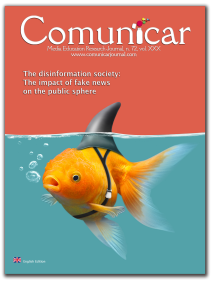
Comunicar, n. 72, v. XXX, 2022 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com



Interdisciplinarity of scientific production on hate speech and social media: A bibliometric analysis

Interdisciplinariedad de la producción científica sobre el discurso del odio y las redes sociales: Un análisis bibliométrico

Dr. Antonia Ramírez-García. Professor, Department of Education, University of Córdoba (Spain) (a.ramirez@uco.es) (https://orcid.org/0000-0001-7574-4854)
Dr. Antonio González-Molina. Coordinator, Scientific Information and Evaluation Unit, University of Córdoba (Spain) (agmolina@uco.es) (https://orcid.org/0000-0002-8223-813X)
Dr. María-del-Pilar Gutiérrez-Arenas Associate Professor, Department of Education, University of Córdoba (Spain) (pilar.gutierrez@uco.es) (https://orcid.org/0000-0002-8869-5550)
Dr. Manuel Moyano-Pacheco. Lecturer, Psychology Department, University of Córdoba (Spain) (mmoyano@uco.es) (https://orcid.org/0000-0001-6745-0936)

ABSTRACT

The impact of hate speech, both on a personal and social level, has increased due to social media. This has made it the focus of interest of numerous scientific journals, which increases the visibility of this global problem. The aim of this research is to analyse the basic descriptive metrics of the scientific production on hate speech and social media, as well as to explore the interdisciplinarity of these approaches. A bibliometric study has been carried out on the basis of the works indexed in the Scopus database related to the binomial 'hate speech' and 'social media' over a period of 20 years (2001 to 2020). The metrics used show that it is from 2017 onwards when this topic begins to arouse greater interest among researchers and that they constitute a sufficient indicator to consider the topic as one of interest to the scientific community. The joint research between both concepts raises its quality levels from a strictly metric point of view. 'Computer Science' and 'Social Sciences' are the two areas that clearly define the scientific production on this subject. The inversion of percentages in terms of the areas of origin of the works and citations in these two areas, is evidence of this interdisciplinarity. The indicators obtained show the relevance and transcendence of a social problem in the face of which proactive measures must be implemented.

RESUMEN

Las repercusiones que tiene el discurso del odio, tanto a nivel personal como social, se han intensificado con las redes sociales. Esto lo ha convertido en centro de interés de numerosas revistas científicas, lo que incrementa la visibilización de esta problemática global. El objetivo de esta investigación es analizar las métricas básicas descriptivas de la producción científica sobre el discurso del odio y redes sociales, así como explorar la interdisciplinariedad de estos enfoques. Se ha llevado a cabo un estudio bibliométrico a partir de trabajos indexados en la base de datos Scopus relacionados con el binomio «discurso de odio» y «redes sociales», en un período temporal de 20 años (2001 a 2020). Las métricas utilizadas demuestran que, a partir del año 2017, esta temática comienza a despertar mayor interés entre los investigadores, constituyéndose un indicador suficiente para considerar el tema como de interés por parte de la comunidad científica. La investigación conjunta entre ambos conceptos eleva sus niveles de calidad desde un punto de vista estrictamente métrico. Las áreas «Computer Science» y «Social Sciences» son las dos que definen claramente la producción científica sobre este tema. La inversión de porcentajes en cuanto a áreas de procedencia de los trabajos y citas en estas dos áreas evidencian esta interdisciplinariedad. Los indicadores obtenidos muestran la relevancia y trascendencia de un problema social ante el que se deben implementar medidas proactivas.

KEYWORDS | PALABRAS CLAVE

Hate speech, bibliometric analysis, social media, interdisciplinarity, scientific production, visibility.
Discurso del odio, análisis bibliométrico, redes sociales, interdisciplinariedad, producción científica, visibilización.

Received: 2021-12-27 | Reviewed: 2022-01-14 | Accepted: 2022-03-02 | OnlineFirst: 2022-05-15 | Published: 2022-07-01
DOI https://doi.org/10.3916/C72-2022-10 | Pages: 123-134

1. Introduction

Freedom of expression is the cornerstone of the system of rights and freedoms that identify democratic societies. This is applied in numerous different contexts, such as art, literature, religion, and politics, among others. However, as Ballesteros-Aguayo and Langa-Nuño (2018) point out, it is also a two-sided coin that, on the one hand, makes it possible to develop ideological, educational, or religious freedom and, on the other hand, is used with the intention of inflicting harm or undermining the dignity of the person. This is when hate speech arises, understood by the Council of Europe (1997) as those forms of expression that propagate, incite, promote, or justify rational hatred, xenophobia, anti-Semitism, and all other forms of hatred based on intolerance, including aggressive nationalism, ethnocentrism, discrimination, and hostility towards immigrant minorities.

According to Parekh (2006), hate speech has three defining elements: 1) an objectively offensive or degrading message; 2) targeting a specifically identified social group; and 3) risk of exclusion of that group. Along the same lines, Waldron (2012) expressed that hate speech manifests itself as: 1) accusing members of a specific collective of committing unlawful acts in a generalised manner; 2) comparing the collective group with another element that allows its dehumanisation; 3) denigration and offensive characterization of the collective; and 4) specific prohibition according to representative defining features of the collective.

For Gagliardone et al. (2015), the concept also includes expressions that directly encourage the commission of discriminatory acts or hate violence, and it has even been widely used in the media to refer to threats towards specific individuals in a more or less offensive way. Regarding these two concepts - freedom of expression and hate speech - Western societies hold different positions, especially in the United States (inclined towards not limiting freedom of expression) and European states which, although they express different conceptions regarding freedom of expression and its limits, according to Gascón (2019: 64), they consider that "hate speech is inadmissible in a democratic society that protects human rights and fights against discrimination".

This fact has led the European Union to establish legislative measures with the intention of regulating these types of messages, given the difficulty of distinguishing them from other manifestations. These include the European Convention for the Protection of Human Rights and Fundamental Freedoms (Ministry of Foreign Affairs, 1999), the Recommendation of the Committee of Ministers of the Council of Europe (1997) no. R 20 and General Recommendation no. 15 on Lines of Action to combat hate speech (Ministry of Foreign Affairs and Cooperation, 2016). Likewise, a series of parameters has been defined, included in the so-called Strasbourg Test, which allow the delimitation of hate speech (subject matter of the message, sender of the message, intention of the sender, target group of the speech, geographical area of dissemination of the message and the channel used to disseminate the message).

