Revista Latina de Comunicación Social
ISSN 1138-5820 / No. 82, 01-18.
https://doi.org/10.4185/rlcs-2024-2205

Application of AI tools as methodology for the analysis of toxicity in social media: A case study of Spanish politics on Twitter

Aplicación de herramientas de IA como metodología para el análisis de la toxicidad en redes sociales: Estudio de caso de la política española en Twitter

Uxía Carral
Carlos III University of Madrid. Spain.
ucarral@hum.uc3m.es

Carlos Elías
Carlos III University of Madrid. Spain.
carlos.elias@uc3m.es

Funding: The present text appears within the framework of the two projects granted to members of Madrid's Carlos III University from which this research receives funding: (1) 'Pseudoscience, conspiracy theories, fake news and media literacy in health communication'(Ref: PID2022-142755OB-I00) within the National Plan 'Knowledge Generation Projects' of the Ministry of Science and Innovation and (2) 'EU, disinformation and fake news' (Ref: 610538-EPP-1-2019-1-ES-EPPJMO-CHAIR), a Jean Monnet Chair of the European Commission.

Acknowledgements: The authors would like to sincerely thank Rubén Míguez Pérez, PhD in Telecommunications Engineering from the University of Vigo, who also made this research possible.

How to cite this article / Standardized reference:

Carral, Uxía, & Elías, Carlos (2024). Application of AI tools as methodology for the analysis of toxicity in social media: A case study of Spanish politics on Twitter [Aplicación de herramientas de IA como metodología para el análisis de la toxicidad en redes sociales: Estudio de caso de la política española en Twitter]. *Revista Latina de Comunicación Social*, 82, 01-18. https://www.doi.org/10.4185/RLCS-2024-2205

Receipt Date: 04/08/2023
Acceptance Date: 26/12/2023
Publication Date: 24/01/2023

ABSTRACT

Introduction: A new artificial intelligence (AI) methodology is analyzed with the understanding that communication is one of the most important fields of work for its application. In addition to the content collection and production phases, other areas within the world of communication such as distribution, and specifically the moderation of comments (on social networks and in the media) are also experiencing a period of innovation, but in a less obvious way for the audience. **Methodology:** To find out how various AI tools can measure the quality of the conversation and combat toxicity in communicative spaces, we analyzed 43,165 tweets published from October 18th to October 24th 2021 corresponding to seven Spanish politicians and the cascade of user responses. **Results:** The most significant outcome reveals insults as the predominant toxic category in the comments, regardless of ideology. In addition, the conversations have an average of 21% of bots. **Discussion:** Therefore, this research shows how new AI methodologies can account for a hitherto qualitative term such as toxicity and contradicts previous findings on bots as its spreaders, with real users generating the most. **Conclusions:** In the specific study of politics, there is a perceived difference in behaviors between horizontal conversation among peers and vertical conversations with politicians. Therefore, these tools help to make visible new realities such as toxicity, with the ultimate aim of eradicating it and cleaning up the online debate.

Keywords: Artificial intelligence; Toxicity;Bots; Social media; Twitter; Methodology; Perspective API.

RESUMEN

Introducción: Se analiza una nueva metodología de inteligencia artificial (IA), entendiendo que la comunicación se presenta como uno de los campos de trabajo más trascendentes para su aplicación. Además de las fases de recolección y producción de contenido, otras áreas dentro del mundo de la comunicación como la distribución, y en concreto la moderación de comentarios (en redes sociales y en medios) también están viviendo un período de innovación, pero de forma menos evidente para la audiencia. **Metodología:** Se procede a conocer cómo diversas herramientas de IA pueden medir la calidad de la conversación y combatir la toxicidad en espacios comunicativos. Se han analizado 43.165 tuits publicados del 18 al 24 de octubre de 2021 correspondientes a siete políticos españoles y a la cascada de respuestas de los usuarios. **Resultados:** Las principales consecuencias apuntan a los insultos como la categoría tóxica predominante en los comentarios, independientemente de la ideología. Además, las conversaciones cuentan con un promedio del 21% de usuarios *bots*. **Discusión:** Visto lo anterior, esta investigación muestra cómo nuevas metodologías de IA pueden contabilizar un término hasta ahora tan cualitativo como la toxicidad y contradice los hallazgos previos sobre *bots* como difusores de toxicidad, siendo los usuarios reales quienes más toxicidad generan. **Conclusiones:** En el estudio concreto de política, se percibe una diferencia de comportamientos entre la conversación horizontal entre pares y la vertical con los políticos. Por tanto, estas herramientas ayudan a visibilizar nuevas realidades como la toxicidad, con el fin último de llegar a erradicarla y sanear el debate online.

Palabras clave: Inteligencia artificial; Toxicidad; *Bots*; Redes sociales; Twitter; Metodología; Perspective API.

1. Introduction

In 1997, Goldhaber theorized about the "attention economy" applied to digitized communication processes. He pointed to the existence of a virtual sphere as a trigger for the growth of data monitoring to the point that mankind would not be able to cope with such an amount of information (Goldhaber, 1997). This phenomenon was already coined by Toffler (1975) in the 1970s as information overload, i.e., large amounts of information that are difficult to take in and, consequently, attention has become a scarce, limited and precious commodity to the detriment of information..

This new media scenario focused on capturing the attention of audiences has consolidated an ecosystem prone to the generation of toxicity (Southwell et al., 2018; Blanco-Alfonso et al., 2019). However, Silverman (2014) advocated the use of technology as a tool in the search for true information. That same conception shares the focus of this research work: to determine the level of toxicity and the nature of the audience in the online conversation through a new experimental methodology with various artificial intelligence (AI) tools.

For this purpose, the topic has been limited to the study of the Spanish political conversation on Twitter. However, other discursive arguments could be chosen in networks such as flat earthers or anti-vaccinationists, and even

"Con With an 80% CAP provided by the Botometer algorithm itself"

specific temporal contexts such as the assault on Capitol Hill in the United States. In short, this case study is a mere example that serves to test the proposed instruments, in an exploratory way, taking into account their possible limitations and errors, but also opening the door to the emergence of a new methodological technique for the study of social networks.

1.1. The origin of social media toxicity

Although the concept of fake news became popular during the election of Trump as president of the USA, Wardle and Derakhshan (2017) insist on extending the spectrum of 'misinformation' also to content with a false, impolite, hurtful, manipulated and/or impostor context created deliberately to do harm. This generic term 'toxicity'' thus encompasses fake news, but also other phenomena such as misinformation, hate speech, harassment, discrimination or cyberbullying, among others. Such toxicity can be transmitted through any means of communication, however, it is preferably spread through social networks and instant messaging, as they are the most effective channels for mass dissemination (Casero-Ripollés et al., 2016). Therefore, platforms such as Twitter, conceived as "the best democratization tool that our societies had developed", have come to be considered, a decade after their birth, "a great threat to Western democratic systems" (Magallón-Rosa, 2019, p. 53).

