alto-analytics.com/en_US/the-construction-of-anti-immigration-messages-in-italy/

6. The methodology of disinformation campaigns

Those responsible for the communicative action have strategies in social networks of automated accounts (bots) which disseminate messages with great magnitude and segmentation.

Actions of *hacking*, manipulation or disruption of public opinion through disinformation actions follow a constant pattern, both in offensive actions by nation-states and by sub-national groups. This methodology is based on a combination of different disciplines, such as sociology, politics, literature, journalism or technology.

The four (4) steps required to implement a successful digital and communications disruption campaign to destabilize a state's public opinion are as follows:

6. The methodology of disinformation campaigns

1 **Analysis and detection of a country's social and political vulnerabilities.**

Disinformation actions are based on exploiting the breeding ground for issues that already generate confrontations in public opinion in the internal dynamics of a country.

The alteration of public opinion does not consist in creating new confrontations in the internal politics of the country considered as an adversary. On the contrary, the strategy is focused on fuelling debates and political confrontations that are already in place and which polarize a society.

In the case of the United States, the alleged communicative disruption generated, supposedly by Russia, fed existing political and social debates. For its part, Daesh carried out an in-depth social and political analysis of the Sunni communities in Iraq and Syria with the aim of detecting and exploiting the main vulnerabilities of these states and proposing a narrative aimed at generating a new social contract with the Sunni citizens of these countries.

2 **Creation of "transmedia" narratives.**

Once detected in a country the spaces where scenarios of social or political confrontation take place, those responsible for a disruptive action of disinformation create effective and transmedia scripts and narratives, capable of generating a cultural resonance and a mobilization among their potential audiences.

These narratives are designed to be distributed through different characters, formats and communication platforms, adapted to the cultural and aesthetic particularities of the collectives to which they are addressed.

3 **Creation of their own media network.**

Since the first decade of the 21st century, the distribution of digital marketing content has been based on a three-pronged strategy called "owned media, paid media and earned media". There is a consensus in the communication sector to accept that the combination of these three (3) elements is the basis for the effective dissemination of the contents of a marketing campaign.

Owned media refer to the platforms and communication channels that a brand or product creates to directly communicate with its audiences; paid media refer to the advertisements and spaces that the brand inserts in foreign communication channels in exchange for an economic

consideration, and the earned media refer to the information, comments or unpaid communications that users and influential agents make about the brand.

Narratives constructed to channel and feedback a state's socio-economic and political conflicts are introduced into the digital debate through their own media network hierarchically controlled by the people responsible for a disruption mechanism for disinformation purposes.

4 **Creation of automated distribution channels.**

The final step in disruption is direct, automated and segmented distribution of narratives to potential audiences in digital environments.

To this end, **those responsible for the communicative action have strategies in social networks of automated accounts (bots) which disseminate messages with great magnitude and segmentation,** without waiting for the user to access voluntarily the platform where the content is published.

The existence of these automated distribution tools has been demonstrated in cases of interference by state actors[31] as well as non-state actors[32].

31. Javier. «Why did Russian social media swarn the digital conversation about Catalan Independence?» [en línea], en Washington Post. 22 noviembre 2017. Available at web: https://www.washingtonpost.com/news/monkey-cage/wp/2017/11/22/why-did-russian-social-media-swarm-the-digital-conversation-about-catalan-independence/?noredirect=on&utm_term=.126764742aa3.
32. BERGER, J.M.; MORGAN, Jonathon. «The ISIS Twitter Census. The Brookings Project on U.S. Relations with the Islamic World» [en línea], en Brookings, Analysis Paper. N.º 20. Marzo 2015. Available at web: https://www.brookings.edu/wp-content/uploads/2016/06/isis_twitter_census_berger_morgan.pdf.

7. The 10 key elements of a disinformation campaign

The European Commission identifies the phenomenon of disinformation with fake news. According to this institution, *"disinformation - or fake news - consists of verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public, and may cause public harm"*[33].

This definition clearly sets out the nature of offensive communication actions. However, it can also be demonstrated that *fake news* is just one more tool in the complex machinery behind a disinformation campaign.

Offensive communication actions are complex phenomena which use different tools and procedures (not just *fake news*) to deliver to citizens messages that cause chaos and confusion in the public opinion of a country considered an adversary.

Below, we will identify some of the most commonly tools used in disinformation campaigns. The first two (2) techniques are related to the creation of content or narratives and the next eight (8) are related to the process of content distribution (through media, social networks or algorithms).

Offensive communication actions are complex phenomena that use different tools and procedures to reach citizens with messages that cause chaos and confusion.

33. COMISIÓN EUROPEA. 2018. "Fake News and Online Disinformation". https://ec.europa.eu/digital-single-market/en/fake-news-disinformation

7.1 Fake news

Fake news can be accepted as credible or believable by a large number of citizens and can provoke serious political and security crises in a state.