Hatred is a drive or emotion that has accompanied humanity throughout time. Its danger lies, according to Garton (2017), in that it can be constructed, encouraged, inculcated, propagated and, ultimately, applied. In our opinion, in today's post-modern society, there is a context prone to the dissemination of this type of emotion and, therefore, of its corresponding discourse. An environment mediated by technology and digitalisation has thus emerged in which there are millions of prosumers of emotions and feelings willing to visualise, create and share them through social media.

In this regard, in 2016 the European Union signed a Code of Conduct to combat online hate speech with the technology companies responsible for social media such as Facebook, Microsoft, Twitter and YouTube, extending in 2018 to Instagram, Google+, Snapchat and Dailymotion. The aim of this Code is for these intermediaries and online communication platforms to act immediately in cases of online hate speech and make a series of public commitments to: 1) establish clear and effective procedures that would prohibit such speech; 2) generate a procedure to remove such speech in less than 24 hours; 3) educate and raise awareness among users; 4) provide information on reporting procedures when communicating with authorities; 5) increase collaboration among themselves, with other intermediaries to achieve the best practices, as well as with civil society; and 6) develop and promote alternative speech. Ultimately, this Code seeks to prevent the spread of hate speech (European Commission, 2020).

Despite the signing of this Convention, a number of issues need to be highlighted. Firstly, social media is not subject to the professional ethics that have regulated traditional social networks. Secondly, these

networks are intermediaries in digital communication, so they can decide what is or is not published under their own publication policies. Thirdly, they play a dual role, since, as Ben and Matamoros (2016) state, on the one hand, they officially prohibit explicit manifestations of hate and, on the other hand, they offer their infrastructure for the proliferation of associations and collectives that can incite hatred.

The European Union's concern about the presence of hate speech on social media and the establishment of mechanisms to regulate it has led to the emergence of various European projects. Among others, the "Preventing, redressing, inhibiting hate speech in new media" (BRaVE, 2019), documents such as the Raxen reports (Info Raxen, n.d.) that warn about the growth of hate speech on the Internet and social media as well as research on Facebook as a network that favours discrimination among its users (Gillespie, 2010) and the proliferation of negative feelings in the comments of this social network (Jaramillo et al., 2015) or Twitter and the instantaneous expression of emotions and moods (Burnap & Williams, 2015), as well as the treatment of immigration on this network (Merino-Arribas & López-Meri, 2018). Likewise, there has been a growing interest in this topic in the academic sphere. Wright et al. (2021: 22) state that "it is a central and highly relevant scientific and social issue", which has even generated its own concept, 'cyberhate'.

For Chakraborti et al. (2014), cyberhate is any digital act of violence, hostility and intimidation towards people motivated by their identity or difference. In this sense, Wachs and Wright (2019) specify that this expression of hatred against 'the others' is produced through offensive texts, speeches, videos, or images. In our opinion, the relevance of Wright et al. (2021) for this theme could be motivated by several factors. Firstly, due to the interest shown by the scientific community in social media, since, immediately after their emergence, studies on the matter are published. As can be seen in Table 1, not even two years pass between the appearance of a certain social network and a publication corresponding to it.

Table 1. Social Media and First Publications

Social Media	Origin	Millions of users (2021)	First Publications	
			JCR	Scopus
Facebook	2004	2,740	Lashinsky (2005)	Lashinsky (2005); Gross et al. (2005)
YouTube	2005	2,291	Woolley (2006a; 2006b)	Australian Computer Society (2005)
Twitter	2006	353	McFedries (2007)	McFedries (2007), Greene (2007), Green (2007)
WhatsApp	2009	2,000	Kim (2011)	Abanmy et al. (2012); Nikou et al. (2012a); Nikou et al. (2012b)
Instagram	2010	1,221	Kvalnes (2010); Barnes et al. (2010)	Barnes et al. (2010)
TikTok	2016	689	Liu et al. (2018)	Shafer (2016)

Secondly, the number of network users. Data provided by Galeano (2021) show that more than half of the world's population uses social media (53.6%), or 4.2 billion people, with a year-to-year increase of 13.2% over the previous year, probably as a result of the pandemic. Table 1 shows the number of users of the most widely used social media. Therefore, an added increase in the average time spent using social media (2 hours and 25 minutes) must be added. Social media therefore brings together millions of prosumer users in real time who can respond spontaneously, instantaneously, and impulsively, under cover of anonymity, to messages, images and/or videos impregnated with hate.

Thirdly, the characteristics of social media itself, which not only constitute a new dissemination channel (Losada-Díaz et al., 2021), but also create new scenarios and forms of development, including 'Flaming' (strong, 'inflammatory' opinions using offensive language) and 'Trolling' (Khosravinik & Esposito, 2018). Trolling includes a list of actions such as in-game insults, tasteless and dangerous jokes, threats, rape, and murder in which absurd and inflammatory comments are used, the aim of which is to provoke an equally aggressive reaction and enjoy the conflict that is generated (Hardaker, 2013). Added to this is the proliferation of 'haters', who are people who engage in obsessive verbal attacks and aggression.

Finally, the repercussions that hate speech can have, including direct emotional or psychological damage to the person and/or group, as well as indirect consequences such as the perpetuation of discriminatory stereotypes, dehumanisation of groups, marginalisation, reduction of empathy, silencing effect on victims and, according to Marabel (2021), even the proliferation of hate crimes, risk to public order, and the modelling of totalitarian societies. Hate speech, then, has become the focus of interest of many institutions, and scientific journals are no strangers to this. As Martínez-Nicolás and Saperas

(2011) state, these are configured as the main channel for the dissemination of scientific production. These journals act as trend-setting agents through the monographs they propose, the articles they select for publication, and the reviews they include in their publications, among other aspects. If scientific journals are also well positioned in quality rankings (Journal Citation Reports, Scimago Journal Rank), their influence is much greater. Therefore, the leadership they have among the scientific community would make it possible to increase the visibility of this global problem and contribute to the social responsibility to which they are also called.

