

ESTRELLA GUALDA  
(Editora)



# TEORÍAS DE LA CONSPIRACIÓN Y DISCURSOS DE ODIO EN LÍNEA EN LA SOCIEDAD DE LAS PLATAFORMAS

Comparación de pautas en las narrativas y redes sociales sobre COVID-19, inmigrantes, refugiados, estudios de género y personas LGTBIQ+

*Dykinson, S.L.*



# **TEORÍAS DE LA CONSPIRACIÓN Y DISCURSOS DE ODIOS EN LÍNEA EN LA SOCIEDAD DE LAS PLATAFORMAS**

**Comparación de pautas en las narrativas y redes sociales sobre  
COVID-19, inmigrantes, refugiados, estudios de género y  
personas LGTBIQ+**



ESTRELLA GUALDA  
(Editora)

# TEORÍAS DE LA CONSPIRACIÓN Y DISCURSOS DE ODIOS EN LÍNEA EN LA SOCIEDAD DE LAS PLATAFORMAS

Comparación de pautas en las narrativas y redes sociales sobre  
COVID-19, inmigrantes, refugiados, estudios de género y  
personas LGTBIQ+

ESTRELLA GUALDA  
EMILIA E. MARTINEZ-BRAWLEY  
EMILIANA MANGONE  
STELLAMARINA DONATO  
MARIELBA SILVA DE ZACARIAS  
WANDA JUARES  
SILVANA FERREYRA  
ALFONSO AMENDOLA  
MARTINA MASULLO  
EMMANUEL BORTHIRY  
ALBA TABOADA-VILLAMARÍN  
CAROLINA REBOLLO-DÍAZ

ISIDRO MARÍN GUTIÉRREZ  
MÓNICA HINOJOSA BECERRA  
ANGEL LUIS TORRES TOUKOUMIDIS  
ELENA RUIZ-ÁNGEL  
PATRICIA RUIZ-ÁNGEL  
FRANCISCO JAVIER SANTOS FERNÁNDEZ  
FRANCESCA CUBEDDU  
MARTA SCOCCO  
LAURA CABRERA ÁLVAREZ  
AILÍN CANILLO  
MARIANNA COPPOLA  
GIUSEPPE MASULLO



**coideso**  
Pensamiento Contemporáneo e  
Innovación para el Desarrollo Social



Grupo de Investigación  
ESTUDIOS  
SOCIALES  
E  
INTERVENCIÓN  
SOCIAL

Grupo de Investigación  
**i2c**  
Ingeniería de la Información  
y el Conocimiento

**CITES**  
centro de investigación

**CISCO LAB**  
Laboratorio de Investigación en Inteligencia Artificial y Big Data

*Dykinson, S. L.*

No está permitida la reproducción total o parcial de este libro, ni su incorporación a un sistema informático, ni su transmisión en cualquier forma o por cualquier medio, sea este electrónico, mecánico, por fotocopia, por grabación u otros métodos, sin el permiso previo y por escrito del editor. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (art. 270 y siguientes del Código Penal).

Diríjase a Cedro (Centro Español de Derechos Reprográficos) si necesita fotocopiar o escanear algún fragmento de esta obra. Puede contactar con Cedro a través de la web [www.conlicencia.com](http://www.conlicencia.com) o por teléfono en el 917021970/932720407.

Este libro ha sido sometido a evaluación por parte de nuestro Consejo Editorial  
Para mayor información, véase [www.dykinson.com/quienes\\_somos](http://www.dykinson.com/quienes_somos)

© Copyright by  
Los autores  
Madrid

Editorial DYKINSON, S.L. Meléndez Valdés, 61 - 28015 Madrid  
Teléfono (+34) 91 544 28 46 - (+34) 91 544 28 69  
e-mail: [info@dykinson.com](mailto:info@dykinson.com)  
<http://www.dykinson.es>  
<http://www.dykinson.com>

ISBN: 978-84-1170-913-2  
Depósito Legal: M-1978-2024  
DOI: 10.14679/2749

*Maquetación:*  
[german.balaguer@gmail.com](mailto:german.balaguer@gmail.com)

# ÍNDICE

<b>PRESENTACIÓN</b> .....	15
---------------------------	----

## I.

### **TEORÍAS DE LA CONSPIRACIÓN Y DISCURSOS DE ODIO EN LÍNEA: APROXIMACIONES TEÓRICAS**

<b>MICRONARRATIVAS DE ODIO EN LÍNEA, TEORÍAS DE LA CONSPIRACIÓN Y DESINFORMACIÓN EN LA LLAMADA “SOCIEDAD DE LAS PLATAFORMAS”</b> .....	23
--	----

ESTRELLA GUALDA

1. INTRODUCCIÓN: MICRONARRATIVAS EN LÍNEA Y SU RELEVANCIA ACTUAL.....	24
2. MICRONARRATIVAS Y COMUNICACIÓN EN LA SOCIEDAD DE LAS PLATAFORMAS.....	25
3. LA ERA DE INTERNET, LOS PROSUMIDORES Y SUS MICRONARRATIVAS.....	27
4. MICRONARRATIVAS EN LÍNEA COMO CONSTRUCTORAS DE HISTORIAS: ASPECTOS FORMALES Y ESTRUCTURALES .....	29
4.1. Brevidad .....	29
4.2. Formatos y contenidos atractivos .....	30
4.3. Difusión (o viralidad) que alcanzan los mensajes.....	30
4.4. Empleo de recursos comunicativos y estilísticos .....	31
5. LAS MICRONARRATIVAS EN LÍNEA COMO VEHÍCULOS DE TEORÍAS DE LA CONSPIRACIÓN Y DISCURSOS DE ODIO: EL CASO DE LA COVID-19, INMIGRANTES, REFUGIADOS Y PERSONAS LGBTIQ+.....	35
5.1. Micronarrativas en línea, conspirativas y de odio .....	35
5.2. Micronarrativas de odio en línea y teorías de la conspiración: COVID-19, inmigrantes y refugiados y personas LGBTIQ+ .....	36
6. A MODO DE CONCLUSIÓN: APUNTES PARA CONTRARRESTAR LAS MICRONARRATIVAS NOCIVAS Y FRAUDULENTAS EN UN CONTEXTO DE AUGE DE LA INTELIGENCIA ARTIFICIAL [IA].....	38
REFERENCIAS.....	39

**HATE SPEECH. A BIRD'S EYE VIEW OF A CONUNDRUM OF INTERNATIONAL EPIDEMIC PROPORTIONS .....43**

EMILIA E. MARTINEZ-BRAWLEY

1. INTRODUCTION .....	43
2. CYBERSPACE, THE INTERNET AND THE BROAD NETWORK OF ELECTRONIC COMMUNICATION .....	44
3. SOCIO-CULTURAL AND HISTORICAL CONSIDERATIONS.....	44
4. THE HISTORICAL PROBLEM OF DEFINITIONS .....	47
5. THE EPIDEMIC NATURE OF HATE SPEECH VIA THE INTERNET. TWO SALIENT FACILITATING FACTORS. ....	51
5.1. Manipulation .....	51
5.2. The contagion effect of online communication .....	52
6. SOME CONCLUDING REFLECTIONS .....	54
REFERENCES .....	57

**ONLINE HATE SPEECH AS AN UNINTENDED AND PARADOXICAL OUTCOME OF CANCEL CULTURE.....61**

EMILIANA MANGONE / STELLAMARINA DONATO

1. CANCEL CULTURE AND ITS RELATIONSHIP TO HATE SPEECH....	61
2. RELATED CONCEPTS AND UNINTENDED CONSEQUENCES OF CANCEL CULTURE .....	65
3. FROM CANCEL CULTURE TO HATE SPEECH: THE DYNAMICS FOR THE LGBTQIA+ COMMUNITY AND MIGRANTS .....	68
REFERENCES .....	71