Fake news is informative messages disseminated to the public and which do not correspond with any true or scientifically or historically demonstrable fact. Despite its null relation with the truth, fake news can be accepted as credible or plausible by many citizens and cause serious political and security crises within a state.

Believing this type of news to be true is due to the following characteristics:

- **They're based on some real elements.** These pieces of news are based on some real character, place or current phenomenon.
- **They're surprising.** These pieces of news are presented to the reader in a very attractive and sensationalist way, usually including provocative or surprising titles that make them want to read them.
- **They come from recently created or poorly traceable media.** Fake news is originally disseminated through unknown media, either because the medium is new (created ad-hoc to initiate such actions) or because it originates in foreign countries where it is difficult to investigate the origin and traceability of the medium.
- **Absence of sources.** Fake news is included in texts where no reliable or recognized sources are identified or mentioned. In the case that some source is presented, it is only to give voice to those that reaffirm the theory of the news and no opportunity of expression is given to voices or thoughts doubting the thesis maintained in the information.

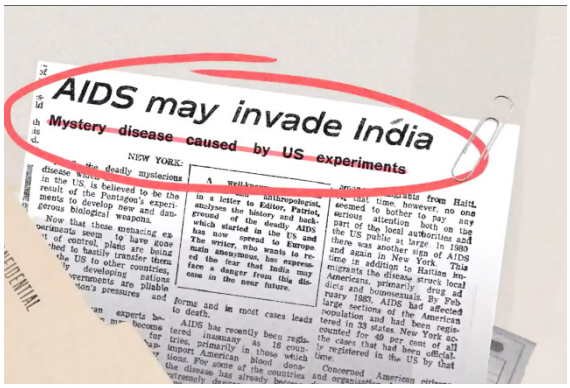
7.1 Fake news

They trust the long term. Fake news can have a short-term impact, but they can also develop their offensive capacity in the long term. Sometimes false pieces of information are published in an unknown medium in a foreign country, hoping that over time they enter the distribution chain of news in more reliable media and appear to be credible.

Likewise, news published in unknown media can end up being cited in academic research or in reference websites of massive use (such as Wikipedia) and, in this way, acquire after months or years the appearance of credibility and contaminate the process of shaping public opinion.

One of the most successful disinformation operations using fake news was the so-called "Operation Infektion"[34]. This campaign was launched by the Russian government in the early 1980s with the aim of spreading the word that the AIDS virus had been created by the United States government to eliminate the African-American and homosexual population living in the country. This story was first published in July 1983 in an unknown local newspaper in India and went almost unnoticed.

Nevertheless, as time went by, this information was quoted and mentioned by media in other countries, until, on 30th March 1987, the U.S. network CBS echoed the information in a primetime program and turned a false story, created four years earlier by Soviet intelligence services in India, into a topic of public and political debate in the United States.


AIDS may invade India

Mystery disease caused by US experiments

NEW YORK:

Illustration 11.- Original article published in India in July 1983 in which it was reported that AIDS had been created by the United States government.

34. NEW YORK TIMES. 2018. "Operation Infektion. Russian Disinformation: from Cold War to Kanye", by Adam B. Ellick and Adam Westbrook. https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html

A more recent success story, in late October 2016, is when **anonymous Facebook accounts and poorly credible media** began to spread the news that Democratic presidential candidate Hillary Clinton was part of a child sexual exploitation network based at a well-known pizzeria in Washington, D.C.

This information spread very quickly and thousands of citizens gave it credibility. In fact, barely a month after the publication of this news, a U.S. citizen entered the restaurant armed and began shooting with the intention of releasing the allegedly sexually exploited children[35].