In this context, different authors (Carneiro-Barrera et al., 2019; Cabrera, 2020) advocate the exploration of the publications that have been made on a particular topic over a given time. In this way, it is possible to find out who has made contributions to the topic, what collaborative structures have been configured, or in what context it has been produced. It is therefore necessary to resort to bibliometric studies, considered as a branch of scientometrics (Marín-Aranguren & Trejos-Mateu, 2019). These studies are highly regarded for their contributions to the quantification of written communication processes (Mingers & Leydesdorff, 2015) through the application of statistical and mathematical methods (Rehn & Kronman, 2008), which make it possible to describe the internal and external properties of a body of scientific knowledge (Estabrooks et al., 2004).

In the same way, the major providers of scientific information databases (Clarivate Analytics and Scopus) include among their analysis tools (InCite and Scival, respectively) bibliometric indicators endorsed by the scientific community as useful metrics to describe, among other issues, the characteristics of scientific production. In this scenario, and as a concept that has been well studied over the last few years, we find the interdisciplinarity of science, which allows us to carry out analyses of different objects such as large scientific fields (Chen et al., 2014; Khosrowjerdi & Bayat, 2013; Porter & Rafols, 2009), academic collaboration (Repiso-Caballero et al., 2016), journals (Leydesdorff & Rafols, 2011), comparison of perspectives (Avila-Robinson et al., 2021), and purposes (Rinia et al., 2002), which aim to find solutions to complex social problems, such as hate speech. The response to this phenomenon cannot be approached from a single scientific field, nor from an exclusive methodological proposal; it requires a multifaceted study that provides specific evidence of this social reality.

Thus, Tontodimamma et al. (2021) analysed the topics of interest on hate speech between 1992 and 2018, highlighting the influence exerted by social media, and Mishra (2021) focused her descriptive study on the type of publications, research areas, countries, affiliation, and keywords on hate speech between 1962 and 2021, but without linking it to social media. Therefore, this paper complements and updates previous studies, shows the basic descriptive metrics of the scientific production on hate speech and social media, and explores the interdisciplinarity of the approaches, based on the study of the classification of production by thematic areas, similar to the methodology by scientific categories (Montero-Díaz et al., 2018) and keyword analysis (Leydesdorff & Nerghes, 2017; Vargas-Quesada et al., 2017), both of the output and of the citing papers.

2. Material and methods

Although the study presented here does not correspond to a typical systematic review, as it is scientometric research, characterised by the analysis of scientific literature, it is advisable to ensure a rigorous methodological process that facilitates understanding by readers who are not familiar with this type of work. For this reason, the methodology proposed by PRISMA (2020) has been adapted for this article (Figure 1).

The two sources traditionally used for bibliometric studies are Web of Science (WoS, from Clarivate Analytics) and Scopus (Elsevier). Although both databases can cover the information needs for the present study, Scopus has been chosen because of the greater coverage at the level of journals analysed and the total citation volume (Singh et al., 2021; Martín-Martín et al., 2021). A simple search was carried out on the term 'hate speech' to retrieve the total number of documents analysed. Regarding the document typology, all the types coded in the database were considered, taking into account the possible disciplines involved in the study of the subject of hate speech, and the different publications as well as citation patterns of the researchers according to their study areas.

At a formal level, the very clear definition of the concept 'hate speech' has made the retrieval of documents entirely satisfactory. In the same way, the clear identification of each of the platforms or social media and the concepts directly related to 'social media' (social network, social media) has allowed us to establish the search equations shown in Figure 1 (search strategy).

Figure 1. Methodological indicators of the study

Indicator	Description	Justification
Eligibility criteria	1. Timeframe: 20 years (2001-2020) 2. Document typology: all coded in the database.	1. Home search: 3 years before the emergence of 1st social network 2a. Respond to the possible disciplines involved 2b. Different publication and citation patterns of researchers by area of knowledge
Information sources	Scopus	Increased journal coverage Increased total citation volume
Search strategy	**B1 ("hate speech")**: TITLE-ABS-KEY ("hate speech") and PUBYEAR > 2000 AND PUBYEAR < 2021 **B2 ("hate speech" + "social networks")**: TITLE-ABS-KEY ("hate speech") AND (TITLE-ABS-KEY ("social network") OR TITLE-ABS-KEY ("social media") OR TITLE-ABS-KEY ("social platform") OR TITLE-ABS-KEY ("facebook") OR TITLE-ABS-KEY ("twitter") OR TITLE-ABS-KEY ("tiktok") OR TITLE-ABS-KEY ("whatsapp") OR TITLE-ABS-KEY ("youtube") OR TITLE-ABS-KEY ("instagram") OR TITLE-ABS-KEY ("snapchat") AND PUBYEAR > 2001 AND PUBYEAR < 2020	B1. General search for the term that is configured as the focus of interest. B2. Inclusion of the social networks with the largest number of users (Galeano, 2021).
Study selection process	1. Consultation date: 15/09/2021 2. Search field: "Title", "Abstract" y "KeyWords" 3. Selection: existence or non-existence of the indexed work in the database	Textual content fields on the article commonly used in bibliographic information retrieval.
Data extraction process	1. Export: csv format for processing in Microsoft Access database and Scival. 2. Post-processing in Microsoft Excel 3. Use of VOSViewer Software 1.6.17 4. Download of the list of journals included in Scopus and integration in the *ad hoc* designed system.	1. Analysis and generation of indicators 2. Dumping of production and category data to generate graphs 3. Generation of the keyword map or network 4. Categorisation of the retrieved papers: 27 major scientific areas and 334 thematic categories.
Data list	Structuring of the data: N papers, N citations, % papers cited at least once, citation ratio per paper, % papers published in first quartile and first decile journals according to SJR[1] indicator, % papers ranked in the top decile by global citation, normalised impact by category with respect to the world [2] and % works carried out in internacional[3] collaboration. $N_{(B1)}$=1713; $N_{(B2)}$=639	1. Gutiérrez-Salcedo et al. (2018); Guerrero-Bote & Moya-Anegón, (2012) 2. Field-Weighted Citation Impact–FWCI 3. Snowball Metrics (n.d.)

The selection of platforms or social networks considered for the study is based on the user data provided by Galeano (2021), and the final choice has depended on the existence or not of any work specifically indexed in the database in the period of analysis considered. The data exported from Scopus were citation information, bibliographic information, abstract, keywords, and other information. Finally, for the categorisation of the retrieved papers, it was necessary to download the list of journals included in the Scopus database, which was also integrated into the ad-hoc system designed.