**AUTOMATIC DETECTION OF HATE SPEECH: RESEARCH CHALLENGES AND OPPORTUNITIES .....75**

MARIELBA SILVA DE ZACARIAS

1. INTRODUCTION .....	76
2. CHALLENGES AND RESEARCH AVENUES.....	76
2.1. What is Hate Speech? .....	76
2.2. The Contextual Nature of Hate Speech .....	77
2.3. Temporal and Dynamic Nature of Hate Speech .....	77
2.4. A Multilingual and Multicultural problem .....	78
2.5. Dataset and Annotation Challenges.....	79
3. THE LARGER PICTURE .....	80
4. CONCLUSIONS.....	82
REFERENCES .....	82

## II. COVID-19, TEORÍAS DE LA CONSPIRACIÓN, INFODEMIA Y ODIO EN LÍNEA

### **MOVILIZACIONES DURANTE LA PANDEMIA DE COVID-19 EN ARGENTINA. UNA EXPLORACIÓN DE MENSAJES Y USUARIOS EN TWITTER DURANTE LOS “BANDERAZOS” (MAYO-OCTUBRE 2020) .....89**

WANDA JUARES / SILVANA FERREYRA

1. INTRODUCCIÓN .....	89
2. ANTECEDENTES.....	91
3. CONSTRUCCIÓN DEL CORPUS .....	93
4. EL CONTENIDO DE LAS PROTESTAS.....	94
5. TRES MIRADAS SOBRE LOS USUARIOS .....	100
6. REFLEXIONES FINALES.....	108
REFERENCIAS.....	109

### **MICRONARRATIVAS NEGACIONISTAS Y TEORÍAS DE LA CONSPIRACIÓN SOBRE LA “PLANDEMIA” EN TWITTER.....113**

ESTRELLA GUALDA

1. TEORÍAS DE LA CONSPIRACIÓN Y “PLANDEMIA” DE COVID-19...113	113
2. OBJETIVOS, HIPÓTESIS Y LÍNEAS DE EXPLORACIÓN .....	116
3. MÉTODOS, TÉCNICAS Y FUENTES DE LA INVESTIGACIÓN.....117	117
3.1. Diseño metodológico .....	117
3.2. Producción de datos: Colección y fuente de datos .....	117
3.3. Preprocesamiento y análisis.....	119
4. EVOLUCIÓN DE LOS TUIOS SOBRE LA PLANDEMIA Y RASGOS BÁSICOS.....	119
4.1. Rasgos básicos.....	119
4.2. Evolución temporal y comparativa de las etiquetas #plandemia y #plandemic .....	120
5. PLANDEMIC & PLANDEMIA: PRINCIPALES TOKENS EN EL DATASET .....	122
6. MICRONARRATIVAS DE HASHTAGS EN TORNO A LA PLANDEMIA.....	124
7. DISCUSIÓN Y CONCLUSIONES .....	132
REFERENCIAS.....	134

### **COVID-19 NARRATIVE THROUGH SOCIAL MEDIA: THE TELEGRAM CASE.....137**

ALFONSO AMENDOLA / MARTINA MASULLO

1. THE COVID-19 NARRATIVE THROUGH SOCIAL: A GENERAL THEORETICAL FRAMEWORK.....	137
--	-----

2. INFODEMIC, COMMUNICATIVE CRISIS AND SPONTANEOUS SELF-MANAGEMENT OF COMMUNICATION .....	139
3. TELEGRAM CASE: OBJECTIVES, METHODOLOGY AND RESULTS ...	142
4. CONCLUSIONS .....	148
REFERENCES .....	148

<b>COVID-19 Y CAMPAÑAS POLÍTICAS EN TWITTER: UN ANÁLISIS DE LOS CANDIDATOS EN LAS ELECCIONES LEGISLATIVAS 2021 EN ARGENTINA .....</b>	<b>151</b>
<b>EMMANUEL BORTHIRY</b>	

1. INTRODUCCIÓN .....	151
2. ANTECEDENTES DE INVESTIGACIÓN .....	152
3. ASPECTOS METODOLÓGICOS.....	154
4. LA EVOLUCIÓN DEL TÓPICO COVID-19.....	157
5. CANDIDATOS COVID-19.....	159
6. TÓPICOS Y ENFOQUES PARTIDARIOS.....	164
7. CONCLUSIONES.....	167
REFERENCIAS.....	168

### III.

#### TEORÍAS DE LA CONSPIRACIÓN Y DISCURSOS DE ODIO EN LÍNEA SOBRE INMIGRANTES Y/O REFUGIADOS

<b>CONSPIRACY THEORIES ON THE ORIGIN OF THE CORONAVIRUS AND ITS IMPACT ON THE PERCEPTION OF CHINA AND CHINESE ON TWITTER.....</b>	<b>175</b>
<b>ALBA TABOADA-VILLAMARÍN / CAROLINA REBOLLO-DÍAZ</b>	

1. INTRODUCTION .....	175
2. METHODS.....	178
2.1. Case study selection and data collection .....	178
2.2. Data analysis.....	179
3. MAIN ISSUES SURROUNDING THE ORIGIN OF THE CORONAVIRUS .....	181
3.1. China's role in the debate on the origin of the SARS-CoV-2 virus...181	
3.2. China's impact on the perception of the pandemic in the United States .....	182
3.3. India-China Geopolitical Rivalry in the Context of the Health Crisis.....	182
3.4. Evaluation of Chinese vaccination strategies and their global impact .....	183
4. MAIN SENTIMENTS AND CONSPIRACY THEORIES.....	184
5. CONCLUSIONS.....	187
REFERENCES .....	188

**CHEQUEANDO DISCURSOS DE ODIOS EN LÍNEA SOBRE LA MIGRACION VENEZOLANA EN ECUADOR .....193**

ISIDRO MARÍN GUTIÉRREZ / MÓNICA HINOJOSA BECERRA / ANGEL LUIS TORRES TOUKOUMIDIS

1. INTRODUCCIÓN .....	193
2. MARCO TEÓRICO .....	194
3. LA CRISIS MIGRATORIA VENEZOLANA.....	195
4. METODOLOGÍA .....	196
5. RESULTADOS .....	197
6. CONCLUSIONES.....	198
REFERENCES.....	200

**IV.**

**ESTUDIOS DE GÉNERO, DISCURSOS DE ODIOS EN LÍNEA Y CIBERVIOLENCIAS**

**UNA APROXIMACIÓN AL ORIGEN DE LA MISOGINIA Y EL DISCURSO DE ODIOS EN LÍNEA HACIA LAS MUJERES .....205**

ELENA RUIZ-ÁNGEL / PATRICIA RUIZ-ÁNGEL / FRANCISCO JAVIER SANTOS FERNÁNDEZ

1. INTRODUCCIÓN .....	205
2. APROXIMACIÓN AL ORIGEN DEL DISCURSO MISÓGINO .....	207
3. CONCEPTUALIZACIÓN DE LA MISOGINIA.....	210
4. EL FENÓMENO DE LA MANOSFERA Y LA PROLIFERACIÓN DEL DISCURSO DE ODIOS EN LÍNEA HACIA LAS MUJERES.....	213
REFERENCIAS.....	216

**ONLINE HATE SPEECH AS A FORM OF CYBERVIOLENCE: COMMUNICATION PRACTICES, SOCIAL PERCEPTION AND VULNERABILITIES IN THE EXPERIENCE OF YOUNG PEOPLE .....221**