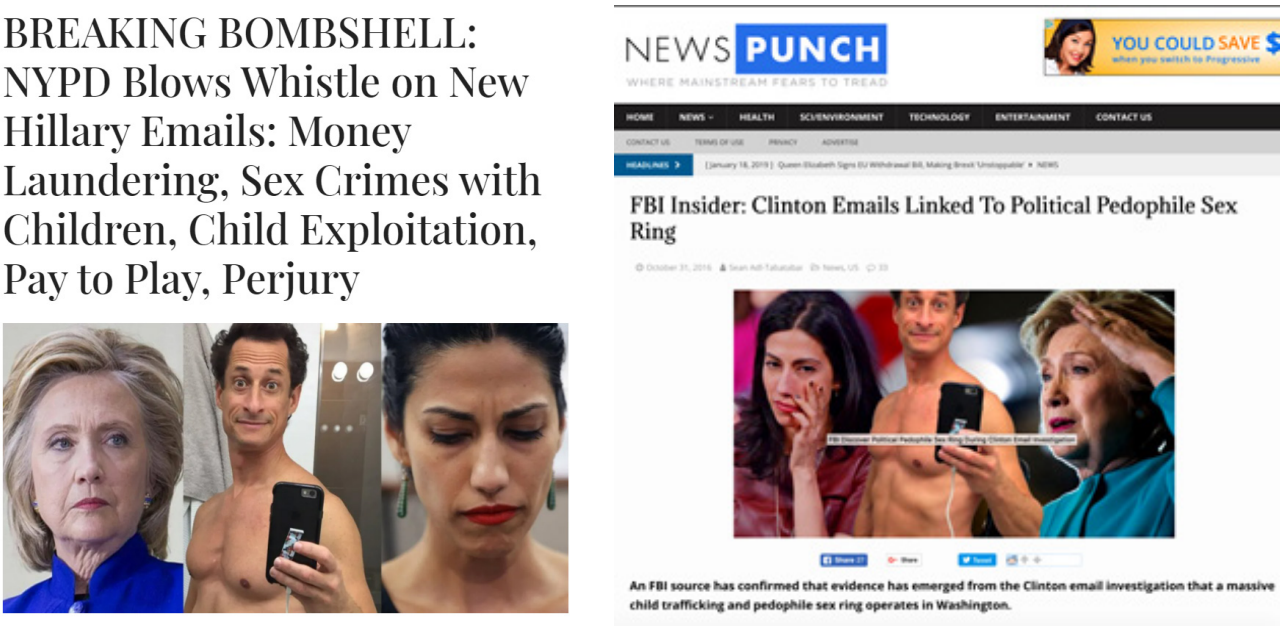


Illustration 12.- Screenshots of some of the diaries that disseminated in late October 2016 the fake news about Hillary Clinton's implication in an alleged child sexual exploitation network[36].

Another example of fake news, in this case with regards to Spanish news, occurred on 28th October 2017, when the Russian medium RT News published a headline announcing the presence of tanks in the streets of Barcelona (the information was shared on Facebook by 11,800 users).

35. ROLLING STONE. 2017. "Anatomy of a Fake News Scandal", by Amanda Robb. https://www.rollingstone.com/politics/politics-news/anatomy-of-a-fake-news-scandal-125877/

36. You can read these news items by clicking on the following links: https://newspunch.com/fbi-clinton-email-pedophile-ring/ https://truepundit.com/breaking-bombshell-nypd-blows-whistle-on-new-hillary-emails-money-laundering-sex-crimes-with-children-child-exploitation-pay-to-play-perjury/

"Tanques en las calles de Barcelona": España y Cataluña al borde de un desenlace violento

Publicado: 28 oct 2017 02:56 GMT

El analista político John Wight considera que el movimiento separatista está siguiendo un camino peligroso, en medio de la determinación de Madrid de "desatar la violencia".

Illustration 13.- Screenshot of the information published by RT News suggesting the presence of tanks in the streets of Barcelona in October 2017[37].

In The Line Of Fire @Vuvuzela_Infos · 4 feb.
Dictator #Macron & #EU mercenaries were caught speaking different language than french during Saturday manifestation in #Paris. #GiletsJaunes #YellowVests are seriously THANKING #Macron for employing THUGS to kill innocents rather than solving social demands.
Traducir Tweet

Politique4.0#10 retwitteó
Positively Joan @PositivelyJoan · 4 min
RT@Vuvuzela_Infos
[Revealed] Militia Army sent by EU to help dictator #Macron during Sat 2nd feb manifestation in #France against #GiletsJaunes #YellowVests.
Traducir Tweet

Illustration 14.- Screenshot of two disinformative digital profiles on Twitter claiming that foreign militiamen are "murdering innocents" on the streets of France during the French yellow vest crisis.

37. Available at: https://actualidad.rt.com/actualidad/253812-espana-cataluna-violencia-conflicto

7.1.1 Deep fake news

Within the phenomenon of fake news, it is important to highlight the evolution and popularization of the so-called deep fake news. This is a technology that can add even more complexity and offensive capacity to this technique and consists in the use of software to create real video images of authorities and politicians but with the audio and movement of the mouth modified.

Accordingly, speeches or statements of public authorities that look real, but whose message is totally or partially adulterated, can be disseminated.

7.2 The approach

Approaches manipulate the perceptions that can be generated by headlines and photographs of information.

Disinformation campaigns are not only based on false content. On other occasions, information disseminated with malicious intentions is based on a real fact but is presented to public opinion with an approach elaborated and constructed in such a way that the end user interprets that fact in a way that does not coincide with reality.

These approaches manipulate the perceptions that can be generated by headlines and photographs which are part of the information, known as the "first reading level" of a story. This first reading level not only attracts the first attention of the user, but also, with the current habits of informative consumption in mobile devices, constitutes, in many occasions, the only information that is received from a news item.