Based on this transformation, the digital public sphere has also seen its meaning redefined, giving way to distrust and the questioning of legitimacy. Bernard Williams explains in Truth and Truthfulness (2002) how contemporary thought is constantly on alert not to be deceived, so that the inclination to know the ins and outs of the facts is on the rise: "this interest in truthfulness promotes a process of criticism that undermines confidence in an objective truth at all". It also influences 'the conspiracist paradox' (Elías, 2019), whereby those who pay more active attention to manipulation coming from the media are more prone to be manipulated, as they reject conventional ways and tend to interact more often with toxic content, amplifying the scope of toxicity in online conversations.

Likewise, the public is skeptical of the information provided by the media due to polarization and misleading processes such as clickbait, born as a result of the digital communicative competition that arises in the environment of social networks. The result of this phenomenon Daniel Innerarity (2018) defines it as the 'uberization of truth', that is, the "deprofessionalization of information work" that has weakened the classic monopolies- university and press- to the benefit of social networks. With the implementation of the prosumer-user model in a participatory culture (Jenkins, 2006), the monopoly of the public agenda by the media and journalistic credibility is beginning to crack, giving rise to the post-truth era.

In this context, what is important is not the facts, what is important is the narrative (Frankfurt, 2006), i.e., objective facts are less influential in the formation of public opinion than the appeals to emotion and beliefs (Keyes, 2004) that we transmit in the narrative itself. By not dialoguing with reference to facts, citizens lose awareness of what the truth is. By debating opinions, we pervert what makes up the true reality and the parallel reality created by post-truth. In short, as expressed by the American philosopher, Lee McIntyre, in his work *Posverdad* (2018, p. 24): "if what is preached is the renunciation of values such as truth (...) or objectivity, the way is largely cleared to impose on citizens the interests of who is lying", encouraging the spread of lies, hatred and violence to certain groups, i.e., facilitating the intoxication of online communication.

Another fundamental change in the current social context has been the rise of algorithmic systems as content curation systems in social networks. Platforms such as Twitter have become the new gatekeepers, employing complex algorithms to select content based on the perceived behavior, engagement and relevance of their users, altering the way information flows and is consumed (Zuiderveen et al., 2018). This algorithmic control determines what content users see with the effective goal of shaping their online experiences. However, authors such as Noble (2018) and Barocas et al. (2023) highlight the potential risks associated with algorithmic gatekeeping noting that these algorithms may unintentionally amplify toxic content related to racial and gender traits.

1.2. The dilemma of technology control

A study, published in the scientific journal Science in 2018, processed 126,000 rumors spread by approximately 3 million people between 2006 and 2017. They concluded that, "in terms of daily rates of dissemination, fake news is 70% more likely to be retweeted than true news." It takes a true post "6 times longer to reach 1,500 users than a fake one" and, if it is a cascade of viral RTs, "a fake thread gets 10 to 20 times more interaction than a string of true facts" (Vosoughi et al., 2018, p. 1149). In addition, MIT Lab researchers pointed to the use of technology and, specifically, artificial intelligence tools as one of the main strategies by which fake content turned successfully viral. However, the so-called "high technology" (López-García and Vizoso, 2021) has a crucial impact on both the digitalization phase of newsrooms and the debates in journalism today.

The term "artificial intelligence" (AI) was born in 1956 during the Darmouth Conference, an ad hoc meeting of twenty intellectuals to develop machine learning. John McCarthy, one of the organizers, is credited with defining AI. According to a Stanford University compilation (2007, p. 2), the author himself calls AI "the science and engineering of making intelligent machines, especially intelligent computer programs, (...) and understanding human intelligence." Six decades after its invention, Google has concretized McCarthy's conceptualization by explaining which characteristics of human intelligence it aspires to replicate. In 2018, during the presentation of the ethical principles and limits of AI for the company, its CEO, Sundar Pichai, highlighted "problem solving, creativity and adaptability" as the capabilities that "a non-human mechanism" must demonstrate.

McCarthy's description highlights the importance of the human factor in the creation of machines, incorporating an element of "social constructivism". Bijker and Pinch (1987) argue that machines learn from a corpus biased by the cultural practices and values of the society that creates them, which gives control to human action. Technological advancement, therefore, is influenced by social change, in line with Kaplan (2009). On the other hand, the Google CEO's view advocates "technological determinism," where machines demonstrate their learning autonomously. McLuhan (1996) and Ellul (1962) also support the idea that technology determines social change. In technology, possible implies necessary; that which is ever available, will necessarily be used (Diéguez, 2005). This is how the arrival of personalizing algorithms such as Spotify has largely replaced music radios or how streaming services have changed our forms of consumption (binge watching) with respect to television and cinema.

Similarly, in the early days of social networks, the messages published had no impact on the media agenda; today, however, the platforms have acquired such relevance that a viral post (Carral et al., 2023) can open a newscast, change electoral results or trigger demonstrations against political power (Arab Springs, 15M in Spain, gilets jaunes). Therefore, many authors have questioned the role of human gatekeepers in a new media environment, where the content curation process in social networks, search engines or news aggregators is driven by algorithms, whose criteria derive from individual users (Jürgens et al., 2011; Meraz and Papacharissi 2013). In this way algorithms redistribute and channel information promoting new patterns in the flow of news (Wallace, 2017), shaping what is considered of public interest (Napoli, 2014) and participating in the process of construction of the new social reality (Just and Latzer 2016).

In the case of social networks, moreover, most of the spaces that offer decentralized gatekeeping mechanisms are open to everyone (Wallace, 2017). Control comes once the comment has already been published and remains at the expense of the audience reporting it in case it is offensive. Therefore, any toxic side of the conversation is partially born in consequence of that post-moderation process of the social network. However, human action also accompanies this transition from human to technological moderation. Positively, by making decisions such as the creation of legislation or the invention of other devices that reduce these pernicious effects, and negatively, by encouraging and amplifying incorrect behavior (harassment, insults, falsehoods, discrimination) on social networks.

Toxic discourse causes people to be less likely to participate in digital public spaces. Specifically, this silencing effect impacts more marginalized voices in society such as, for example, the LTGBIQ+ community (Martínez, 2022). Dialogue spaces for civic engagement such as the comments section in media articles or conversation threads in social networks and forums, are contaminated. In this way, toxic practices like disinformation campaigns or insults resulting from the high level of polarization endanger independent journalism, lead to lower user loyalty ratios, discourage advertisers, increase the costs to the media for hiring moderators or, to the extreme, cause the closure of these spaces for debate, preventing citizen participation, ideological pluralism and two-way media communication (Fuchs, 2021).

In this sense, the distribution area has benefited the most from the implementation of technology in the communication process with improvements in content personalization and recommendation, new content delivery formats (newsletters) and interaction and assistance to the audience (chatbots, wizards) (Sánchez-García et al., 2023). However, comment moderation is also undergoing a period of innovation, but in a less perceptible way for the audience due to its internal operation in the newsroom. Artificial Intelligence tools have been created to remove toxic content from comments and encourage healthier conversation. For example, Perspective API collaborates with The New York Times, EL PAÍS or Le Monde, among other media in the field of journalism, creating an intelligent artifact that finds patterns to detect abusive language and thus qualify comments according to toxicity. In this way, the task of classifying comments performed by the professional moderator is streamlined so that the media outlet can update the comments section in real time to its prosumer audience (Jigsaw, 2016).