FRANCESCA CUBEDDU / MARTA SCOCCO

1. INTRODUCTION .....	221
2. HATE SPEECH ON THE WEB AND SOCIAL MEDIA.....	223
3. THE SOCIAL IMPACTS OF ONLINE HATE SPEECH.....	225
4. METHOD .....	227
5. ONLINE HATE SPEECH AND CYBERVIOLENCE: KNOWLEDGE OF THE PHENOMENON AND RISK PERCEPTION AMONG YOUNG PEOPLE .....	228
6. INDIVIDUAL AND SOCIAL GROUP VULNERABILITIES IN ONLINE FORMS OF VIOLENCE.....	230
7. CONCLUSIONS .....	231
REFERENCES.....	233

**LA CULTURA DE LA CANCELACIÓN Y SUS CONEXIONES CON LAS TEORÍAS DE LA CONSPIRACIÓN Y LOS DISCURSOS DE ODIOS EN LÍNEA.....237**

LAURA CABRERA ÁLVAREZ

1. INTRODUCCIÓN .....237
2. LA CULTURA DE LA CANCELACIÓN Y SUS CONEXIONES CON LOS DISCURSOS DE ODIOS Y LAS TEORÍAS DE LA CONSPIRACIÓN .....239
3. EL FENÓMENO DE LA CULTURA DE LA CANCELACIÓN: DE LA PRÁCTICA SOCIAL DE “CANCELAR” A SUS CONSECUENCIAS.....241
- REFERENCIAS .....243

**V.**

**NARRATIVAS, DISCURSOS E IMAGINARIOS SOCIALES SOBRE PERSONAS LGTBIQ+**

**BLOGS LESBIANOS: UNA EXPLORACIÓN DE SUS USOS Y CONTENIDO POR EL ACTIVISMO LGTBIQ+ A MEDIADOS DE LOS DOS 2000S EN ARGENTINA.....249**

AILÍN CANILLO

1. INTRODUCCIÓN .....249
2. ANTECEDENTES.....250
3. METODOLOGÍA .....252
4. RESULTADOS Y ANÁLISIS .....254
5. DISCUSIÓN Y CONCLUSIONES .....269
- REFERENCIAS.....271

**DISCRIMINATION AND HATE SPEECH AMONG TRANSGENDER INDIVIDUALS IN AN ONLINE COMMUNITY: BETWEEN PASSING AND TRANS-NORMATIVITY .....273**

MARIANNA COPPOLA / GIUSEPPE MASULLO

1. INTRODUCTION .....273
2. HATE SPEECH TOWARDS AND AMONG LGBTQ+ INDIVIDUALS...275
3. SEXUAL NORMATIVITY AND TRANS-NORMATIVITY .....276
4. THE RESEARCH.....278
  - 4.1. The Informa Trans Italia community.....279
5. RESULTS AND DISCUSSION.....280
  - 5.1. Defining and passing: the main forms of intra-categorical discrimination within the Informa Trans Italia community. ....280
  - 5.2. Discriminated users and targets of hate speech among binary and non-binary transgender individuals: a typology .....283
6. CONCLUSIONS.....286
- REFERENCES .....287

<b>HOMOFOBIA 2.0: UNA APROXIMACIÓN TEÓRICA AL DISCURSO DE ODIO LGTBFÓBICO EN LÍNEA Y LAS ESTRATEGIAS PARA DESARTICULARLO</b> .....	289
FRANCISCO JAVIER SANTOS FERNÁNDEZ / ELENA RUIZ-ÁNGEL / PATRICIA RUIZ-ÁNGEL	
1. INTRODUCCIÓN .....	289
2. HOMOFOBIA.....	290
3. LA RADICALIZACIÓN EN LÍNEA HACIA LAS PERSONAS LGTBIQ+ .....	291
4. CONSECUENCIAS DEL DISCURSO DE ODIO HOMOFÓBICO EN LÍNEA .....	293
5. ACTIVISMO DIGITAL Y NUEVOS MOVIMIENTOS SOCIALES LGTBIQ+ .....	294
6. LA CONTRANARRATIVA COMO ESTRATEGIA DE SUBVERSIÓN DEL DISCURSO DE ODIO EN LÍNEA .....	296
7. EL USO DEL HUMOR Y LA REAPROPIACIÓN DE SÍMBOLOS COMO CONTRA-NARRATIVA HACIA LA HOMOFOBIA EN LÍNEA: EL CASO DE GAYSPER.....	297
REFERENCIAS.....	300



## PRESENTACIÓN

Este libro, que lleva por título *Teorías de la conspiración y discursos de odio en línea en la sociedad de las plataformas*, reúne un conjunto de contribuciones internacionales en torno a estos dos ejes temáticos desde una perspectiva tanto teórica, en sus primeros capítulos, como aplicada, a través de la presentación de varios aspectos donde sus autores exploran cómo se concretan las narrativas conspirativas y el odio en línea en conexión con (a) la pandemia de COVID-19, (b) los inmigrantes y/o refugiados, (c) el odio y las ciberviolencias dirigidas principalmente a mujeres, (d) o las narrativas, discursos e imaginarios sociales sobre personas LGTBIQ+.

Este libro es uno de los productos que resultan de la ejecución del proyecto de I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGTBIQ+ [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDER A way of making Europe”<sup>1</sup>. Si bien este proyecto se adscribe oficialmente a la Universidad de Huelva, en España, y al centro de investigación “Pensamiento Contemporáneo e Innovación para el Desarrollo Social, COIDESO”, los investigadores que participan en el mismo se vinculan a este centro y a los grupos de investigación “Estudios Sociales E Intervención Social, ESEIS” e “Ingeniería de la Información y del Conocimiento, I2C”, el “Centro de Investigación en Tecnología, Energía y Sostenibilidad, CITES” y el “Laboratorio de Ciencias Sociales Computacionales Aplicadas (CISCOA-Lab).

Igualmente, contamos con el privilegio de que, junto al equipo interdisciplinar de la Universidad de Huelva, un conjunto de investigadores de las áreas de Ciencias Sociales, Humanidades y Técnicas de las Universidades de Málaga, la Universidad Autónoma de Madrid (España), la Università degli Studi di

---

<sup>1</sup> Otras publicaciones o actividades de transferencia del proyecto se pueden seguir de forma actualizada en la página web del proyecto: <http://www.uhu.es/nonconspirahate/>

Salerno (Italia), la Universidade do Algarve (Portugal) y Universidad Nacional Mar del Plata, CONICET/INHUS (Argentina) participen en este proyecto. El libro que tiene en sus manos cuenta con contribuciones de investigadores que trabajan en los centros anteriores y que están implicados en la ejecución del proyecto citado, pero se han sumado igualmente contribuciones importantes de personas que, sin estar vinculadas directamente a dicho proyecto (Arizona State University, USA; University of Roma Tre, University of Macerata y LUISS University Rome, Italy; Universidad Politécnica Salesiana de Ecuador - sede Cuenca) cuentan con gran experiencia en las temáticas de los capítulos que firman, lo que nos ofrece un panorama más rico para la discusión del estado de la cuestión relativo a las teorías de la conspiración y los discursos de odio, así como para el estudio de casos desde una perspectiva internacional.