FUERTES VIDEOS: La brutal represión de la Policía contra los votantes del referéndum catalán

Publicado: 1 oct 2017 12:18 GMT

Los oficiales han irrumpido en numerosos centros de votación y han disparado balas de goma y utilizado otros medios violentos contra las multitudes de votantes.

Illustration 15.- News item published by RT News on 1 October 2017[38].

The headline is accompanied by a photograph in which a person lies flat on the floor with closed eyes and a bleeding wound on the forehead which has been *pixelated*. The information included in the text does not lie explicitly, hence the news item could not be presented as a case of fake news. The photograph is real and shows a person injured in a confrontation with the police during the referendum on illegal independence in Catalonia.

Still, the first level of reading, composed of the headline ("strong videos, brutal repression") and the photograph (a person lying down with eyes closed and a bleeding wound on the forehead) may lead the reader to the conclusion that there were people killed on 1st October 2017 in Catalonia (the information has been shared on Facebook by 22,800 people up till January 2018).

38. Available at: https://actualidad.rt.com/actualidad/251682-represion-brutal-policia-espa%C3%B1ola-referendum-catalan

7.3 New media

New media serve as platforms for the initial publication of malicious information.

The phenomenon of fake news has become popular and has increased exponentially at the beginning of the 21st century due to the ease and effectiveness with which actors with their own interests, covert political agents or foreign countries can create media and communication platforms which look professional and credible, and in various languages.

These new media serve as platforms for the initial publication of malicious information and, in a matter of days or months, can compete in influence with media that have been established for decades in the public opinion of a society.

They usually share some of the following features:

- Recent creation and absence of recognized professional trajectory.
- Absence of signature on the news.
- Lack of information about the publishers of the contents and about the financial managers or shareholders.
- Lack of information about the physical headquarters of the company.
- Linked to foreign governments which hide their governmental relationship and even try to mimic local or national media to generate confusion.

Quality media are implementing measures to differentiate themselves from new digital platforms aimed at disseminating disinformation campaigns. One of these projects is the *Trust Project,* to which several media from all over the world have adhered, as well as several academic institutions.

7.3 New media

This project establishes eight (8) confidence indicators which guarantee that the process of production and dissemination of their information abide by criteria of quality and confidence[39]:

- **Best practices:** What are the principles of the medium? Who finances it? What is its mission? It also implies the inclusion of a code of ethics, a commitment to diversity, rigour, corrections and other standards.
- **Journalist's experience:** Who wrote this article? Information about the author, including his background and published articles.
- **Type of work:** What journalistic genre does the article belong to? Labels that distinguish the texts of opinion, analysis or advertising from the news.
- **Quotes and references:** What is the source? For research or in-depth stories, access to the sources behind the facts and statements.
- **Methods of work:** How was it built? Also, for in-depth stories, information about why reporters decided to follow a story and how they approached the process.
- **Is it local?:** Identifying when the story arose in a place where the medium has in-depth knowledge of the local context or the community to which it is directed.
- **Diversity:** What is the editorial commitment to bring diverse perspectives?
- **Readers' Comments:** Providing spaces to encourage readers' participation and feedback.

39. https://elpais.com/estaticos/que-es-the-trust-project/

7.4 Social forums

Disinformation campaigners spread malicious messages in the hope that other users will give credence to the information and spread the content.

Online discussion forums (whether open, on public platforms, or closed, on deep web pages) are also recurrently used for the dissemination of disinformation campaign messages.

The success of these forums lies precisely in the fact that users can comment on any current issue in a completely anonymous manner and there is no control or censorship over the veracity of the shared messages.

People behind a disinformation campaign spread malicious messages in the hope that other users will give credibility to the information and disseminate the content on other open digital platforms as well as on personal social networks.

7.5 Malicious digital profiles

The dissemination of malicious or fake news does not only take place through new media of low credibility.

The dissemination of malicious or fake news does not only take place through new media of low credibility. They also do this through the manipulation or falsification of digital profiles on social networks of real people or institutions with the aim of making the public believe that they have made statements which they actually never made.

The manipulation of these digital profiles can adopt different aspects or modalities:

7.5 Malicious digital profiles

Malicious simulation of real accounts and profiles. The most effective method of supplanting the digital profile of an individual or institution is through the use of software or websites that allow you to recreate the look and feel of a real social network but adding invented content.

Once the recreation is done, a photograph or *screenshot* of the publication is taken and shared through social networks such as WhatsApp among close contacts, until the false message is *viralized*.

Creation of false digital profiles or parodies. Another way to spread malicious information through social networks is by creating a digital profile of a person or institution without their consent and supplanting their identity.

To avoid this type of situation, some digital platforms offer the possibility of verifying and authenticating that an individual's digital profile matches his or her true identity.