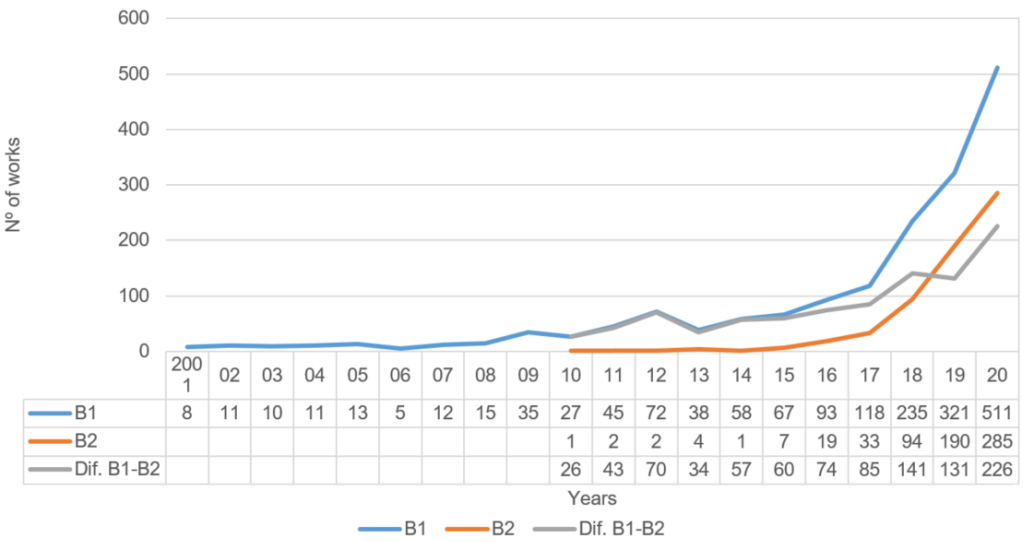
3. Results

The execution of query B1, the most inclusive query, located all papers that included the term 'hate speech' in any of the established search fields. A total of 1,713 papers were retrieved, regardless of whether the terms related to 'social media' appeared. Query B2, specific to the observation under study,

retrieved a total of 639 papers. Due to the connection procedures between the Scopus database and the Scival analytical tool, there is an error inherent to the synchronisation of these tools that affected the total count, with a final output retrieved for query B1 of 1,705 papers and for B2, 638 papers, which will be the final sample under study. This same problem is transferred to the set of jobs resulting from the Boolean difference of B1-B2 (B1 not B2).

Figure 2 shows the evolution of production over time. The first publication in which the concepts 'hate speech' and some of those related to 'social media' appear together is in 2010, specifically with the term 'social media'. It was not until 2011 that this association appeared with the 'Facebook' platform. As can be seen in Figure 2, the research where the concepts 'hate speech' and 'social media' are integrated occurs in 2019, although it is in 2017 when the trend changes and research on the topic studied arouses greater interest among researchers.

Figure 2. Evolution of scientific production on datasets B1, B2 and their differences



	2001	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
B1	8	11	10	11	13	5	12	15	35	27	45	72	38	58	67	93	118	235	321	511
B2										1	2	2	4	1	7	19	33	94	190	285
Dif. B1-B2										26	43	70	34	57	60	74	85	141	131	226

With regard to the documentary typology of the information analysed (set B2), 50.23% of the works belong to the Conference paper type, 39.28% to the Article type and in lower percentages Book Chapter with 3.43%, Conference review 3.13% and the rest, Conference review, Review, Book and Note in percentages of 3.13%, 2.5%, 0.78%, and 0.47% respectively.

Table 2 shows the metrics relating to the 2010-2020 output, a range in which there are papers already published in the B2 dataset and a comparison of each of the indicators can be made. Column B1-B2 includes the metrics of the papers not included in B2 that are in B1, i.e., the papers where the term 'hate speech' appears but none of the terms established to recover the papers related to 'social media' appear. As shown, the relative metrics, both quantitative (volume of papers) and qualitative (related to citation) of the B2 dataset, have higher values with respect to both the B1 dataset and the difference. In this sense, the contribution of the joint research on hate speech and social media shows an increase in its quality levels from a strictly metric point of view.

On the other hand, the values for the percentages of cited papers, international collaboration and the FWCI normalised impact are worth highlighting. 67.1% of the research papers related to hate speech and social media are cited by third party researchers at least once. This is corroborated by the international collaboration indicator of the same dataset, B2. The FWCI, as an indicator that relates citation to the volume of papers considering the publication and citation behaviour of the different areas, is a parameter that describes the status of research in relation to the world. The reference value for this indicator is 1, for

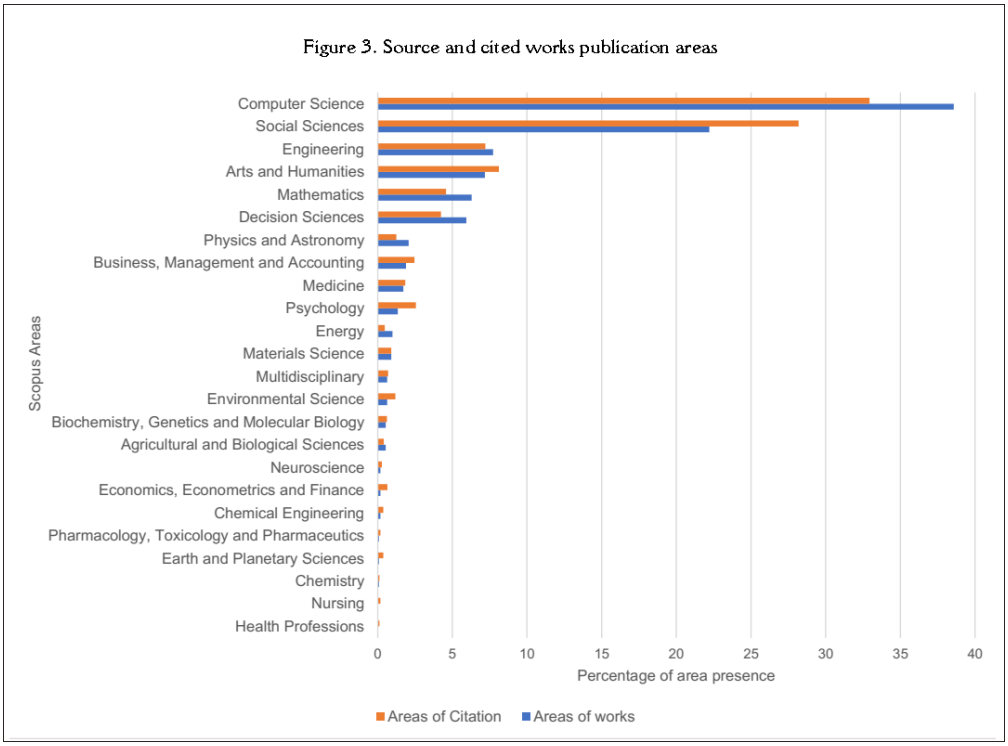
the area of Computer Science it is 1.05 and for the area of Social Science it is 1.23. If we compare these reference values with those obtained in this study, we can say that the scientific production related to hate speech and social media together, is cited 173% more than the world average, a value well above the 74% relating to the works that include the term 'hate speech' without any relation to the search terms related to social media. As for the percentage of papers published in the first quartile journals, although it is true that there is a more moderate increase in the B2 dataset, if the indicator for the first decile is considered, it can be affirmed that these papers still constitute excellent science. The same aspect is reinforced by the value, 15%, of the indicator for papers in the top 10% (first decile) of the world's most cited papers, compared to 8.9% for the B1 dataset.