Quizás uno de los aspectos más novedosos de este libro es reunir contribuciones en torno a dos cuestiones que conectan entre sí, las teorías de la conspiración y los discursos del odio, escritas desde una diversidad de ciencias y de disciplinas que miran interdisciplinariamente su objeto de estudio, propiciando, a través de conjugar la teoría con los estudios de caso, una mejor comprensión de dos problemáticas que tienen una gran centralidad en las sociedades actuales y que a veces están marcando el ritmo de fenómenos sociales y políticos. El foco principal del libro, como se aprecia en su título, es examinar precisamente las narrativas conspirativas y los discursos de odio en el contexto de la sociedad de las plataformas presentando una atención especial a estos fenómenos cuando se despliegan en plataformas, como Twitter o X, Instagram, Telegram, u otras, así como a través de servicios de mensajería o páginas web tales a Facebook, Youtube, WhatsApp, amén de otras que se encuentran en Internet.

El libro se divide en cinco bloques temáticos. El primero de ellos, titulado **“Teorías de la conspiración y discursos de odio en línea: aproximaciones teóricas”**, sitúa aspectos generales de carácter teórico y conceptual a través de cuatro capítulos. Comienza el libro con un primer capítulo de Estrella Gualda que, prestando atención especial a la definición de las micronarrativas de odio en línea, conecta estas narrativas con las teorías de la conspiración y la desinformación en la “sociedad de las plataformas”, describiendo igualmente algunos parámetros clave de estas sociedades que influyen en la producción de micronarrativas. Emilia Martínez-Brawley, a continuación, se detiene en la conceptualización, teorización e historia del término discurso de odio hasta llegar a la actualidad desarrollando el argumento metafórico de que los discursos de odio cuentan con una naturaleza epidémica, por la forma en que se propagan con la ayuda de Internet. Ante la situación actual, apuesta por el buen juicio, la racionalidad y el sentido común a la hora de frenar estos discursos. Emiliana Mangone y Stellamarina Donato, se detienen en el desarrollo de un argumento más específico conectando el discurso de odio en línea con la llamada cultura de

la cancelación, sugiriendo, como línea de investigación a explorar en el ámbito de la comunicación, que el discurso de odio es un resultado no pretendido y paradójico de la cancelación, un subproducto de esta misma cultura. Para terminar este bloque más teórico, Marielba Zacarias discute sobre aspectos más técnicos respecto a los retos y oportunidades que presenta hoy la detección automática de los discursos de odio, resaltando su complejidad y la necesidad de contar tanto con soluciones tecnológicas como con la colaboración entre investigadores, responsables políticos, *stakeholders* de la industria y organizaciones sociales civiles para lograr sistemas de detección del discurso del odio más eficaces, sin que se olviden los aspectos éticos.

Un segundo bloque, ya con un foco más aplicado, se concentra en la pandemia de **COVID-19, las teorías de la conspiración, la infodemia y el odio en línea**, conectando a través de varios capítulos estos aspectos. Wanda Juarez y Silvana Ferreyra dirigen su atención a las movilizaciones durante la pandemia en Argentina explorando las principales narrativas, especialmente de odio, que se difunden en los primeros momentos de la pandemia en estas movilizaciones. Igualmente, se detienen en la indagación de los usuarios que participaron en las mismas, ofreciendo a los lectores un mapa sobre el activismo digital en Argentina. En el trabajo de Estrella Gualda, con un foco en las micronarrativas negacionistas y las teorías de la conspiración sobre la llamada “plandemia” se exploran los relatos articulados internacionalmente en torno a los hashtags “#plandemia” y “#plandemic”. Igualmente, se abordan las conexiones de las micronarrativas negacionistas con los movimientos anticuarentena, antivacunación y antimascarillas, identificándose en los relatos aspectos como el rechazo a los gobiernos, o el descrédito de fuentes de autoridad, o discursos incívicos, todos ellos que pueden suponer un importante riesgo para las democracias. El trabajo de Alfonso Amendola y Martina Masullo se centra en el caso de Telegram, como otra plataforma clave en la producción de narrativas relativas a la pandemia de COVID-19, y rastrea y analiza algunas de las principales *fake news*, teorías de la conspiración y discursos de odio que se han generado en Telegram en torno a la pandemia, considerando Telegram como una de las plataformas más responsables del auge de noticias falsas y de desinformación sobre la COVID-19 en Italia. En este bloque, por último, el capítulo de Emmanuel Borthiry, que vuelve de nuevo a Argentina de la mano del estudio de las campañas políticas en Twitter, se detiene en el rol clave que para todos los candidatos tuvo, en el momento de la campaña de las elecciones legislativas del año 2021, participar en los debates, confrontar o expresar opiniones respecto a la situación epidemiológica, las medidas de cuidado, la economía o la campaña de vacunación.

Los siguientes capítulos del libro se enmarcan en el bloque sobre **teorías de la conspiración y discursos de odio en línea sobre inmigrantes y/o refugiados**.

El capítulo de Alba Taboada-Villamarín y Carolina Rebollo-Díaz, que conecta también con la pandemia de coronavirus, explora una de las teorías de la conspiración más difundidas durante la pandemia, esto es, la que se centra en atribuir el origen del coronavirus a China, y sobre todo, el impacto que se puede rastrear en Twitter respecto a las percepciones sobre China y sobre la población China destacándose en los datos la presencia de emociones negativas como el miedo y la desconfianza hacia China y hacia la Organización Mundial de la Salud. Por otra parte, el trabajo de Isidro Marín Gutiérrez, Mónica Hinojosa Becerra y Angel Luis Torres Toukoumidis, nos presenta, más centrado en la desinformación y el *fact-checking*, una experiencia de comprobación de los discursos de odio en línea sobre la migración venezolana en Ecuador, indagando en las noticias sobre los venezolanos expuestas desde la agencia oficial de verificación ecuatoriana, *Ecuador Chequea*, para contrastar las tipologías de noticias falsas que se presentan junto con el discurso que se encuentra en las mismas y reflexionar sobre cómo la desinformación ha afectado el comportamiento de la población ecuatoriana hacia los venezolanos y cómo las agencias de *fact-checking* pueden combatirla.

Los dos bloques que siguen presentan trabajos enmarcados en los estudios de género, donde las teorías de la conspiración y los discursos de odio, e incluso la violencia, se dirige hacia mujeres o hacia personas LGTBQ+. De esta forma, el cuarto bloque, sobre **Estudios de género, discursos de odio en línea y ciberviolencias** contiene capítulos más centrados en aspectos relativos a las mujeres, mientras que el siguiente bloque, lo hace sobre las personas LGTBQ+. Se encuadran en el cuarto bloque los trabajos de Elena Ruiz-Ángel, Patricia Ruiz-Ángel y Francisco Javier Santos Fernández que llevan a cabo una aproximación al origen de la misoginia y el discurso de odio en línea hacia las mujeres, resaltando aspectos como que el discurso de odio hacia las mujeres ha aumentado en las redes sociales y en la manofera durante y después del confinamiento de la pandemia del COVID-19. Francesca Cubeddu y Marta Scocco concentran su atención en el discurso de odio en línea como una forma de ciberviolencia estudiando las prácticas comunicativas, la percepción social y las vulnerabilidades que experimentan los jóvenes, concluyendo que es necesaria la promoción de la educación para la ciudadanía digital en las generaciones más jóvenes, siendo la vía educativa una estrategia real para combatir estos fenómenos. Así mismo, se sugiere la necesidad de la protección de los derechos y se destaca el activismo social que cada uno de nosotros puede movilizar para prevenir y contrarrestar los riesgos y peligros de la digitalización. El trabajo de Laura Cabrera Álvarez, que conecta la cultura de la cancelación con las teorías de la conspiración y los discursos de odio en línea, refleja de nuevo una de las vulnerabilidades que presenta la sociedad de las plataformas, en la medida en que la digitalización plantea un escenario donde la práctica de la cancelación emerge y crece exponencialmente, como un mecanismo que se conecta a veces con la expresión del odio y la divulgación de teorías conspiratorias que pueden

afectar profundamente a las personas u organizaciones e instituciones que son canceladas.