Hacking digital profiles. A more sophisticated way is by hacking or stealing the user's passwords from a digital profile to control it maliciously for a period of time. Some of these actions have caused losses and damage not only to the reputation of companies, but also to the global economy.

One of the most damaging actions carried out using this methodology occurred on 23rd 2013, when the group calling itself "Sirius Electronic Army" took control of the Associated Press (AP) Twitter account and reported a shooting at the White House. In a matter of a few minutes this message caused the U.S. stock market to plummet more than 150 points for 5 minutes[40].

A tweet from the Associated Press was false, and the AP said it had been hacked.
SCREENSHOT

Illustration 16.- Twitter account of the Associated Press agency.

40. USA TODAY. 2013. "AP Twitter feed hacked; no attack at White House". https://www.usatoday.com/story/theoval/2013/04/23/obama-carney-associated-press-hack-white-house/2106757/

7.6 Automated accounts of non-human behaviour

The creation of anonymous accounts, managed in an automated way in social networks, is another mechanism used to spread mass messages in a disinformation campaign.

The creation of anonymous accounts, managed in an automated way in social networks, is another mechanism used to **spread mass messages in a disinformation campaign.**

Most digital social communication platforms allow the creation of anonymous accounts and the use of software that automates the management of these profiles. This allows covert groups to launch disinformation campaigns which reach a large volume and dissemination on social networks.

The detection of automated digital profiles in a disinformation campaign depends on the level of sophistication used in the creation of such profiles. In any case, it is possible to define some characteristics to determine if there is a real person behind a digital account or we are in front of an automated tool to massively spread a certain message.

No correspondence with a real, verifiable person. The first feature of an automated digital profile is that neither the name, nor the photograph, nor the account information relates to a real person identifiable in other sources.

Likewise, in the content history of the profile there is no detail or comment regarding the personal life of an individual or a real institution.

High or unusual daily activity. Automated accounts stand out for their unusual temporary behaviour. Hundreds of messages on the same subject during a single day; an unusually high number of messages (more than fifty per day); messages posted 24 hours a day, seven days a week...

Absence of followers or followers suspected of being automatic accounts. The digital profiles used in disinformation campaigns stand out for the scarce number of followers they have in their profile or, alternatively, for having a high number of followers that could also be automatic bot accounts.

7.6 Automated accounts of non-human behaviour

- **Unilateralism.** Automated digital profiles don't usually dialogue on social networks, they just spread messages in conversations where their potential audiences are.
- **Absence of original content.** One of the main characteristics of digital non-human behaviour accounts is the absence of original content. Most of these profiles share or interact with content created by other profiles.
- **Little thematic variety.** Automated digital profiles focus on publishing and disseminating as widely as possible those political or social messages for which they have been created.
- **Scarce variety of sources.** Similarly, these disinformation campaign accounts exclusively use those sources that are part of the same strategy and create and disseminate similar messages with the same approaches.

Similarly, the **automated use of social networks** is, for instance, the technique commonly used by the terrorist group DAESH to disseminate its campaigns on social networks. Terrorists generate an average of 100 new digital profiles, with hardly any followers, to disseminate each of their communication campaigns, and manage these accounts through bots that automate the distribution of messages.

The *viralization* of contents is done by parasitizing the most popular hashtags among the conversations of their potential audiences.

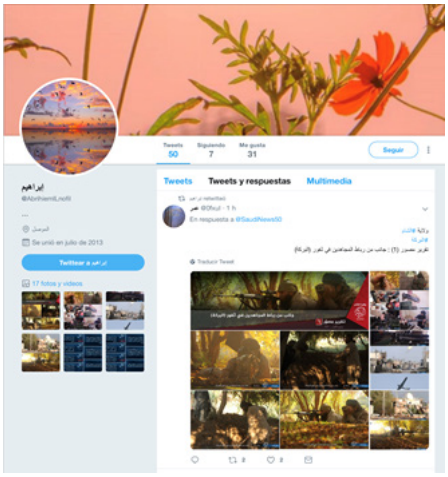


Illustration 17.- Example of bot used by the terrorist group DAESH to disseminate its propaganda through social networks.

7.7 Digital coverages or hybrid accounts

Disinformation campaigns are increasingly using more complex strategies to make it more difficult to detect automated accounts responsible for spreading their messages.

Disinformation campaigns are increasingly using more complex strategies to make it more difficult to detect automated accounts responsible for spreading their messages. Those responsible for offensive disinformation strategies are aware that automated accounts with the features described above are easily detectable and lack the credibility necessary to make a message *viral* in an efficient manner.

For this reason, **disinformation campaigns are increasingly employing digital profiles that appear to be human behaviour,** but are actually part of an army of false identities created and controlled by an organised group working undercover.