Innovation has also reached other social worlds such as gaming. FACEIT, the leading platform in the industry also incorporated Perspective API to improve the work of human moderators, while encouraging new ways for the community to interact with each other free of harassment, but without stripping their users of their personalities (Jigsaw, 2019b). There is, however, a side of content moderation visible to the audience and focused on social platforms such as YouTube, Facebook, Twitter, Reddit, and Disqus. Given its decentralizing algorithmic system in moderation, through another experimental AI tool called Tune, users have an extension with which they can set the level of toxicity of conversations in social networks (Jigsaw, 2019a).

2. Objectives

In that sense, this research tries to know the performance of two AI tools through the study of a specific case such as the Spanish political conversation on Twitter. This case is a mere example to test the tools, still in the experimental phase, however, we could choose topics as varied as the analysis of the media, different athletes or clubs, other discursive arguments in networks such as anti-vaccine or climate change deniers, and even study specific temporal contexts such as 1-O in Catalonia or January 6, 2021 when the assault on Capitol Hill in the United States took place.

This topic has been specifically chosen for various reasons. Firstly, political leaders along with journalists are the most influential figures in defining the limits and contents that make up the public debate (Casero-Ripollés et al., 2016). Secondly, Twitter not only allows the study of the messages disseminated, but also the analysis of the online circles of support created to defend, promote or attack (Chadwick, 2013). Thirdly, the interest in the study of toxicity through the Perspective tool. A debate characterized by offensive language and fallacious arguments discourages citizen participation and increases the political polarization of extremist ideologies (Stryker et al., 2016).

Since the specific use of the Perspective tool in the world of communication is such a recent topic, there is hardly any exploratory background (Hosseini et al., 2017; Jain et al., 2018; Guerrero-Solé and Philippe, 200; Rieder and Skop, 2021) from which to extrapolate conclusions about the subject matter and to serve as a suggestion when stating hypotheses. For this reason, two exploratory research questions (RQ) have been proposed instead in relation to the objective explained above:

- RQ1 How toxic is the conversation about Spanish politics on Twitter?
- RQ2 What characterizes the audience that shapes these online discussions?

3. Methodology

3.1. Subject matter

In order to answer these questions, a period of time belonging to the context of an ordinary online conversation has been selected for the study. Specifically, the authors have avoided the choice of a time period conditioned by determining events such as an electoral campaign, since it would not be appropriate to extrapolate the results of such an agitated and unprecedented time as a representative conclusion of the dynamics of the daily debate. Likewise, seeking the greatest possible neutrality in the choice of the study cases, the authors determined to analyze all the Twitter profiles of the leaders of the parties that would have obtained more than 1% of the votes in the general elections in Spain on November 10, 2019: Santiago Abascal (Vox), Inés Arrimadas (Cs), Ione Belarra (Unidas Podemos), Pablo Casado (PP), Íñigo Errejón (Más País-EQUO), Gabriel Rufián (ERC) and Pedro Sánchez (PSOE).

Therefore, faced with a scenario of such plural themes and voices, the authors chose to use a hybrid quantitative methodology consisting of computational and manual procedures for the analysis of an object of study n=43,165 tweets. The corpus was composed of the tweets published from October 18th to 24th, 2021 corresponding to seven Spanish politicians and the cascade of replies from their audiences that make up the totality of each thread of conversation. It is worth clarifying that the authors decided to focus on a daily tweet, and on the one with the most replies, because the interest of the research does not lie in the politician's message, but in the audience's comments on each publication. Previous studies (Lampe and Johnston, 2005) show that replies "increase the chances of starting a conversation and creating a community" of "engaged and motivated users for future participation" (Burke et al., 2010, p. 4).

It should also be noted that only messages published in Spanish were analyzed, since the natural language processing tools used are not yet available for languages such as Catalan or Basque. For this technical reason, Mertxe Aizpurua (EH Bildu), Laura Borrás (JuntxCat), Aitor Esteban (PNV) and Mireia Vehí (CUP) were excluded. However, despite the fact that the collection and the first analysis was completed with the support of different tools, due to its experimental operation in Spanish, the process required human supervision to correct certain errors caused by biases in Botometer and to make up for the limitations of the free version of Communalytic. Likewise, the crossing of variables analyzed by different tools also involved manual calculations.

3.2. Tools and techniques for analysis

Once the sample was selected, the authors proceeded to collect the study material and ran the toxicity analysis on each dataset, with prior authorization for the use of Google's API, Perspective. This AI tool relies on machine learning models in order to measure the level of toxicity of a post and the impact it could have on the conversation. The platform's own categorization and detection capabilities produce biases that the creators of Perspective API have been reducing since its launch in 2016. False positives and false negatives represent the priority problem, with "identity-related messages receiving inappropriately high or low toxicity scores." Terms such as 'black', 'Muslim', 'feminist', 'woman' or 'gay' "tend to have higher scores because comments about those groups are overrepresented in abusive and toxic reviews" and, conversely, "male", "white" or "straight" receive lower scores (Jigsaw, 2021). In addition, the tool has certain limitations: its experimental nature in non-English languages and the inability to interpret specific irony or toxicity from slang.

In the next step and using Botometer, the authors examined the users who participated in all the conversation threads of the object of study. The machine learning algorithm was developed in 2017 by the Observatory of Social Media (OSoMe), in the University of Indiana. It is trained by comparing several machine learning models, whose algorithms extract the characteristics (features) of the account profile, friends, social network structure, temporal activity patterns, reach, language and message sentiment. Finally, Botometer gives a score between 0 and 1 to each user with the ultimate goal of determining whether behind each profile hides a real person (0) or a bot (1), asserting it with a CAP (complete automation probability) of 80%. In other words, 20% out of all the accounts identified as bots in the sample may be false positives. This percentage includes cases such as corporate accounts that publish through desktop applications, since the algorithm can lead to confuse these actions with those typical of a bot. However, a manual human verification of the sample has managed to reduce this margin of error.

However, one of the goals of this work is not only focused on the nature of the users, but also on their function. For this reason, the authors decided to implement the six categories of analysis in Spanish that Botometer offers, thus obtaining a score between 0 and 1 for each of them:

- Echo chambers (eco-chamber bot): accounts that participate in monitoring groups and share or delete political content in high volume.
- Fake followers: bots bought to increase the number of followers of a profile.
- Financial: bots that post using cashtags.
- Self-declared: bots created from the botwiki.org website.
- Spammer: accounts that massively send data sets or information.
- Other: as it is not specified under which criteria it is acceptable to flag accounts as bots, the authors decided not to work with this category in order to keep the transparency and rigor of the analysis.

After confirming the presence of bots among the users participating in the conversations on political leaders' tweets, the authors proceeded to study how these automated bots relate to non-bot users. To this end, the authors have proposed two categories for each politician: a ranking of the 10 users most likely to be bots according to Botometer and another top 10 users who have interacted the most in their threads.