El último bloque, siguiendo la línea de los estudios de género, se concentra específicamente en **narrativas, discursos e imaginarios sociales sobre personas LGTBIQ+**. Comienza con el capítulo de Ailín Canillo sobre blogs lesbianos, explorando los usos y contenido de este activismo LGTBIQ+ en Argentina caracterizando el tipo de activismo que representan estos blogs para determinar si estaban más enfocados en la defensa de derechos, en la promoción de cambios políticos o en el reconocimiento de identidades. Marianna Coppola y Giuseppe Masullo presentan un trabajo sobre la discriminación y discurso de odio entre personas transexuales en una comunidad en línea preguntándose cuestiones como: ¿cuáles son las características fenotípicas, psicológicas y sociales en las que se fija idealmente la comunidad transexual a efectos de normalización?, ¿en qué medida se apartan estas características del modelo cis-heteronormativo? O ¿qué estrategias argumentativas utilizan los transexuales para elogiar o desvalorizar a las personas que no se ajustan a los modelos dominantes? La respuesta a estas preguntas se aborda a través de una etnografía virtual en grupos de Facebook de temática transgénero en Italia. Por último, el trabajo de Francisco Javier Santos Fernández, Elena Ruiz-Ángel, Patricia Ruiz-Ángel, sobre Homofobia 2.0 presenta una aproximación teórica al discurso de odio LGTBfóbico en línea y las estrategias para desarticularlo y cierra el libro planteando no solo algunos de los riesgos y problemas que genera la homofobia o discursos de odio homofóbico en línea, sino también presentando un estudio de caso (Gaysper) donde por la vía del activismo digital y con el uso del humor se puede observar que es posible contrarrestar el daño que hacen los discursos de odio en línea.

Estrella Gualda

*Responsable del Proyecto PID2021-123983OB-I00*



**I.**  
**TEORÍAS DE LA CONSPIRACIÓN  
Y DISCURSOS DE ODIO EN LÍNEA:  
APROXIMACIONES TEÓRICAS**



# MICRONARRATIVAS DE ODIOS EN LÍNEA, TEORÍAS DE LA CONSPIRACIÓN Y DESINFORMACIÓN EN LA LLAMADA “SOCIEDAD DE LAS PLATAFORMAS”<sup>1</sup>

ESTRELLA GUALDA

*Catedrática de Sociología*

*Universidad de Huelva, ESEIS/COIDESO, España*

DOI: 10.14679/2750

## Sumario:

1. Introducción: micronarrativas en línea y su relevancia actual.
  2. Micronarrativas y comunicación en la llamada “sociedad de las plataformas”.
  3. La Era de Internet, los prosumidores y las micronarrativas.
  4. Micronarrativas en línea como constructoras de historias: aspectos formales y estructurales.
    - 4.1. Brevedad.
    - 4.2. Formatos y contenidos atractivos.
    - 4.3. Difusión o viralidad que alcanzan los mensajes.
    - 4.4. Empleo de recursos comunicativos y estilísticos.
      - 4.4.1. Los tropos y la retórica: de las metáforas a la ironía.
      - 4.4.2. Las etiquetas (#): elementos clave en algunas narrativas que se difunden en Internet.
      - 4.4.3. Emojis, símbolos, imágenes o vídeos: buscando una mayor expresividad y emotividad de las micronarrativas.
  5. Las micronarrativas en línea como vehículos de teorías de la conspiración y discursos de odio: el caso de la COVID-19, inmigrantes, refugiados y personas LGBTIQ+.
    - 5.1. Micronarrativas en línea, conspirativas y de odio.
    - 5.2. Micronarrativas de odio en línea y teorías de la conspiración: COVID-19, inmigrantes y refugiados y personas LGBTIQ+.
  6. A modo de conclusión: apuntes para contrarrestar las micronarrativas nocivas y fraudulentas en un contexto de auge de la inteligencia artificial.
- Referencias.

---

<sup>1</sup> Este trabajo forma parte del Proyecto I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGBTIQ [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDER A way of making Europe” (<http://www.uhu.es/nonconspirahate/>). También agradecemos el apoyo de los centros de investigación de la Universidad de Huelva, el grupo de investigación “Estudios Sociales e Intervención Social, ESEIS” y el centro “Pensamiento Contemporáneo e Innovación para el Desarrollo Social, COIDESO”. El capítulo se ha beneficiado igualmente de mi participación en el proyecto de excelencia andaluz: “Impacto de la desinformación en Andalucía: análisis transversal de las audiencias y las rutinas y agendas periodísticas. DesinfoAnd” (ProyExcel 00143, <https://www.disinformationresearch.uma.es>).

«In fact, the fictional narratives conveyed on social media can be defined as micronarratives such as representations which 1) possess some of the conditions of narrativity, 2) are composed of autonomous multimodal fragments, spread across the digital narrative space, and 3) require the user's cognitive action of interpretation in order to be connected to other fragments» (Venditti, Piredda & Mattana, 2017: S279).

## 1. INTRODUCCIÓN: MICRONARRATIVAS EN LÍNEA Y SU RELEVANCIA ACTUAL

Este capítulo reflexiona sobre los cambios que se han producido en los procesos de comunicación cuando esta ocurre en el escenario de la llamada “sociedad de las plataformas” [*The Platform Society*], desarrollada en el contexto de las sociedades occidentales (Van Dijck et al., 2018). Específicamente, presentamos en estas páginas qué entendemos por “micronarrativas en línea”, como una de las formas de comunicación que, con el auge de las plataformas de redes sociales (Gualda, 2018), ha crecido exponencialmente. Las micronarrativas en línea, como narrativas que se expresan en formato muy breve, se multiplican en un contexto tecnológico donde las plataformas de redes sociales imponen limitaciones, en cuanto a la concisión o el número de caracteres o palabras que como máximo es posible introducir en un mensaje, un comentario, o un *post* que se publica en plataformas o páginas web como Twitter, Instagram, TikTok, Facebook, Youtube, u otras, en cada una de las cuales se limita el número máximo de caracteres de cada mensaje. Paralelamente, en la medida en que muchos usuarios leen, comentan o interaccionan con estos textos a través de, por ejemplo, un *smartphone*, esta práctica social contribuye igualmente a que se potencie la brevedad como formato más cómodo en el que se produce la comunicación. De forma similar ocurre cuando se trata de imágenes o vídeos.

La brevedad de las micronarrativas en línea en la sociedad de las plataformas puede entenderse tanto como una forma cuanto como un formato de expresión. La concisión, en cuanto a la forma, es un recurso estilístico mediante el cual los autores usan el lenguaje como vía de comunicación para transmitir un mensaje, con el objetivo de provocar algún efecto en particular, por ejemplo, un mayor impacto del mensaje a transmitir, focalizando la atención en algún aspecto clave. Como formato de expresión, la brevedad se refiere a la extensión de la narrativa que, en el caso de las plataformas de redes sociales, es estipulado habitualmente por la misma plataforma. Recordemos el límite de 140 caracteres que Twitter en su momento impuso a los tuits, y que modificó posteriormente a 280, condicionando la manera en que podían expresarse los mensajes en esta red. Esta pauta coercitiva, que limita el número de caracteres en la comunicación en diferentes redes sociales o en páginas web tiene, de acuerdo con la investigación desarrollada, impactos significativos en los contenidos que se

crean y que se comparten. Igualmente, abre una discusión sobre los vínculos entre el número de caracteres y el compromiso, el *engagement* en marketing, entre los factores que tienen que ver con la mejor legibilidad en las redes sociales (Cuevas et al., 2022; Gkikas et al, 2022; Shahbaznezhad, 2021).