These digital profiles try to hide the characteristics of automated accounts; they seem to be associated with real people with first names, surnames, photographs, hobbies, dates of birth; they have a considerable number of followers; they have a considerable temporal trajectory and sometimes they interact with other users. Nonetheless, this is all part of the carefully crafted coverage to gain the trust of other users and effectively distribute misinformative messages.

7.7 Digital coverages or hybrid accounts

An example of this type of account is the profile from which the fake news linking Hillary Clinton to a child exploitation network in a pizzeria was spread. The information that came out of anonymous digital forums gained credibility the moment they were posted on a social network account that appeared to represent a real person: David Goldberg (@davidgoldbergNY), a New York City lawyer, who even presented himself with a photograph.

However, no real person was behind that account, which was only used to spread false information about the U.S. presidential candidate. The profile was deleted by the social network Twitter when it was proven false.

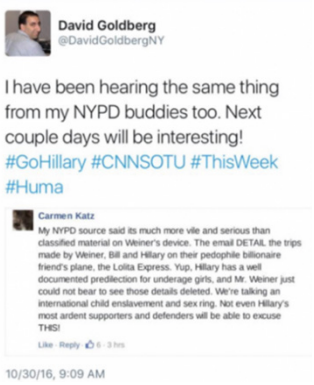


Illustration 18.- Message published from @davidgoldbergNY account.

7.7 Digital coverages or hybrid accounts

Another example of false social network accounts, which appeared to represent real people, occurred during the dissemination of news related to the holding of the illegal referendum of independence in Catalonia in October 2017.

Among the most active digital profiles distributing news criticizing the attitude of the Spanish government was the profile of Ivan (@ ivan226622), an Asian citizen (with real photographs) who claimed to be passionate about technology, business and news. He had 1,287 followers and had posted 580,000 messages since November 2012.

However, Ivan was not a real person, but part of a network of fake digital profiles that, in a coordinated way, was dedicated to spreading negative news about Spain during the first weeks of October 2017. In fact, it was possible to see how @ivan226622 tweeted the same negative news items about Spain at the same time as other accounts such as @rick888 or @bobbit2266, what suggests that all these accounts were managed by the same person or organisation. Within weeks of being detected, the social network Twitter removed these accounts when they were proven false.

Illustration 19.- Twitter messages with the publication of the same piece of news from different accounts.

In addition, digital coverages or hybrid accounts are also associated with "coordinated inauthentic behaviour"[41], a term used by Facebook to refer to pages and profiles that work together to deceive users about who they are or what they do. For ideological or financial reasons, these accounts are believed to be used from a part of the world that differs from their real origin.

In this sense, the social network Facebook has also detected the creation of false profiles on its platform, accounts that has been removed once identified their origin. On 31st January 2019[42], on its official website, the social network reported the closure of 207 Facebook pages, 800 profiles and 456 groups, *"linked to the online union Saracen Group of Indonesia"*. According to Facebook, these accounts were removed because of their deceptive behaviour and not because of the content they posted.

7.8 The stars invited

The collaboration of influential people, who have certain credibility among potential audiences.

Another element used to spread messages in a disinformation campaign is through the **collaboration of influential people, who have certain credibility among potential audiences** and who, above all, do not appear to have any kind of direct link or interest in the political or social issues under discussion.

This apparent distance endows the messages emitted by these influential characters with greater credibility and objectivity. Nevertheless, there is a possibility that these people have political or economic ties to undercover agents and that their messages are part of a disinformation strategy which the final audiences are unaware of.

41. GELICHER, Natahniel, 'Coordinated Inauthentic Behaviour´, Facebook, https://newsroom.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/
42. GELICHER, Nathaniel, `Taking down coordinated inauthentic behavior in Indonesia', Facebook, https://newsroom.fb.com/news/2019/01/taking-down-coordinated-inauthentic-behavior-in-indonesia/

7.9 Algorithms, resonance chambers and trusted networks

The debate disappears and digital communities in dispute just radicalize and reaffirm their positions instead of dialoguing in a reasoned debate.

The algorithms used by new digital communication platforms, such as social networks, have involuntarily become allies of disinformation campaigns. These technologies do not make it a priority for the end user to receive varied and plural information on current political, social or economic affairs. On the contrary, the algorithms used by these technological companies are designed so that the user receives and interacts only with messages that potentially can be of their liking and reaffirm their ideas.

In this way, communication platforms offer information to their users according to their political preferences, and not the quality, plurality or veracity of the contents.

This circumstance favours the creation of echo chambers in the digital conversation, where **the debate disappears and digital communities in dispute just radicalize and reaffirm their positions instead of dialoguing in a reasoned debate.**

7.10 Paid advertisements

Disinformation campaigners have found it effective to create content on the most contentious issues in a country's digital debate and promote it through paid campaigns.