Table 2. Bibliometric indicators

Indicator	B1	B2	B1-B2
Nº of works	1585	638	948
Nº of citations	11,514	5,949	5,565
Percentage of works cited	66.1	67.1	65.3
Ratio of citations per work	7.3	9.3	5.9
Percentage of works in Q1 by SJR	40.4	40.7	40.2
Percentage in the first decile according to SJR	16.9	19	15.8
Percentage of cited papers in the top global citation decile	8.9	15	4.7
FWCI (Field-weighted Citation Impact)	2.14	2.73	1.74
Percentage of works carried out in international collaboration	13.3	21.4	7.8
Percentage of works carried out in national collaboration	12.6	16.9	9.7
Nº of areas	24	22	22
Nº of categories	142	120	111

By means of various operations with the database defined 'ad hoc', with the information from the B2 set, the categorisation of the papers was carried out based on the cross-referenced information with the list of Scopus journals.

Figure 3. Source and cited works publication areas



The result was a 65% match, which is too low considering the total volume of papers retrieved. This aspect must also be analysed from the point of view of the majority of the documentary typology

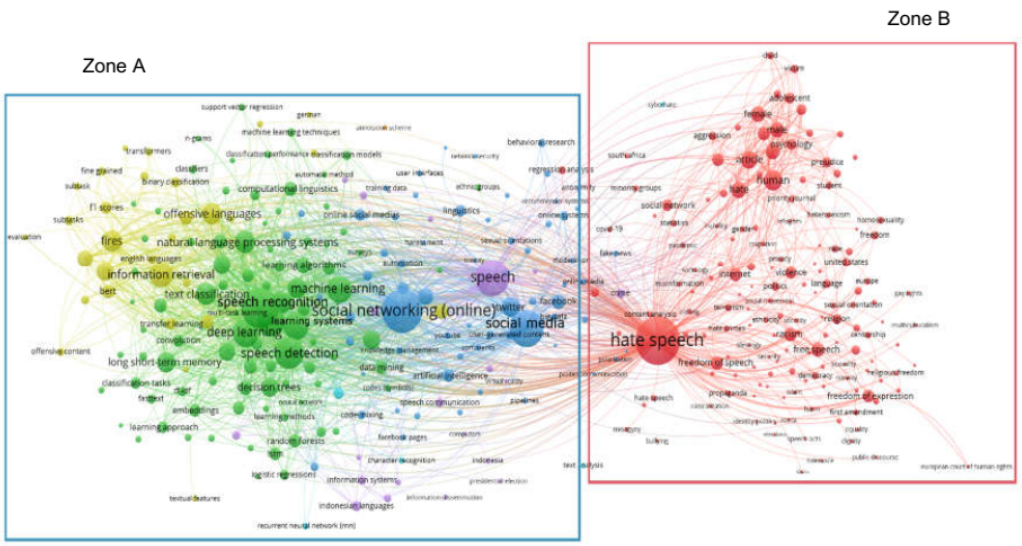
(conference paper), which produces a certain lack of solidity due to the very nature of the information in databases of this type. This fact motivates the use of the area classification system for the analysis of interdisciplinary approaches to hate speech research.

Figure 3 shows the percentages of the scientific production analysed ascribed to the Scopus subject areas of both the B2 papers, 'source' papers in this case, and the citing papers as a whole. Graphically, it can be seen how the first two classification areas, Computer Science and Social Sciences, clearly define the scientific production analysed, although there are papers in practically all areas. It should also be noted that in these two areas, the percentages of the area of work and citation are inverted, demonstrating the need for interdisciplinarity in the approach to hate speech.

If the previous classification offered a macro approach (scientific areas) to the possible approaches used when studying the concept of hate speech and social media, an analysis from the point of view of methodologies such as keyword co-occurrence analysis (Leydesdorff & Nerghes, 2017; Wang et al., 2012) shows at a micro level (keywords) the existing relationships between the works.

Figure 4 shows a graph made from the keywords of the B1 works. It has been generated under the default parameters of the software used, VOSViewer, taking into account a minimum occurrence of terms of 5. Two well-defined zones, A and B, with 6 and 1 clusters each, are clearly visible. Zone B, which includes the red clusters, represents approaches to research on hate speech and social media from a social science point of view. Zone A represents works with computer science approaches, including, in this case, aspects of computational methodologies, machine learning, text mining, offensive language detection, algorithmic, etc. The positioning of the central node being hate speech, supports the network due to the search methodology used. However, it is important to consider the relationships, although weak, of certain peripheral nodes that establish connections between the two approaches to the research carried out.