Las micronarrativas en línea, en las plataformas de redes sociales, emergen en formato breve condicionadas por la propia plataforma en la que el *micropost* o mensaje breve se publica en una plataforma de microblogs tipo Twitter. Igualmente, las encontramos en forma de comentarios sucintos a blogs, en foros o en diversidad de páginas webs. Sea por las limitaciones en número de palabras o caracteres que se imponen por la plataforma, sea por las prácticas comunicativas que progresivamente van adquiriendo los usuarios al consumir y producir mensajes breves, sea por el ahorro en tiempo que supone esta forma de comunicación, o por otros factores, las micronarrativas en línea se han convertido en una de las principales formas de expresión hoy en día, con gran impacto en diferentes áreas, tales a la comunicación política, la social, la científica, y otras.

En este capítulo describimos inicialmente algunos rasgos clave del contexto social y tecnológico en el que emergen las micronarrativas en línea, para después detenernos a argumentar cómo estas micronarrativas se han convertido en constructoras de relatos y argumentarios y cuáles son los rasgos formales y de formato que las caracterizan. Posteriormente, nos referimos a la relevancia adquirida por las micronarrativas conspirativas, aquellas que difunden teorías de la conspiración en línea, y las micronarrativas de odio en línea, como dos formas de expresión por excelencia de mensajes en las redes sociales asociados en gran medida al auge de la desinformación. A continuación, el capítulo se detiene en algunos ejemplos para explicar cómo las micronarrativas son frecuentemente usadas como soporte o vehículo conductor para difundir teorías de la conspiración y odio en internet. Terminamos sugiriendo algunas recomendaciones para neutralizar o contrarrestar la amplia difusión de estas micronarrativas en el contexto actual. La reflexión sobre cómo las micronarrativas en línea se han convertido en un elemento clave en la comunicación que tiene lugar en la sociedad de las plataformas nos da algunas pistas para comprender cómo se articulan específicamente las micronarrativas conspirativas y de odio.

## 2. MICRONARRATIVAS Y COMUNICACIÓN EN LA SOCIEDAD DE LAS PLATAFORMAS

Los avances tecnológicos habitualmente suponen un marco de posibilidades para el cambio social. En sociedades donde Internet es accesible a gran parte de su población, así como los teléfonos móviles o *smartphones*, u otras tecnologías que permiten impulsar la interconectividad, somos testigos del cambio

fundamental que se produce en algunas pautas de comunicación, así como en las herramientas, estrategias o formatos que usamos a estos efectos.

Una práctica común en la historia de la sociología y de otras ciencias sociales ha sido la de denominar o “apellidar” a la sociedad atendiendo a algunos de sus rasgos característicos. De esta forma, junto a nociones ya clásicas como las de “sociedad postindustrial” de Bell (1973), la “sociedad de la información” de Masuda (1980), la “tercera ola de la información” de Toffler (1980), o la “sociedad del conocimiento” de Naisbitt (1982), donde ya se iban presentando las potencialidades y la capacidad revolucionaria de una tecnología como Internet, textos más recientes van dando forma a la diversidad de impactos que la World Wide Web (la web) iba a provocar, decantándose diversos autores por subrayar algunos rasgos de las sociedades contemporáneas frente a otros, a través de “bautizar” sintéticamente a las sociedades que describen. Algunos de los “apellidos” más populares han destacado aspectos críticos a raíz de la llegada de Internet, como que nos encontramos ante una “sociedad de la vigilancia” (Lyon, 1994, 2007; Zukoff, 2019), o ante una “sociedad líquida” (Bauman, 2000), o incluso de la “vigilancia líquida” (Lyon, 2018; Lyon y Bauman, 2019).

Junto a los anteriores, otros autores han destacado el impacto que Internet tiene para la sociedad. Ritzer, por ejemplo, cuando alude a la “sociedad mcdonaldizada” (Ritzer, 1996) destaca que Internet contribuye a que la mcdonaldización de la sociedad crezca, y con ello al aumento de algunas características negativas como la pérdida de la diversidad, la alienación y la despersonalización, al desarrollarse aspectos como estandarización, la previsibilidad, el control y la eficiencia para ofrecer productos y servicios a los clientes.

Manuel Castells (1997) elabora el concepto “sociedad red”, muy ligado al desarrollo de Internet. En la sociedad red las personas están conectadas a través de redes digitales y esta gran conectividad produce cambios en la organización social y una diversidad de transformaciones en diferentes esferas. Más recientemente, la idea de la “sociedad de las plataformas” de Van Dijck *et al.* (2018) profundiza precisamente en la interacción de las personas y cómo esta se ve modificada por Internet, y muy especialmente por la participación de las personas en plataformas de redes sociales, con consecuencias a veces negativas como la polarización social o la pérdida de la privacidad. Otros autores, como Narula (2022), que elabora la idea de “sociedad virtual” ponen encima de la mesa algunos cambios que podrían tener lugar con el desarrollo del metaverso. La sociedad virtual, que se describe como la evolución natural de la sociedad red, se caracterizaría por la existencia del metaverso, como espacio digital inmersivo en el que las personas pueden interactuar entre sí de forma virtual experimentando el mundo de una manera diferente.

Las diversas etapas por las que la sociedad contemporánea va pasando, al tiempo que las tecnologías asociadas a Internet evolucionan, presentan igual-

mente cambios en la manera en que se produce la comunicación. De esta forma, hay cambios en cómo se articulan las micronarrativas en línea, entendidas como narrativas breves que se comparten en Internet. En la sociedad red, caracterizada por la conectividad entre personas y organizaciones a través de internet, y en la posibilidad (especialmente en regímenes democráticos) de expresarse libremente a través de la red o de promover nuevas formas de participación ciudadana (Castells, 1997), se inaugura un modelo de comunicación donde las personas pueden compartir sus historias de forma rápida y sencilla, lo que potencia el dinamismo, la diversidad y la creatividad de las micronarrativas a través del desarrollo de blogs, foros, mensajería, redes sociales o nuevas tecnologías como la realidad virtual o la realidad aumentada, que presentan diversos soportes mediante los cuales se pueden narrar historias.

En cambio, en la sociedad de las plataformas (Van Dijck et al., 2018), se ha avanzado en la estandarización por influencia de las empresas tecnológicas que obligan en mayor medida a los usuarios a usar unos formatos y normas determinadas, lo cual limita la creatividad y la diversidad de las narrativas. Igualmente, las empresas tienen mayor control y pueden restringir la libertad de expresión, o con apoyo de la algoritmia y las intenciones de comercialización, afectar a su autenticidad. Respecto a la formulación de la sociedad virtual (Narula, 2022), el metaverso cuenta con el potencial de transformar la forma en que las personas viven, trabajan y se relacionan pudiendo dar lugar igualmente a nuevas formas de contar historias, con narrativas más personalizadas y adaptadas a los usuarios, acordes con los mundos virtuales que se generen.

Como rasgo clave de la sociedad contemporánea, hoy es el escenario de la sociedad de las plataformas aquel en el que se produce la difusión por excelencia de micronarrativas en línea, que tienen lugar tanto en el contexto de plataformas de redes sociales como de páginas web. El mayor grado de interactividad es quizás uno de los signos distintivos respecto a la sociedad red. En estos escenarios, la interacción cotidiana entre los ciudadanos y las ciudadanas es un estímulo a la producción continua de micronarrativas, algunas de las cuales, según explicaremos más adelante, son micronarrativas conspirativas y de odio.