The creation of these echo chambers by the algorithms has favoured the massive dissemination of disinformation actions by covert agents. Those responsible for disinformation campaigns have proven the effectiveness of creating content on those most controversial issues in a country's digital debate and **promote them through paid campaigns between each of the echo chambers created in these debates.** In this way, each of the conflicting communities identifies with this content, interacts with it and social polarization and confrontation between public opinion is fostered.

This was the strategy that, according to a report by the Intelligence Committee of the U.S. House of Representatives, Russia used to interfere in the 2016 presidential elections. According to this research, the Russian-based Internet Research Agency (IRA) paid Facebook to promote more than 3,000 unsigned and unidentified messages[43] aimed at polarizing U.S. citizens around the country's most controversial debates.

Thus, the IRA was able to covertly finance campaigns for the integration of the Muslim community and, at the same time, campaigns promoting Islamophobia, as shown in the images bellow, included in the report of the U.S. House of Representatives.

43. THE WALL STREET JOURNAL. 2018. Release of Thousands of Russia-Linked Facebook Ads Shows How Propaganda Sharpened, by Deepa Seetharaman, Georgia Wells and Byron Tau. https://www.wsj.com/articles/full-stock-of-russia-linked-facebook-ads-shows-how-propaganda-sharpened-1525960804

7.10 Paid advertisements

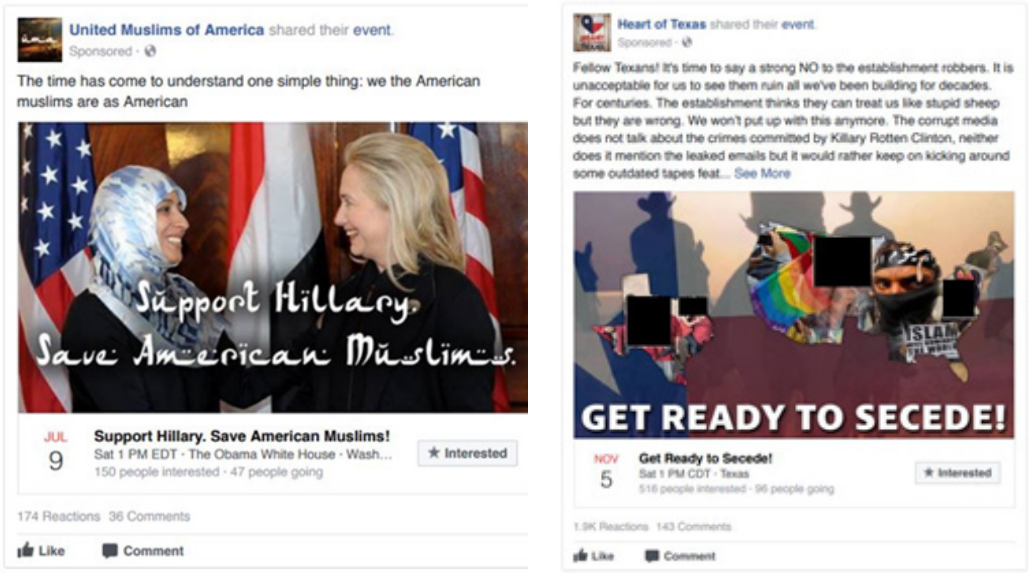


Illustration 20.- Images included in the report of the U.S. House of Representatives.

8. Decalogue of recommendations

Avoiding being the victim of a disinformation campaign is not the responsibility of a single agent. **Public institutions have the obligation to develop the necessary capacities to prevent, detect and neutralize disinformation offensives** against a state.

Private companies also have an obligation to prevent their digital platforms from becoming tools used in malicious campaigns against citizens and legitimate government systems.

On the other hand, **the Academy must continue to investigate this new phenomenon and generate scientific evidence on the methodologies and consequences** that disinformation campaigns have on public opinion and governance.

However, the first and last victims of communication wars are the citizens. For this reason, **digital media users** need to **be prepared to detect a disinformation campaign and acquire the necessary skills to avoid being manipulated.**

Public institutions have the obligation to develop the necessary capacities to prevent, detect and neutralize disinformation offensives.

Security Decalogue against disinformation campaigns

1. **Analyse the source of the news you receive and consume:** dozens of communicative impacts are received daily in our mobile devices with news that surprise us, outrage us or excite us. Sometimes, this news comes from "non-traditional" digital platforms with little or no transparency. It is important to know which media publishes a news item, what its trajectory is, and what journalists, companies or countries are behind the publication[44]. In this sense, it is advisable to have links that redirect the information to its original sources or to other texts that validate the data.