4. Results

Authors such as Moreno-López and Arroyo-López (2022) have previously studied the issue of toxicity and hate speech in networks through a qualitative methodology such as in-depth interviews and questionnaires. This research, on the other hand, aims at reviewing those concepts in a quantitative way and highlighting a new methodology with AI tools. To answer the first research question, the authors analyzed certain negative characteristics affecting the content. A second, more thorough review allowed the authors to recognize the particularities of the most relevant categories when it comes to preventing a good quality online conversation. Then, in order to answer the second research question, the authors searched for possible bots that interfere with the actual conversation. The authors then compared these data with the statistics of the users who were the backbone of the multiple discussions that arose as a result of the politician's tweet. This procedure aims to find out possible relationships between bots, core prosumers and toxicity.

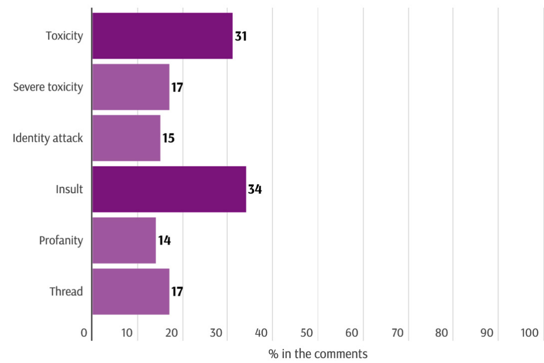
4.1. RQ1: How toxic is the Spanish political conversation on Twitter?

'Toxicity' is defined as the quality of a comment to be rude, disrespectful or unreasonable to point of making another user decide to leave the conversation. This is how the term is described by the very AI tool used in this research, Perspective, created to facilitate conversation and prevent online toxicity and harassment. However, in Spanish, Perspective offers five more attributes in which to categorize user comments: severe toxicity, when the comment incites hatred or is aggressive; identity attack, expressly directed for identity reasons; insult, understood as an inflammatory and negative comment against an individual or group; blasphemy, if obscene or profane language is used; and threat, when it is considered that there is an intention to inflict pain or violence against others.

Therefore, in the first phase of the study, the authors have focused on the attributes. It should be noted that the analysis only includes textual units, so that photographs, gifs, emoticons and audio and/or video content are outside the scope of inspection. The machine learning model does not detect which spelling errors or letters separated by orthographic signs can be understood as an insult, enhancing the bias of 'false negatives', as already demonstrated by Jain et al., (2018). Also, it should be remembered that the tool has been trained with negative comments in any domain, not just political ones.

If the six attributes are analyzed in isolation (Figure 1), the levels of profanity (14%), identity attacks (15%), severe toxicity (17%) and threats (17%) are below the overall average. However, if attention is focused on the degree of toxicity, the result increases to 31% of toxic comments on average. In the same upward direction, Perspective also seems to point out that insults are the most worrying category, with one in three posts containing negative and inflammatory messages against an individual or group on average. However, it should be noted that the results could be biased by the algorithm's own detection capabilities. That is, it may be less sensitive to detect one category than another, leading us, in this case, to overestimate the presence of insults compared to other types.

Figure 1: *Attributes identified by Perspective and their percentage.*



Source: Elaborated by the authors.

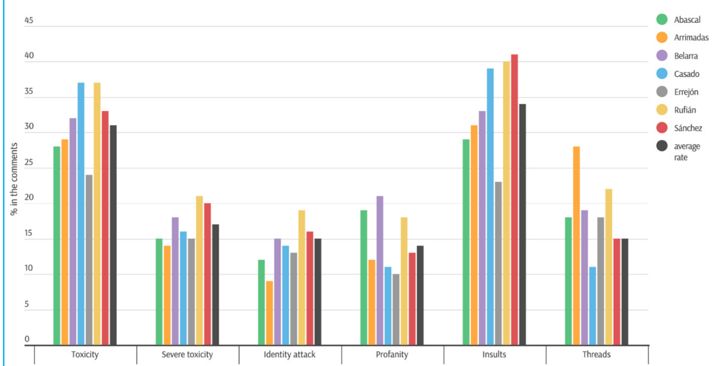
In a second phase of analysis, the authors compared the averages of the attributes of each individual politician. This reveals an upward percentage trend, but one that is consistent with the dynamics predicted in the previous stage. In fact, in up to six of the seven profiles analyzed, the attribute with the highest average is linked to the appearance of insults in the users' responses to the initial tweet, as suggested by Perspective. And in the exceptional case, that of Errejón, its maximum average is associated with toxicity, the second attribute with the highest average. As we observe in Figure 2, the metrics at the individual level indicate a similar propensity to the joint trends, although two striking particularities are worth highlighting.

First, the conversations that occur as a result of the tweets of the leader of ERC, Gabriel Rufián, exceed the overall averages in any of the six attributes. Even in the category of insults, if the overall average reached 34%, his personal average rises to 40%. Second, the statistics of the President of the Government also exceed the joint averages in four of the six attributes. Specifically, Pedro Sánchez's profile has the highest average percentage (41%) of any attribute, with some days in which 50% of the messages, i.e. half of all replies to his tweet, contain insults.

In relation to this attribute, when exploring the critical points, i.e., the maximum level indexes of any of the six attributes, Perspective's enunciation of insults is consolidated. The highest percentage of negative comments in as many as six of the seven subjects analyzed is linked to insults. Moreover, the maximum proportion of insults suffered by each politician is in all cases higher than the average for the category, established at 34%.

Once again, Rufián, Sánchez and Casado stand out, with alarming percentages. In up to two days of the week studied, 50% of the

Figure 2: *Percentage of each attribute in the comments.*

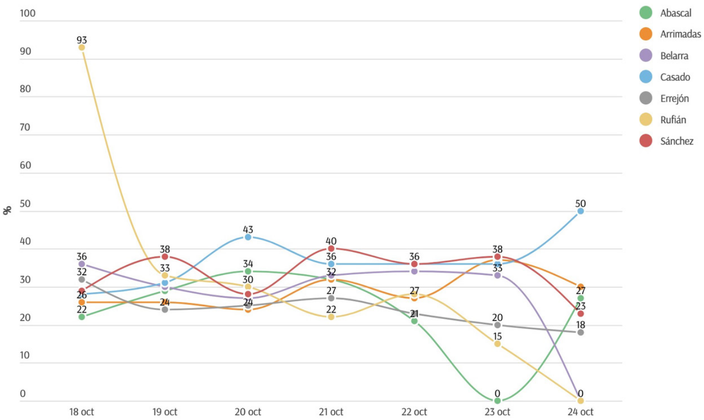


Source: Elaborated by the authors.

comments regarding the initial tweet of both the leader of the opposition at the time and the president of the Government were considered insults by Perspective. The exception falls on Belarra, whose maximum peak is associated with the category of 'threats'. The leader of "Podemos" received intimidating comments in 40% of the replies to her tweet, in which she valued the apologies pronounced by Arnaldo Otegui in reference to ETA's crimes. However, the subject matter of Belarra's publication fits perfectly with the messages of the rest of the political leaders with more negative comments.