### 3. LA ERA DE INTERNET, LOS PROSUMIDORES Y SUS MICRONARRATIVAS

La llamada Web 2.0 (Ackland, 2013), se distinguía de su predecesora (la Web 1.0) por tratarse de una red colaborativa. Es en el momento de tránsito de la Web 1.0 a la Web 2.0 cuando se desarrollan con más intensidad las “micronarrativas en línea”, como narrativas que igualmente van unidas a la interacción comunicativa que se produce en Internet. Igualmente, aunque en los tiempos de la llamada Web 1.0 se podían encontrar micronarrativas en Internet, la

posibilidad de la interacción en tiempo real entre los usuarios en Internet marca un hito significativo.

La Web 1.0 y la Web 2.0 se diferencian por la manera en la que los usuarios interactúan con el contenido. En la Web 1.0 los usuarios eran fundamentalmente consumidores de contenido en sitios web estáticos creados por profesionales en los que los usuarios tenían poco que decir en cuanto a la creación o modificación del contenido (por ejemplo, una web con noticias, o con bibliografía). En cambio, cuando se desarrolla la Web 2.0, los usuarios pueden tanto crear como consumir contenidos y se potencia la participación social en los sitios web (redes sociales, blogs, foros y otros). Esto supone un cambio radical respecto a las posibilidades de interacción cuando se inaugura la Web 2.0. Con ella se posibilita una interacción más fluida, bidireccional, la creación de comunidades en línea o incluso compartir información en tiempo real. Esta etapa se caracteriza por una participación más activa en la creación o modificación del contenido.

En cambio, en la primera fase el *prosumidor*, como usuario que es tanto productor como consumidor de contenidos (Ritzer, 2015; Ritzer y Jurgenson, 2010), aún no había emergido como actor clave de Internet. El prosumidor se sitúa en la Web 2.0. y es la acción de este la que aporta esencialmente un constante dinamismo en la web al crear y compartir contenido diverso como textos, imágenes, vídeos o audio. Los contenidos se comparten en línea a través de redes sociales y en otros sitios webs como foros o blogs. Es viable participar en línea de forma diversa, comentando noticias, conversando en las redes, apoyando una campaña, o de muchas otras formas. Este tipo de actividades se pueden hacer en lugares como Instagram, Twitter, Facebook, TikTok, Youtube, y otras plataformas o páginas web. El *prosumo* va unido a la faceta colaborativa en Internet. Esta faceta, la de que los consumidores son también productores de información en Internet, es clave para la multiplicación de las micronarrativas en línea, así como para el desarrollo del área de investigación basada en los macrodatos sociales.

De esta forma, las nuevas herramientas que se desarrollan con la Web 2.0, y las que se van introduciendo posteriormente hasta el día de hoy, permiten que cualquier usuario sin conocimientos técnicos de informática pueda crear, publicar, difundir o compartir contenidos a través de la web. Esto facilita la colaboración, el diálogo y la participación en Internet. Las nuevas herramientas que van desarrollándose para la interacción de los usuarios, cuya gama con el tiempo ha ido ampliándose y diversificándose, potencian la sociabilidad humana, como tendencia de los seres humanos a relacionarnos con otros, pero ahora también en escenarios virtuales. La publicación de micronarrativas o microrrelatos es clave en el transcurso de esta interacción, como una de las formas de comunicación más significativas de la sociedad de las plataformas. La producción y difusión de micronarrativas en línea es, por tanto, el resultado en

gran medida de la acción e interacción que se produce al amparo del avance de Internet. Las micronarrativas se nutren substancialmente de la interacción humana. No obstante, en sí mismas, como piezas comunicativas, no necesariamente implican siempre interactividad en el desarrollo de los relatos, si bien cuando las encontramos en línea, la interacción o la conversación suele ser un rasgo frecuente.

En este contexto, el de una sociedad en la que una parte importante de la comunicación se encuentra en plataformas de redes sociales, la población prosumidora, especialmente la más joven, se ha acostumbrado a la comunicación vehiculizada a través de formatos breves donde la comunicación se condensa eficazmente en menos palabras, sin que ello quiera decir que otras formas de comunicación mueran.

#### 4. MICRONARRATIVAS EN LÍNEA COMO CONSTRUCTORAS DE HISTORIAS: ASPECTOS FORMALES Y ESTRUCTURALES

Conforme los desarrollos tecnológicos vinculados a Internet se han hecho más profundos y se ha extendido el uso de los teléfonos móviles u otras herramientas que permiten la interactividad (ordenadores, tabletas, etc.), las micronarrativas en línea, fácilmente difundidas actualmente a través de plataformas de redes sociales o de páginas web (Twitter, Instagram, TikTok, Telegram, WhatsApp, Youtube u otras), se pueden entender cada vez más como una “forma contemporánea de comunicación” (Venditti et al., 2017, p. S273). Argumentamos a continuación, por añadidura, que las narrativas en las redes sociales funcionan como microrrelatos con características similares a otros tipos de microrrelatos (por ejemplo, en la literatura), pero se adaptan a las posibilidades que ofrece la tecnología actual y la web o plataforma particular (o red social) de la que se trate en cada momento. De esta forma, por ejemplo, un aspecto como la brevedad de los mensajes, es condicionado por los mismos límites que impone cada plataforma. Algunas características clave de las micronarrativas en línea son las siguientes:

##### 4.1. Brevedad

Las micronarrativas en línea son muy breves en comparación con otro tipo de narrativas con las que han trabajado en el pasado los científicos sociales, de ahí que las denominemos “micro” narrativas o que podamos referirnos a los “micro” relatos o “micro” historias que se encuentran comúnmente hoy en las redes sociales o en páginas web que permiten la interacción entre sus usuarios. Se pueden producir estas narrativas breves, por ejemplo, a través de comentar vídeos en Youtube, participando en un canal de televisión en línea

como Twitch, comentando alguna jugada o interactuando en un canal de deportes electrónicos en línea o haciendo una crítica breve a una película o una serie en IMDb. Cuando se trata de micronarrativas visuales, no solo de textos, esto se traduce en aspectos como la importancia que los vídeos muy cortos adquieren en redes sociales como TikTok, o que tienen los memes o los emojis en muchos mensajes.

#### **4.2. Formatos y contenidos atractivos**

Las micronarrativas comúnmente se construyen intentando contar con formatos y contenidos atractivos, especialmente cuando se han elaborado con cierta planificación. Se incluye aquí la gran frecuencia con la que los mensajes emplean imágenes [emojis, fotos, memes] o vídeos, o incluso GIFs animados. La importancia de que los contenidos puedan llegar a ser atractivos hace que en ocasiones estos sean diseñados o mejorados en su presentación por diseñadores profesionales, o incluso que se desplieguen herramientas específicas para el diseño de contenidos atractivos o sugerentes, al objeto de que puedan llamar la atención o impactar. Es frecuente que el diseño de micronarrativas atractivas responda a la estrategia de mantener la atención de los usuarios, o de mantener su motivación y su compromiso (Venditti *et al.*, 2017). En el caso, por ejemplo, de narrativas conspirativas o de narrativas de odio, no es infrecuente intentar seducir o convencer a los usuarios de redes sociales a través de narrativas muy atractivas para intentar ganar adeptos o seguidores.