Figure 4. Co-word map



4. Discussion and conclusions

The results offered show the exponential increase in scientific interest in the binomial of hate speech and social media, concurring with the interest and social relevance that this phenomenon has recently acquired in society. From a strictly metric point of view, the initial findings show the best scenario defined by the indicators for hate speech research when linked to social media (B2 dataset) in recent years. The large increase in research output related to hate speech and social media is a sufficient indicator to consider the topic of interest for the scientific community. This fact is also motivated by the unstoppable

development of information and communication technologies. The scientometric indicators show a certain imbalance between the datasets analysed. This imbalance is clearly caused by the increased values in the indicators related to global research on hate speech and social media as linked concepts. Thematic contextualisation makes it possible to see in the same way the interest that the research community has in this, even in works that constitute the science of excellence, i.e. the highly cited (Bornmann, 2014).

In the current system of science, collaborations between researchers are essential because, on the one hand, it has been proven that scientific collaboration favours visibility in terms of citation (Guerrero-Bote et al., 2013) and, on the other hand, because of the necessary interdisciplinarity of science, especially in a subject of such importance as hate speech. Regardless of theoretical considerations and the studies that the literature provides to measure the interdisciplinarity of science (Ávila-Robinson et al., 2021), it is a fact that, as has been shown in this research, there is an approach to the subject of analysis from practically all the thematic areas established by Scopus. The classification of journals according to broad areas of knowledge allows the analysis of scientific production in order to carry out analyses of large domains, as has been done here. The division into lower units of these areas (categories) also provides one of the pillars traditionally used for the analysis of these scientific domains (Bornmann et al., 2011).

For the purposes of this study and given its intention to approximate the interdisciplinary representation of hate speech research, it is not considered necessary to include the graph metrics analysis. However, it would be useful to further explore the relationship between interdisciplinarity and increased scientific impact. On the other hand, the clear definition of 7 well-defined clusters and the grouping into two well-configured zones visually shows the two main approaches to hate speech research. Although the works in the area of Computer Science are higher than those in Social Sciences, the inversion of percentages in terms of the areas of origin of the works and citations in these two predominant areas shows the need to resort to other areas of knowledge in order to understand a social problem of the magnitude of hate speech.

In this sense, a critical analysis such as the one conducted by Viseu (2015) could be necessary for a reconfiguration of the concept of the research team in the field of social sciences through the integration of experts in computer science, jurists, and psychologists, among others. Hate speech in cyberspace represents the tip of the iceberg of a broader structural problem, its normalisation being a breeding ground for incidents of inter-group conflict, polarisation of social groups, dehumanisation of certain groups and processes of violent radicalisation of individuals and groups. From an applied point of view, the indicators obtained could be considered a proxy for the relevance and transcendence of a social problem in the face of which proactive measures must be implemented. For all these reasons, it is necessary to continue to make progress in the adoption of comprehensive and preventive measures in the face of a challenge in which technology, communication, and education converge, as in few others. As possible new lines of research to complement this study, it would be interesting to carry out a content analysis of hate speech in the sources analysed, as well as the possibility of carrying out a comparison between the WOS/Scopus databases.

Authors' Contribution

Idea, A.R.G.; Literature review (state of the art), A.R.G.; Methodology, A.R.G., A.G.M., M.P.G.A.; Data analysis, A.G.M.; Results, A.G.M.; Discussion and conclusions, A.R.G., A.G.M., M.P.G.A., M.M.P.; Drafting (original draft), A.R.G., A.G.M., M.P.G.A.; Final revisions, A.R.G., A.G.M., M.P.G.A., M.M.P.; Project design and sponsorship, A.R.G., A.G.M., M.P.G.A., M.M.P.

Funding Agency

This study is part of the "Sport for Prevention of Extremism in Youth (SPEY)" project, co-funded by the European Commission and led by the Unió de Federacions Esportives de Catalunya (SANT-609085F8).

References

Abanmy, N.O., Al-Quait, N.A., Alami, A.H., Al-Juhani, M.H., & Al-Aqeel, S. (2012). The utilization of Arabic online drug information among adults in Saudi Arabia. *Saudi Pharmaceutical Journal*, *20*(4), 317-321. https://doi.org/10.1016/j.jsps.2012.07.001

Australian Computer Society (Ed.) (2005). *Conferences in research and practice in information technology*. Seventh Australasian Computing Education Conference. https://bit.ly/3kyB0iA

Ávila Robinson, A., Mejia, C., & Sengoku, S. (2021). Are bibliometric measures consistent with scientists' perceptions? The case of interdisciplinarity in research. *Scientometrics*, *126*(9), 7477-7502. https://doi.org/10.1007/S11192-021-04048-0

Ballesteros-Aguayo, L., & Langa-Nuño, C. (2018). *Movimientos populistas en Europa: La actualización del discurso totalitario en los medios de comunicación actuales y su repercusión en la opinión pública*. Egregius. https://bit.ly/3tqgA0R

Barnes, R.I., Earl, G.F., Papazoglou, M., Burchett, L., & Terzuoli, A.J. (2010). The Instagram: A novel sounding technique for enhanced HF propagation advice. In *IEEE National Radar Conference* (pp. 1446-1449). https://doi.org/10.1109/RADAR.2010.5494387

Ben, A., & Matamoros, A. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, *10*, 1167-1193. https://bit.ly/3Bkevok

Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Research Evaluation*, *23*(2), 166-173. https://doi.org/10.1093/reseval/rvu002

Bornmann, L., Leydesdorff, L., Walch-Solimena, C., & Ettl, C. (2011). Mapping excellence in the geography of science: An approach based on Scopus data. *Journal of Informetrics*, *5*(4), 537-546. https://doi.org/10.1016/j.joi.2011.05.005

BRaVE (Ed.) (2019). *Preventing redressing and inhibiting hate speech in new media*. https://bit.ly/3euQUrP

Burnap, P., & Williams, M.L. (2015). Cyber hate speech on Twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & Internet*, *7*(2), 223-242. https://doi.org/10.1002/poi3.85

Cabrera, J.F. (2020). Producción científica sobre integración de TIC a la Educación Física. Estudio bibliométrico en el periodo 1995-2017. *Retos*, *37*, 748-754. https://doi.org/10.47197/retos.v37i37.67348

Carneiro-Barrera, A., Ruiz-Herrera, N., & Díaz-Román, A. (2019). Tesis doctorales en Psicología tras la adaptación al Espacio Europeo de Educación Superior. *Revista de Investigación en Educación*, *17*(1), 32-43. https://bit.ly/2ZuI45t