Since the results seem to confirm the serious problem of insults in the conversation about Spanish politics on Twitter, as pointed out by Perspective, the authors have examined the data concerning the attribute of 'toxicity' in detail. Following the trend set by the global indicators, we can group some political leaders according to the behavior of the audience. On the one hand, the analysis of the degree of toxicity shows that Arrimadas (29%), Abascal (28%) and Errejón (24%) are below the average of toxic comments (31%). Even the leader of Vox and the leader of Más País-EQUO have the lowest percentages of all the leaders in their most negative publications, with 34% and 32% of toxicity.

Figure 3: *Evolution of the toxicity degree in the tweets of each political leader.*



Source: Elaborated by the authors.

On the other hand, Rufián heads the group of politicians with the most toxic conversation threads. His tweets have an average toxicity rate of 37%, although it is worth noting that, at its peak, Perspective considered that up to 93 out of every 100 comments were toxic. It is relevant to take into account the number of publications on which these estimates have been made. This reveals how the profiles of Rufián and Belarra are the ones that receive the most responses, exceeding 10,000 replies. Casado and Sánchez get half; Abascal and Errejón, three times less and Arrimadas, only one tenth. Taking into account these data, a parallel relationship is observed in the audience's behavior towards Rufián and Belarra. Both have the personal average on the degree of toxicity above the average (Figure 3) and, in addition, both have the highest toxicity minimums: in their least toxic publications, the volume of negative comments does not fall below 27%.

Casado, like Rufián, has an average toxicity rating of 37% and his most convulsive tweet also exceeds the group average of toxicity. In fact, 1 out of every 2 responses in the most heated publication of the popular candidate contained toxicity. However, Casado's results show a normal statistical distribution, unlike those of Rufián, which result in a positive asymmetry. While for the Catalan politician there is a difference of 78 percentage points between his minimum and maximum toxicity, for the popular politician the range is reduced to 22 points. This data shows how Casado has a high degree of toxicity, but stable and continuous over time, while Rufián suffers tremendously toxic jolts (maximum toxicity: 93%) alternating with other rather calmer spells (minimum: 15%).

4.2. RQ2: What characterizes the audience that shapes these online discussions?

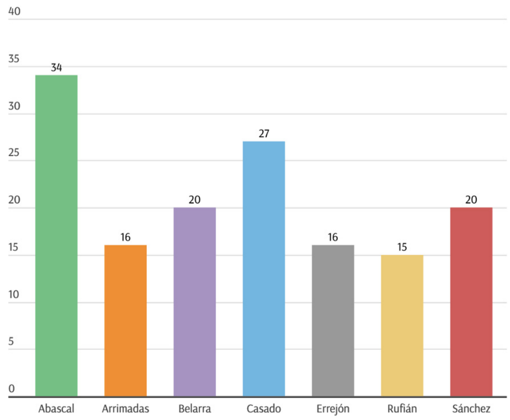
In the first research question, the authors studied how the audience behaves in relation to politicians' messages. Next, the authors will proceed to find out who is this prosumer audience, i.e., who is this audience that consumes the messages as well as participates in the conversation by producing other tweets. To this end, the authors have tried to determine whether they are real users or, on the contrary, bots programmed to disseminate hurtful or confusing content; how they relate to each other and at what point in the conversation most toxicity is created.

Con With an 80% CAP (complete automation probability) provided by the Botometer algorithm itself, the conversations studied between Monday, October 18th and Sunday, October 24th, 2021 had an average of 21% bot users. There is no universally agreed definition to designate what a bot is due to the wide range of behaviors they can have. However, the creators of Botometer (Yang et al., 2020) determine that a bot is a "social networking account controlled, at least in part, through software." Despite the fact that there are many types of bots, malicious ones can be used to manipulate social network users by amplifying misinformation, creating the appearance that some people or ideas are more popular, committing financial fraud, or spreading spam, among others (Schuchard et al., 2019). For this reason, the authors decided to study both the presence of bots and the role they can play thanks to Botometer's six categories of analysis in Spanish.

Thus, it has been found that Abascal is the politician with the highest presence of possible bots in his conversations (average 34%). In second place is Casado, with an average of 27% of bots participating in his threads (Figure 4). And, tied with an average of 20%, Sánchez and Belarra. The leader of "Podemos", however, tops the ranking among self-declared bots, as one in five accounts commenting on her are bots created through the botwiki.org project. All the other political leaders have a lower average of self-reported bots, but none lower than Errejón's 12%. In fact, as with the levels of toxicity, the leader of Más País-EQUO barely arouses relevance among the audience of bots (16%). Contrary to what happens with Rufián who, despite his high levels of toxicity, has the lowest average with 15% of bots among his audience.

In this sense, it is paradoxical to see how the tables have turned. On the one hand, Rufián is also the leader with the fewest false followers (3%), the only one below the average of 5%. Opposite, Arrimadas and Sánchez have 7% and 6% respectively; which indicates that the leader of Ciudadanos has almost 50,000 bots followers and the President of the Government is followed by 96,000 fake users. On the other hand, the leader of Vox, who hardly stood out in the toxicity analyses, is once again the first, this time, reaching the maximum peak of bots. On October 22nd, up to 86% of users commenting on Abascal's tweet, according to Botometer, could be considered bots. Casado's tweet on October 18th also had 81% of bots among the users who commented on it. These two examples are noteworthy since, in comparison, the maximum number of bots of the other politicians never exceeded 50%.

Figure 4: *Number of bots commenting on each politician's tweets.*



Source: Elaborated by the authors.

Finally, despite the fact that Abascal has a larger number of bots among those who comment on him, Casado is the leader with the most bots that function as echo chambers. Some 22% of the profiles that respond to him or participate in the conversations resulting from his thread have the intention of sharing or deleting political content on a large scale. This piece of data means that the popular leader has more than twice as many eco-chamber bots as Sánchez (11%) or, for example, seven times more than Arrimadas (3%). In general terms, it has been concluded that the 10 most participative users of each politician are not identified with a bot profile, in other words, there is no evidence that bots are the ones who comment the most in the conversations created as a result of tweets published by politicians.

Also, given that in profiles such as Casado's, eco-chamber bots account for up to 22%, the authors wondered whether their function of sharing or deleting political content in high volume may involve toxic content. Combining the data obtained from the two AI tools, Perspective and Botometer, the authors have investigated whether it is bots who publish the most toxic messages. It is worth noting that the 10 users that

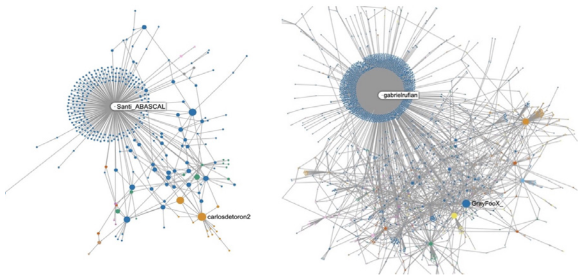
"50% of the comments regarding the initial tweet of both the leader of the opposition at the time and the president of the Government were considered insults by Perspective"

Botometer considers most likely to be bots were selected, on the one hand, and, on the other hand, the replies to the tweet containing at least 80% toxicity. As a result, only 9% of these toxic publications were issued by a bot account included in the top 10 of each politician.