2. **Doubt of the screenshots you receive through social networks:** when you receive news in image format, it is advisable to always show a dose of caution and scepticism. There are many computer programs and software, very easy to use, that can retouch or modify images with false headlines of traditional media or accounts and profiles in social networks of real people. In the same way, it is very common to take images out of context, dissociating the real shot with the headline, and giving credibility to a false story. If you doubt the reality of some of these messages, it is advisable to check the original source with its Internet links or do what is known as reverse search in order to know if a photo was already published on the Internet. It is also possible to check whether an image is original or has been copied through EXIF information.

3. **Who shared the news with you and in what context?** Don't give credibility to all the messages you read on social networks, especially to messages or comments published by anonymous accounts and profiles. Ask yourself, even if it was sent to you by a friend, what the date of the information is, who is the source and what other media have spread it. It is even advisable to search for the headline in some search engine, because, if it is true, other media will have picked it up. Give credibility only to news shared by real sources[46].

4. **Watch out for false "human" accounts:** accounts with an apparent human aspect are appearing on social networks more and more frequently. However, they are actually managed by robots or by third parties in charge of controlling different profiles. Before continuing or relying on the content published by a digital profile, analyse how many people they follow, how many people follow them, whether they publish their own content, if they use excessively the social network... All these are indicators to detect false digital profiles in social networks[47].

5. **Don't be part of the algorithm:** the digital platforms we use daily to communicate and get informed are based on a complex algorithm which offers us personalized information according to our supposed tastes, hobbies or opinions. In this way, digital platforms offer us the information that the software considers to be to our liking. If we want to develop a well-formed, critical and contrasted opinion, it is advisable to obtain alternative sources of information to those that, by default, the algorithms of the communication platforms show us.

6. **Read the fine print:** the most successful disinformation actions are those based on half-truths. Fake news is relatively easy to detect and dismantle. Nevertheless, on many occasions, disinformation promoters use true photos and data which, presented suggestively in a headline and accompanied by a photograph, can be misinterpreted. When reporting in digital formats, don't resign to just have the feeling that a headline and photo can generate. Read the full story and analyse if the data are contrasted and if the quotes and opinions include plurality of opinions.

7. **Be on the lookout for sponsored content of unknown origin:** digital platforms earn revenue in exchange for users sponsoring certain content so that it appears prominently in the profile of a given audience. Be wary of any political or controversial content that appears to be sponsored by anonymous profiles or not identified with real associations, political parties or institutions.

8. **Do not trust guest stars:** there are often cases where relevant political, social or cultural actors are actively involved in political or social discussions in foreign countries. Freedom of expression and opinion is one of the most respectable assets of a liberal democracy. However, it should also be taken into consideration that some of these influential actors participate in certain discussions on the basis of very specific political and economic agendas not known to the final audience.

9. **Critical thinking and a cool head:** certain political bodies, both national and sub-national, use digital communication to confront public opinion in a foreign country and mobilize legitimate discontent on the part of its citizens around controversial and debating political, social or economic issues. Participating in political debates enriches democracy and political plurality. Nonetheless, it is advisable to do it from rationality, respect and critical thinking, avoiding generating spirals of hatred and disqualifications, which sometimes may be promoted by agents or covert groups.

10. **You can stop a conflict:** contemporary disinformation actions are based on the speed and virality with which news, rumours and comments spread. We are all part and parcel of disinformation campaigns. It is important to be aware that we can be used as pawns of strategies sponsored and managed by unknown agents with undeclared political interests. For this reason, it is important to always be aware of the information we receive daily on our computers or mobile devices and not to contribute to disseminating information that is not verified or has a doubtful traceability and origin.

CCN centro criptológico nacional

CCN-CERT centro criptológico nacional

44. There are various agencies, media and web portals to confirm or deny data. This is the case of the Spanish website Maldito bulo, member of the High Level Expert Group of the European Commission, BSDetector, Snopes or FactCheck.
45. There are several tools that allow it, such as RevEye, which work as an extension for Chrome and Firefox browsers and searches various image banks on the web. Google Images also has a similar function, as well as TinEye. It is also possible to have tools that verify videos such as Google Earth (to compare geography or places of interest where it was recorded) or YouTube DataViewer to know the time when the video was uploaded and that extracts screenshots to know if there are other versions of the video available on the Internet, and if they were published before or after.
46. The social networks themselves, for the most part, have information and good practices to avoid fake news and fall into chains of diffusion. This is the case of Facebook, Google or Twitter
47. Crowdtangle is a tool that Facebook acquired in 2016 and monitors how content is 'moving' on social networks or Foller.me that determines if a profile is misleading or not. Indications are that the profile "converts" with few accounts, if it has very high peaks of activity or its date of creation.

centro criptológico nacional

centro criptológico nacional

www.ccn.cni.es

www.ccn-cert.cni.es

oc.ccn.cni.es