Chakraborti, N., Garland, J., & Hardy, S.J. (2014). *The Leicester hate crime project: Findings and conclusions*. University of Leicester. https://bit.ly/38lTPzQ

Chen, S., Arsenault, C., Gingras, Y., & Larivière, V. (2014). Exploring the interdisciplinary evolution of a discipline: the case of Biochemistry and Molecular Biology. *Scientometrics*, *102*(2), 1307-1323. https://doi.org/10.1007/S11192-014-1457-6

Comisión Europea (Ed.) (2020). *El código de conducta de la UE para la lucha contra la incitación ilegal al odio en Internet (comunicado de prensa IP/20/1134)*. https://bit.ly/32lrxX3

Consejo de Europa (Ed.) (1997). *Recomendación del Comité de Ministros del nº R (97) 20*. https://bit.ly/3mBZASv

Estabrooks, C.A., Winther, C., & Derksen, L. (2004). Mapping the field: A bibliometric analysis of the research utilization literature in nursing. *Nursing Research*, *53*(5), 293-303. https://doi.org/10.1097/00006199-200409000-00003

Gagliardone, I., Gal, D., Alves, T., & Martínez, G. (2015). *Countering online hate speech. Programme in comparative media law and policy*. University of Oxford. https://bit.ly/2XYfE6F

Galeano, S. (2021, January 28). *Cuales son las redes sociales con más usuarios del mundo*. Marketing Ecommerce. https://bit.ly/3DmqAex

Garton, T. (2017). *Libertad de palabra. Diez principios para un mundo interconectado*. Tusquets.

Gascón, A. (2019). La lucha contra el discurso del odio en línea en la Unión Europea y los intermediarios de Internet. In Z. Combalía, M. P. Diago, & A. González-Varas (Eds.), *Libertad de expresión y discurso de odio por motivos religiosos* (pp. 64-86). Ediciones del Licregdi. https://bit.ly/3wkba9G

Gillespie, T. (2010). The politics of platforms. *New Media & Society*, *12*(3), 347-364. https://doi.org/10.1177/1461444809342738

Green, H. (2007). Twitter. All trivia, all the time. *Business Week*, *4028*, 40-40.

Greene, K. (2007). Whats is he doing? *MIT Technology Review*, *110*(6), 44-51. https://bit.ly/3sLu4Em

Gross, R., Acquisti, A., & Heinz-Iii, H.J. (2005). Information revelation and privacy in online social networks. *WPES'05: Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, (pp. 71-80). https://doi.org/10.1145/1102199.1102214

Guerrero-Bote, V.P., & Moya-Anegón, F. (2012). A further step forward in measuring journals' scientific prestige: The SJR2 indicator. *Journal of Informetrics*, *6*(4), 674-688. https://doi.org/10.1016/j.joi.2012.07.001

Guerrero-Bote, V.P., Olmeda-Gómez, C., & De-Moya-Anegón, F. (2013). Quantifying the benefits of international scientific collaboration. *Journal of the American Society for Information Science and Technology*, *64*(2), 392-404. https://doi.org/10.1002/asi.22754

Gutiérrez-Salcedo, M., Martínez, M.A., Moral-Munoz, J.A., Herrera-Viedma, E., & Cobo, M.J. (2018). Some bibliometric procedures for analyzing and evaluating research fields. *Appl Intell*, *48*, 1275-1287. https://doi.org/10.1007/s10489-017-1105-y

Hardaker, C. (2013, August 3). *What is turning so many young men into trolls?* The Guardian. https://bit.ly/2Wod9dF

Info Raxen (Ed.) (n.d.). *Servicio de noticias de movimiento contra la intolerancia*. https://bit.ly/3kua2Zp

Jaramillo, S., Cardona, S.A., & Fernández, A. (2015). Minería de datos sobre streams de redes sociales, una herramienta al servicio de la Bibliotecología. *Información, Cultura y Sociedad*, *33*, 63-74. https://doi.org/10.34096/ics.i33.1182

Khosravinik, M., & Esposito, E. (2018). Online hate, digital discourse and critique: Exploring digitally-mediated discursive practices of gender-based hostility. *Lodz Paper in Pragmatics*, *14*(1), 45-68. https://doi.org/10.1515/lpp-2018-0003

Khosrowjerdi, M., & Bayat, M.K. (2013). Mapping the Interdisciplinarity in Scientometric Studies. *Iranian Journal of Information Processing and Management*, *28*(2), 393-409. https://bit.ly/3CAUiwp

Kim, J.H. (2011). A study on the graphic user interface for emotional communications in the mobile messenger -focused on 'Kakao Talk' and 'WhatsApp' by case analysis. *Korea Open Access Journal*, *18*, 142-155. https://doi.org/10.17246/jkdk.2011.18.015

Kvalnes, O. (2010). *Moral reasoning at work: Rethinking ethics in organizations*. Palgrave pivot. https://doi.org/10.1057/9781137532619.0001