Finally, as in the first research question the authors studied in broad terms whether the audience emits toxic comments towards the tweet published by the politician, in this second question they have tried to delve deeper into the issue by finding out where the most toxicity is created. To answer this question, the authors have taken two moments as a reference and compared them. The first is equivalent to the natural state (state 0) of the conversation, measuring all reactions to the politician and interactions between users. The second is given by the degree of toxicity the authors want to study. Since they have worked with a toxic level of 80% in the rest of the questions, they have also chosen that percentage for the comparison.

Thus, as reflected in Figure 5 of the Abascal and Rufián graphs, in a first moment (state 0) there are relations of toxicity- however small the degree- both in the comments addressed directly to the political leader and in the conversations that arise from the interaction between users. In fact, the creation of certain clusters is evident, a concentration of interconnected users in the second and subsequent levels of conversation, not in the main thread. In Abascal's graph, for example, @carlosdetoron2 stands out, who captains a small community in addition to expressing himself individually and directly towards the leader of Vox. In Rufián's graph, @GrayFooX_ leads the secondary conversations, although other users can also be seen generating small clusters (in yellow, orange, green, pink).

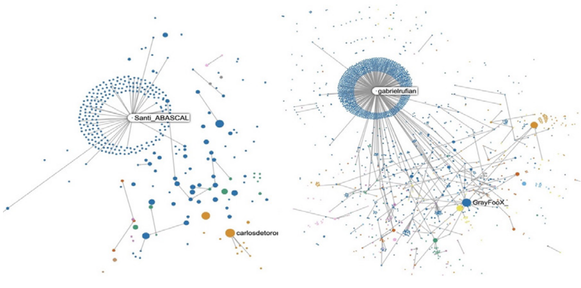
Figure 5: *Interactions in state 0 of Abascal and Rufián.*



Source: Elaborated by the authors.

In contrast, when user participation is limited to relationships with at least 80% toxicity in the comments, it is clear that the most toxic content is focused on direct messages to the leader who posted the tweet. Most of the interactions between users disappear, since their feedback does not contain such high levels of toxicity. Nevertheless, the champions of each cluster remain, even when these users are not the ones issuing toxic comments, since being the centralizers of the secondary conversations, they are included in the discussion -just like the politician- by those who post in a toxic way. Thus, we see in Figure 6 how @carlosdetoron2 retains certain interactions with 80% or more of toxicity in the Abascal thread and the same happens with @GrayFoox as a result of Rufián's tweet.

Figure 6: *Abascal and Rufian´s interactions with at least 80% of toxicity.*



Source: Elaborated by the authors.

5. Discussion and Conclusions

In light of the above-mentioned data, this research shows how new AI methodologies can, first, account for a term so far as abstract as toxicity and, second, help to bring clarity to a topic as deep and unknown as the nature and behaviors of the audience in social networks.

In order to answer RQ-1 on the degree of toxicity of the conversation about Spanish politics on Twitter, it has been necessary to carry out an in-depth study, where the results provided by Perspective reveal that the conversation about political issues on Twitter is circumscribed in a toxic environment, where, for example, around one in three responses are insults (34%). Moreover, it can be confirmed that this is an intrinsic feature of the conversation. In the data for each political leader, the responses containing insults from users to the initial tweet are the most popular category in up to six of the seven profiles analyzed, i.e., we are not dealing with a specific targeting of certain politicians because of their gender or their particular ideology. However, it is important to remark two trends in the prevalence of this phenomenon: there are profiles such as Casado, whose conversation has a high degree of toxicity, stable and continuous over time, while Rufián only suffers shocks with a high degree of toxicity at specific moments. In this sense, a possible line of future research could be to study the reasons of such behaviors, analyzing whether the patterns are due to the subject or whether the topic has the same impact on other individuals.

Likewise, RQ-2 of the present research work asked about the characteristics of the audience that makes up these online discussions. The results indicate the existence of fictitious users in the conversation, since all the Twitter threads analyzed had, on average, 21% of bots. However, it is shown that there is no association between the most active bots in the conversations of each politician and the most toxic posts, as only 9% of the toxic responses were issued by bots. Therefore, these findings support the same theory of previous studies such as Howard et al. (2017), who argue that it is the bots who originally spread more toxic content, but at the same time it is people who share the posts the most. In contrast, the results of this research partially contradict previous findings (Caldarelli et al., 2020; Shao et al., 2018). These accounts have an undeniable spreading capacity, but results do show that it is not the bots, but human users instead who most often engage with content with a higher level of toxicity, thus extending their reach.

Likewise, the granularity of the data provided by these AI tools makes it possible to individually investigate each subject in the sample until it is possible to know, for example, that Abascal is the politician with the highest presence of bots (average of 34%) and that the maximum peak of bots was also recorded in one of his conversations, with 86% of fake users. Thus, the study of bots can be beneficial and useful for areas such as political communication, especially in the creation of marketing strategies. In addition, the core of academic researchers interested in filter bubbles can make use of these tools to analyze the nature of different groups of users. This case study shows, for example, how Casado is the leader with the most bots (22%) that function as echo chambers.

Finally, the same general trend was observed in all the conversations: commenting users direct their toxicity towards the leader who publishes the tweet through direct comments. The higher the degree of toxicity, the less interaction between users. On the one hand, this pattern reflects an amount of negative direct comments against the politician and, on the other hand, a better-quality conversation between peers. Regarding the first behavior, the negative comments against the politician, the customs and cultural contexts of each civilization must be taken into consideration. Contrary to the Western world, studies such as Feldman's (2023) show that, in the East, specifically in Japan, demeaning remarks by Japanese leaders are rarely said face to face to the addressee, but in front of general audiences who are not the target of the demeaning.

In short, these AI tools, still in the experimental phase, help to quantify new realities in social networks. This case is a mere example to test the tools, however, one could choose topics as varied as the analysis of the media, different athletes or clubs, other discursive arguments in networks such as anti-vaccine or climate change deniers, and even study specific temporal contexts such as 1-O in Catalonia or January 6th, 2021 when the assault on Capitol Hill in the United States took place. However, this research rests on a central pillar: the presentation of a new methodology for the academic study within and about social networks through new variables such as toxicity - in all its aspects, and delve into the disclosure of bot algorithms and their connection with such toxicity. Thus, highlighting the variety of uses for both media content moderation and social networks.

6. References

Barocas, S., Hardt, M. y Narayanan, A. (2023). *Fairness and machine learning: Limitations and Opportunities*. MIT Press.

Bijker, W. E. y Pinch, T. (1987). *The social construction of facts and artifacts*. https://acortar.link/Y4RqvA

Blanco-Alfonso I., García-Galera C. y Tejedor-Calvo S. (2019). El impacto de las fake news en la investigación en Ciencias Sociales. Revisión bibliográfica sistematizada. *Historia y Comunicación Social, 24*(2), 449-469. https://doi.org/10.5209/hics.66290

Burke, M., Kraut, R. y Joyce, E. (2010). Membership claims and requests: Conversation-level newcomer socialization strategies in online groups. *Small group research, 41*(1), 4-40. https://doi.org/10.1177/1046496409351936

Caldarelli, G., De Nicola, R., Del Vigna, F., Petrocchi, M. y Saracco, F. (2020). The role of bot squads in the political propaganda on Twitter. *Commun Phys*, 3. https://doi.org/10.1038/s42005-020-0340-4

Carral, U., Tuñón, J. y Elías, C. (2023). Populism, cyberdemocracy and disinformation: analysis of the social media strategies of the French extreme right in the 2014 and 2019 European elections. *Humanit Soc Sci Commun 10*, 23. https://doi.org/10.1057/s41599-023-01507-2

Casero-Ripollés, A., Feenstra, R. A. y Tormey, S. (2016). Old and New Media Logics in an Electoral Campaign: The Case of Podemos and the Two-Way Street Mediatization of Politics. *The International Journal of Press/Politics, 21*(3), 378-397. https://doi.org/10.1177/1940161216645340

Chadwick, A. (2013). *The hybrid media system: Politics and power*. Oxford University Press

Diéguez, A. (2005). El determinismo tecnológico: indicaciones para su interpretación. *Argumentos de Razón Técnica*, 8, 67-87. http://www-formal.stanford.edu/jmc/whatisai.pdf

Elías, C. (2019). *Science on the Ropes. Decline of Scientific Culture in the Era of Fake News*. Springer-Nature. https://doi.org/10.1007/978-3-030-12978-1

Ellul, J. (1962). The Technological Order. *Technology and Culture, 3*(4), 394-421. https://doi.org/10.2307/3100993

Feldman, O. (2023). Challenging Etiquette: Insults, Sarcasm, and Irony in Japanese Politicians' Discourse. En O. Feldman (Ed.), *Political Debasement. The Language of Politics*. Springer. https://doi.org/10.1007/978-981-99-0467-9_5

Frankfurt, H. (2006). *On Bullshit: sobre la manipulación de la verdad*. Paidós

Fuchs, C. (2021). *Social media: A critical introduction*. Sage.

Goldhaber, M. H. (1997). The attention economy and the Net. *First Monday, 2*(4). https://doi.org/10.5210/fm.v2i4.519

Guerrero-Solé, F. y Philippe, O. (2020). La toxicidad de la política española en Twitter durante la pandemia de la COVID-19. *Hipertext.net*, 21, 133-139. https://doi.org/10.31009/hipertext.net.2020.i21.12

Hosseini, H., Kannan, S., Zhang, B. y Poovendran, R. (2017). *Deceiving Google's perspective API built for detecting toxic comments*. Cornell University. https://doi.org/10.48550/arXiv.1702.08138

Howard, P. N., Bradshaw, S., Kollanyi, B. y Bolsolver, G. (2017). Junk News and Bots during the French Presidential Election: What Are French Voters Sharing Over Twitter in Round Two? *ComProp data memo, 21*(3). https://acortar.link/IUZfrN

Innerarity, D. (2018, 31 diciembre). El año de la volatilidad. *El País*. https://elpais.com/elpais/2018/12/28/opinion/1546021545_365361.html

Jain, E., Brown, S., Chen, J., Neaton, E., Baidas, M., Dong, Z. y Artan, N. S. (2018, diciembre). *Adversarial Text Generation for Google's Perspective API*. 2018 International Conference on Computational Science and Computational Intelligence (CSCI), (pp. 1136-1141). IEEE. http://doi.org/10.1109/CSCI46756.2018.00220

Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. NYU Press.

Jigsaw. (2016, septiembre 19). *New York times and Jigsaw partner to scale moderation platform*. Medium. https://acortar.link/t1R0hG

Jigsaw. (2019a, marzo12). *Tune: Control the comments you see*. Medium. https://acortar.link/ACX5fU

Jigsaw. (2019b, octubre 23). *One of Europe's largest gaming platforms is tackling toxicity with machine learning*. Jigsaw. https://acortar.link/KgcA2i

Jigsaw. (2021, febrero 10). *Helping authors understand toxicity, one comment at a time*. Medium. https://bit.ly/3KgjnSz

Jürgens, P., Jungherr, A. y Schoen, H. (2011). Small worlds with a difference: New gatekeepers and the filtering of political information on Twitter. *Proceedings of the 3rd international web science conference*. https://doi.org/10.1145/2527031.2527034

Just, N. y Latzer, M. (2017). Governance by algorithms: reality construction by algorithmic selection on the Internet. *Media, culture & society, 39*(2), 238-258. htpps://doi.org/10.1177/0163443716643157

Kaplan, D. M. (Ed.). (2009). R*eadings in the Philosophy of Technology*. Rowman & Littlefield Publishers.

Keyes, R. (2004). *The Post-Truth Era: Dishonesty and Deception in Contemporary Life*. St. Martin's Press.

Lampe, C. y Johnston, E. (2005). Follow the effects of feedback on new members in an online community. *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. https://doi.org/10.1145/1099203.1099206

López-García, X. y Vizoso, Á. (2021). Periodismo de alta tecnología: signo de los tiempos digitales del tercer milenio. *Profesional de la Información, 30*(3). https://doi.org/10.3145/epi.2021.may.01

Magallón-Rosa, R. (2019). *Unfaking news: cómo combatir la desinformación*. Pirámide.

Martínez Valerio, L. (2022). Mensajes de odio hacia la comunidad LGTBIQ+: análisis de los perfiles de Instagram de la prensa española durante la "Semana del Orgullo". *Revista Latina de Comunicación Social*, 80, 364-388. https://doi.org/10.4185/RLCS-2022-1749

McCarthy, J. (2007). *What is Artificial Intelligence?* Stanford University.

McIntyre, L. (2018). *Post-truth*. MIT Press.

McLuhan, M. (1996). *El medio es el masaje. Un inventario de efectos*. Paidós.

Meraz, S. y Papacharissi, Z. (2013). Networked gatekeeping and networked framing on# Egypt. *The international journal of press/politics, 18*(2), 138-166. https://doi.org/10.1177/1940161212474472

Moreno-López, R. y Arroyo-López, C. (2022). Redes, equipos de monitoreo y aplicaciones móvil para combatir los discursos y delitos de odio en Europa. *Revista Latina de Comunicación Social*, 80, 347-363. https://doi.org/10.4185/RLCS-2022-1750

Napoli, P. M. (2014). Automated media: An institutional theory perspective on algorithmic media production and consumption. *Communication theory, 24*(3), 340-360. https://doi.org/10.1111/comt.12039

Noble, S. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.

Pichai, S. (2018, 7 de junio). AI at Google: Our Principles. *Blog Google.* https://blog.google/technology/ai/ai-principles/

Rieder, B. y Skop, Y. (2021). The fabrics of machine moderation: Studying the technical, normative, and organizational structure of Perspective API. *Big Data & Society, 8*(2), https://doi.org/10.1177/2053951721104618

Sánchez-García, P., Merayo-Álvarez, N., Calvo-Barbero, C. y Diez-Gracia, A. (2023). Spanish technological development of artificial intelligence applied to journalism: companies and tools for documentation, production, and distribution of information. *El Profesional de la Información, 32*(2). https://doi.org/10.3145/epi.2023.mar.08

Schuchard, R., Crooks, A., Stefanidis, A. y Croitoru, A. (2019). Bots fired: examining social bot evidence in online mass shooting conversations. *Palgrave Communications, 5*(1), 1-12. https://doi.org/10.1057/s41599-019-0359-x

Shao, C., Ciampaglia, G.L., Varol, O., Yang, K-C, Flammini A. y Menczer, F. (2018). The spread of low-credibility content by social bots. *Nat Commun*, 9. https://doi.org/10.1038/s41467-018-06930-7

Silverman, C. (2014). *Verification Handbook*. European Journalism Centre.