Lashinsky, A. (2005). Facebook stares down success. *Fortune*, *152*(11). https://bit.ly/3EXzv72
Leydesdorff, L., & Nerghes, A. (2017). Co-word maps and topic modeling: A Comparison using small and medium-sized corpora (N< 1,000). *Journal of the Association for Information Science and Technology*, *68*(4), 1024-1035. https://doi.org/10.1002/asi.23740
Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*, *5*(1), 87-100. https://doi.org/10.1016/j.joi.2010.09.002
Liu, W., Liu, W., Li, M., Chen, P., Yang, L., Xiao, C., & Ye, Y. (2018). Fine-grained task-level parallel and low power h.264 decoding in multi-core systems. In *24th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*. https://doi.org/10.1109/PADSW.2018.8644865
Losada-Díaz, J.C., Zamora-Medina, R., & Martínez-Martínez, H. (2021). El discurso del odio en Instagram durante las Elecciones Generales 2019 en España. *Revista Mediterránea de Comunicación*, *12*(2), 195-208. https://doi.org/10.14198/MEDCOM.19142
Marabel, J.J. (2021). Delitos de odio y redes sociales: El derecho frente al reto de las nuevas tecnologías. *Revista de Derecho UNED*, *27*, 137-172. https://doi.org/10.5944/rduned.27.2021.31076
Marín-Aranguren, E.M., & Trejos-Mateu, F.D. (2019). Sociedad civil en red y gobernanza de la Agenda 2030. *Forum*, *15*, 91-117. https://doi.org/10.15446/frdcp.n15.74544
Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado-López, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *Scientometrics*, *126*, 871-906. https://doi.org/10.1007/s11192-020-03690-4
Martínez-Nicolás, M., & Saperas, E. (1998). Análisis de los artículos publicados en revistas científicas. *Revista Latina de Comunicación Social*, *66*, 101-129. https://doi.org/10.4185/RLCS-66-2011-926-101-129
Mcfedries, P. (2007). Technical Speaking: All A-Twitter. *IEEE Spectrum*, *44*(10). https://doi.org/10.1109/MSPEC.2007.4337670
Merino-Arribas, A., & López-Meri, A. (2018). La lucha contra el inmigracionalismo y el discurso del odio en el metamedio social Twitter. In J. L. González-Esteban, & J. A. García-Avilés (Eds.), *Mediamorfosis. Radiografía de la innovación en el periodismo* (pp. 211-224). https://bit.ly/38jehBv
Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, *246*(1), 1-19. https://doi.org/10.1016/j.ejor.2015.04.002
Ministerio de Asuntos Exteriores (Ed.) (1999). *Resolución de 5 de abril de 1999, de la Secretaría General Técnica, por la que se hacen públicos los textos refundidos del Convenio para la protección de los derechos y de las libertades fundamentales. 6 de mayo de 1999, BOE nº 108*. https://bit.ly/2Wv4OVq
Ministerio de Asuntos Exteriores y de Cooperación (Ed.) (2016). *Recomendación General nº 15 sobre Líneas de Actuación en relación con la lucha contra las expersiones de incitación al odio*. https://bit.ly/3mAEdRz
Mishra, R. (2021). Are we doing enough? A Bibliometric analysis of hate speech research in the selected database of Scopus. *Library Philosophy and Practice*, *5140*. https://bit.ly/3sxcZyj
Montero-Díaz, J., Cobo, M., Gutiérrez-Salcedo, M., Segado-Boj, F., & Herrera-Viedma, E. (1980). A science mapping analysis of 'Communication' WoS subject category (1980-2013). [Mapeo científico de la Categoría «Comunicación» en WoS (1980-2013)]. *Comunicar*, *55*, 81-91. https://doi.org/10.3916/C55-2018-08
Nikou, S., Bouwman, H., & De-Reuver, M. (2012a). The potential of converged mobile telecommunication services: A conjoint analysis. *Info*, *14*(5), 21-35. https://doi.org/10.1108/14636691211256287
Nikou, S., Bouwman, H., & De-Reuver, M. (2012b). Mobile converged rich communication services: A conjoint analysis. *Proceedings of the Annual Hawaii International Conference on System Sciences*, (pp. 1353-1362). https://doi.org/10.1109/HICSS.2012.434
Parekh, B. (2006). Hate speech. Is there a case of banning? *Public Policy Research*, *12*(4), 213-223. https://doi.org/10.1111/j.1070-3535.2005.00405.x
Porter, A.L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, *81*(3), 719-745. https://doi.org/10.1007/S11192-008-2197-2
PRISMA (Ed.) (2020). *Declaración PRISMA*. https://bit.ly/33JwE3x
Rehn, C., & Kronman, U. (2008). *Bibliometric handbook for Karolinska Institutet*. Karolinska Institutet. https://bit.ly/3qmpbQm
Repiso-Caballero, R., Torres-Salinas, D., & López-Cozar, E.D. (2016). Análisis de la relación entre disciplinas a través del uso de tesis doctorales. *Revista Latina de Comunicación Social*, *71*, 874-890. https://doi.org/10.4185/RLCS-2016-1125
Rinia, E.J., Van Leeuwen, T.N., & Van-Raan, A.F.J. (2002). Impact measures of interdisciplinary research in physics. *Budapest Scientometrics*, *53*(2), 241-248. https://doi.org/10.1023/A:1014856625623
Shafer, L. (2016). TiK ToK on the Clock, but the party don't stop, no: The parodic military dance video on YouTube. In D. A. Cunningham, & J. C. Nelson (Eds.), *A companion to the war film* (pp. 320-337). https://doi.org/10.1002/9781118337653.ch19
Singh, V.K., Singh, P., Karmakar, M., Leta, J., & Mayer, P. (2021). The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometric*, *126*, 5113-5142. https://doi.org/10.1007/s11192-021-03948-5
Snowball Metrics (Ed.) (n.d.). *Standardized research metrics - by the sector for the sector*. https://bit.ly/3J6cLmS
Tontodimamma, A., Nissi, E., Sarra, A., & Fontanella, L. (2021). Thirty years of research into hate speech: Topics of interest and their evolution. *Scientometrics*, *126*(1), 157-179. https://doi.org/10.1007/s11192-020-03737-6
Vargas-Quesada, B., Chinchilla-Rodríguez, Z., & Rodríguez, N. (2017). Identification and Visualization of the Intellectual Structure in Graphene Research. *Frontiers in Research Metrics and Analytics*, *0*. https://doi.org/10.3389/frma.2017.00007
Viseu, A. (2015). Integration of social science into research is crucial. *Nature*, *525*(7569), 291-291. https://doi.org/10.1038/525291a

Wachs, S., & Wright, M.F. (2019). The moderation of online disinhibition and sex on the relationship between online hate victimization and perpetration. *Cyberpsychology, Behavior, and Social Networking*, *22*(5), 300-306. https://doi.org/10.1089/cyber.2018.0551

Waldron, J. (2012). *The harm in hate speech*. Harvard University. https://doi.org/10.4159/harvard.9780674065086

Wang, Z.Y., Li, G., Li, C.Y., & Li, A. (2012). Research on the semantic-based co-word analysis. *Scientometrics*, *90*(3), 855-875. https://doi.org/10.1007/S11192-011-0563-Y

Woolley, S. (2006a). Raw and random. *Forbes*, *177*(5). https://bit.ly/3zqtUmu

Woolley, S. (2006b). The YouTube revolution. *Forbes*, *178*(8), 100-100. https://bit.ly/38hphzp

Wright, M.F., Wachs, S., & Gámez-Guadix, M. (2021). Youths' coping with cyberhate: Roles of parental mediation and family support. [Jóvenes ante el ciberodio: El rol de la mediación parental y el apoyo familiar]. *Comunicar*, *67*, 21-33. https://doi.org/10.3916/C67-2021-02