Southwell, B. G., Thorson, E. A. y Sheble, L. (Eds.). (2018). *Misinformation and mass audiences*. University of Texas Press.

Stryker, R., Conway, B. A. y Danielson, J. T. (2016). What is political incivility? *Communication Monographs, 83*(4), 535-556. https://doi.org/10.1080/03637751.2016.1201207

Toffler, A. (1975). *Alvin Toffler*. Pacifica Tape Library.

Tune. (s/f). *Tune experimental*. Google.com. https://acortar.link/bvJqZk

Vosoughi, S., Roy, D. y Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146-1151. https://doi.org/10.1126/science.aap9559

Wallace, J. (2017). Modelling contemporary gatekeeping: The rise of individuals, algorithms, and platforms in digital news dissemination. *Digital Journalism, 6*(3), 274-293. https://doi.org/10.1080/21670811.2017.1343648

Wardle, C. y Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Strasbourg: Council of Europe. https://acortar.link/dsM2G5

Yang, K. C., Varol, O., Hui, P. M. y Menczer, F. (2020). Scalable and generalizable social *bot* detection through data selection. *Proceedings of the AAAI conference on artificial intelligence, 34*(1), 1096-1103. https://doi.org/10.1609/aaai.v34i01.5460

Zuiderveen, F. B., Trilling, D., Möller, J., Bodó, B., De Vreese, C. H. y Helberger, N. (2016). Should we worry about filter bubbles? Internet Policy Review. *Journal on Internet Regulation, 5*(1). https://doi.org/10.14763/2016.1.401

AUTHORS' CONTRIBUTIONS, FUNDING AND ACKNOWLEDGEMENTS

Authors' contributions:

Conceptualization: Carral, Uxía y Elías, Carlos. **Formal analysis:** Carral, Uxía y Elías, Carlos. **Data curation:** Carral, Uxía. **Drafting-Preparation of the original draft:** Carral, Uxía y Elías, Carlos. **Drafting-Revision and Editing:** Elías, Carlos. **Visualization:** Carral, Uxía y Elías, Carlos. **Supervision:** Elías, Carlos. **Project management:** Elías, Carlos. **All authors have read and accepted the published version of the manuscript:** Carral, Uxía y Elías, Carlos.

Funding: The present text appears within the framework of the two projects granted to members of Madrid's Carlos III University from which this research receives funding: 'Pseudoscience, conspiracy theories, fake news and media literacy in health communicatio' (Ref: PID2022-142755OB-I00) within the National Plan 'Knowledge Generation Projects' of the Ministry of Science and Innovation and 'EU, disinformation and fake news' (Ref: 610538-EPP-1-2019-1-ES-EPPJMO-CHAIR), a Jean Monnet Chair of the European Commission.

Acknowledgments: The authors would like to sincerely thank Rubén Míguez Pérez, PhD in Telecommunications Engineering from the University of Vigo, who also made this research possible.

AUTHORS:

Uxía Carral Vilar
Carlos III University of Madrid. Spain.

Predoctoral student in Media Research at Universidad Carlos III de Madrid (UC3M). Her research interests are social networks, innovation in communication and technological humanism. She has made stays at the University Technology of Sydney (UTS) and La Sapienza (Rome, Italy). She has also worked as a journalist in media such as Cadena SER and verification as Newtral. In addition, she has managed the communication of several European projects funded by the European Commission and has collaborated as a data visualization journalist at the *Círculo de Análisis Euromediterráneo* (CAEM). Author of several books and chapters on disinformation, populism, social networks and the European Union.

ucarral@hum.uc3m.es

H-index: 4
Orcid ID: https://orcid.org/0000-0002-2329-3331
Google Scholar: https://scholar.google.es/citations?user=ART9ru0AAAAJ&hl=es&oi=ao
ResearchGate: https://www.researchgate.net/profile/Uxia-Carral-2
Academia.edu: https://uc3m.academia.edu/UxiaCarral

Carlos Elías Pérez
Carlos III University of Madrid. Spain.

Summary of the author's curriculum vitae of 120 words.

Professor of Journalism at Universidad Carlos III de Madrid (UC3M) and Jean Monnet Chair "EU, disinformation and fake news". He specialized in science, technology and public sphere in postdoctoral stays of one year at the London School of Economics and another year at Harvard. He has worked as a journalist at Agencia Efe (politics) and El Mundo (responsible for science). His latest book is Science on the Ropes. Decline of Scientific Culture in the era of Fake News (Springer, 2019). He is director of the master's degree in Corporate and Institutional Communication at UC3M and is a regular contributor to media such as El Mundo, RNE and RTVE.
carlos.elias@uc3m.es

H-index: 20
Orcid ID: https://orcid.org/0000-0002-1330-4324
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=31267503700
https://www.scopus.com/authid/detail.uri?authorId=57693924900
Google Scholar: https://scholar.google.es/citations?user=jiCVGcIAAAAJ&hl=es
ResearchGate: https://www.researchgate.net/profile/Carlos-Elias-3

Revista Latina de Comunicación Social

RELATED ARTICLES

Aramburú Moncada, L. G., López Redondo, I., & López Hidalgo, A. (2023). Inteligencia artificial en RTVE al servicio de la España vacía. Proyecto de cobertura informativa con redacción automatizada para las elecciones municipales de 2023. *Revista Latina de Comunicación Social*, 81, 1-16. https://doi.org/10.4185/RLCS-2023-1550

Deliyore Vega, M. D. (2021). Redes como espacio de comunicación para la educación virtual de estudiantes con discapacidad en Costa Rica en tiempos de pandemia. *Historia y Comunicación Social, 26*(Especial), 75-85. https://doi.org/10.5209/hics.74243

Demuner Flores, M. del R. (2021). Uso de redes sociales en microempresas ante efectos COVID-19. *Revista de Comunicación de la SEECI*, 54, 97-118. https://doi.org/10.15198/seeci.2021.54.e660

Hueso Romero, J. J. (2022). Creación de una red neuronal artificial para predecir el comportamiento de las plataformas MOOC sobre la agenda 2030 y los objetivos para el desarrollo sostenible. *Vivat Academia. Revista de Comunicación*, 155, 61-89. https://doi.org/10.15178/va.2022.155.e1386

Sancho Escrivá, J. V., Fanjul Peyró, C., De la Iglesia Vayá, M., Montell, Joaquín A., & Escartí Fabra, M. J. (2020). Aplicación de la inteligencia artificial con procesamiento del lenguaje natural para textos de investigación cualitativa en la relación médico-paciente con enfermedad mental mediante el uso de tecnologías móviles. *Revista de Comunicación y Salud, 10*(1), 19-41. http://doi.org/10.35669/rcys.2020.10(1).19-41