#### **4.3. Difusión (o viralidad) que alcanzan los mensajes**

La enorme difusión (o viralidad) que pueden llegar a alcanzar algunos mensajes conecta directamente con la facilidad y rapidez con que cualquier mensaje puede ser compartido. El desarrollo técnico de los últimos años permite la difusión fácil de las micronarrativas a través de múltiples plataformas y de la multimodalidad. La capacidad de manejarse en escenarios “multiplataforma” posibilita que una plataforma de redes sociales (por ejemplo, Instagram), pueda funcionar en múltiples dispositivos (portátil, PC de sobremesa, tablet, teléfono) y diversos sistemas operativos (Windows, macOS y Linux), lo que proporciona grandes ventajas en cuanto a accesibilidad y usabilidad. Con la “multimodalidad” lo que ocurre es que una plataforma de redes sociales puede usar diversos formatos de contenido (texto, imágenes, vídeos, audio, código). Esto facilita que los usuarios puedan contar con diversas formas de expresión. Estas dos características también posibilitan la participación y la interacción. Estos factores, junto al abaratamiento de los dispositivos o de la conexión a Internet, permite a los usuarios acceder fácilmente a cualquier plataforma de redes sociales, en cualquier lugar y en cualquier momento. Como consecuencia, las

pautas de comunicación humana, personales y sociales, cambian, potenciando, entre otras cosas, que se multipliquen las micronarrativas en las redes sociales.

Entre las tecnologías que han potenciado los fenómenos de multiplataforma y multimodalidad en las plataformas de redes sociales se encuentran los dispositivos móviles, las redes 5G e incluso la inteligencia artificial, que se utiliza para mejorar la experiencia del usuario en las plataformas de redes sociales, por ejemplo, recomendando contenidos (el caso de TikTok). Ambos aspectos facilitan tremendamente la interconectividad, con lo que habiéndose democratizado el acceso a las tecnologías de la información y comunicación, y contando con que ya el 68 % (el 89,2 % en Europa) de la población mundial cuenta con Internet (Internet World Stats, 2023), así como, habiéndose desarrollado diversos elementos que potencian la conectividad, encontramos el campo abonado para la eclosión de micronarrativas en Internet, micronarrativas que cada vez más se han convertido en un objeto de estudio de interés para las ciencias sociales.

Un aspecto clave también de cara a una fácil propagación y un mayor impacto, es precisamente, de nuevo, la brevedad de estas narrativas, que facilita el hecho de ser compartidas y difundidas, más que otro tipo de narrativas que emplean formatos más extensos o pesados. De esta forma, un tuit que incluya un meme, una foto, un GIF, o un vídeo de un minuto es más fácil que sea compartido al clásico libro completo, película o documental. Además de la facilidad por el menor peso y la mayor rapidez para la transmisión, el ahorro en tiempo de lectura y visualizado son también importantes para lograr un mayor impacto.

#### **4.4. Empleo de recursos comunicativos y estilísticos**

Otra pauta que se puede considerar estructural en las micronarrativas en línea es la manera en la que se presentan al público los relatos, a saber, haciendo uso frecuentemente de estrategias y recursos comunicativos y estilísticos, empleando a veces pautas parecidas a las que encontramos en la retórica, si bien, adaptadas a las posibilidades que ofrecen las plataformas de comunicación donde se difunden estas micronarrativas. Entre los recursos estilísticos que se emplean en las narrativas a través del uso del lenguaje, se encuentran las figuras retóricas, los recursos gramaticales o los recursos de estilo. Este tipo de recursos pueden ayudar a crear un relato, un contexto, un clima o atmósfera. Pueden ayudar a conseguir algún tipo de efecto que se desee. Los recursos comunicativos son una pieza clave para que se comprenda la historia que se relata y se usan para transmitir la información. Tenemos aquí desde la elección de las palabras adecuadas para expresar las ideas, hasta la sintaxis o la manera en que se estructuran las frases, la puntuación o el vocabulario mismo.

#### 4.4.1. *Los tropos y la retórica: de las metáforas a la ironía*

Merece la pena una mención a los tropos y a la retórica, como recursos estilísticos comunes que se encuentran frecuentemente en las micronarrativas en línea. Los tropos son figuras retóricas que usan conceptos de forma figurada para sugerir una comparación o una analogía. Pueden ser muy dañinos en cuanto a las analogías que elaboran. Son muy usados en el discurso social y político y muy comunes en las micronarrativas que se encuentran en las redes sociales. Son una práctica común en la retórica. Han sido muy usados, por ejemplo, al potenciarse la creatividad léxica durante la pandemia (Rodríguez González, 2022), escenario de donde tomamos algunos ejemplos.

Citamos a continuación algunos ejemplos de tropos que se han encontrado en las narrativas de los movimientos negacionistas, antimascarillas, anticuarentena o antivacunas en la pandemia de COVID-19 y que se difunden al objeto de crear confusión o duda, de difundir desinformación, manipular o de presentar oposición al gobierno<sup>2</sup>. La metáfora y la metonimia son centrales en los estudios sobre el discurso (Soledad, 2021). La metáfora se produce cuando se comparan dos cosas que no son parecidas de forma literal. Por ejemplo, cuando se asocia metafóricamente la vacunación con un veneno, se produce alarma en la población. Cuando se asocia la mascarilla con un bozal, de nuevo la metáfora produce información engañosa e incluso despectiva. Siguiendo el ejemplo, frecuentemente, el tono emocional es intenso, notándose la ira o el cabreo hacia los gobiernos:

«#YoNoMeVacuno para eso crearon esta falsa Pandemia. Para obligarnos a meternos ese veneno».

«Eso es Bioterrorismo, infundir Pánicos mentales para que la gente no se quite el bozal, no ande en la calle y los gobiernos sumisos a la China Comunista vuelvan a violentar las Libertades Fundamentales. #Bandidos!!».

En el caso de la metonimia, que es una figura que sustituye una cosa por otra con la que está relacionada, en el siguiente ejemplo el término “vacuna” se usa como metonimia para referirse a los efectos secundarios de las vacunas:

«Mentira. La vacuna mata. #YoNoMevacuno los abuelos no deben inocularse; mientras más se pichan más esparcen el virus! Estudien investiguen obtusos!».

El sarcasmo o la ironía son igualmente figuras muy usadas que pretenden decir lo contrario a lo que se quiere decir. En el caso del sarcasmo se pretende

<sup>2</sup> Los ejemplos de tuits mostrados en este capítulo son tuits literales anonimizados extraídos del *Conspiracy Theories Dataset, 2020-2023 [Twitter]* del proyecto NON-CONSPIRA-HATE en el que se enmarca este libro [<https://eseis.es/investigacion/discursos-de-odio>]

expresar desagrado o burla. La ironía, en cambio, es una figura retórica que consiste en decir algo que tiene el sentido opuesto al que se quiere expresar. En el ejemplo que sigue la palabra “GRACIAS” tiene un sentido irónico:

«Así funciona la DICTADURA GLOBAL por la que hoy tocáis las palmas en los balcones ¿Cómo lo imaginábais? ¿Con un tío bigotón entrando pegando tiros? NO, IGNORANTES! DARÉIS GRACIAS POR ELLA #Plandemia #YoNoSoyZombie <https://t.co/qO5VYwAhq7>».

Igualmente, figuras como la hipérbole, en tanto que exageración, son muy comunes para magnificar algunos aspectos, como en el ejemplo que sigue, donde se exagera la peligrosidad de las vacunas y se las asocia a la muerte. El uso de las mayúsculas al escribir, igualmente, refuerza la exageración:

«EL AUTISMO LLEGÓ A VIETNAM DE LA MANO DE LAS “”VACUNAS”” DE BILL GATES. MUERTE, ESTERILIDAD Y ENFERMEDADES SON EL OBJETIVO DE LAS VACUNAS PARA CONTROLAR LA DEMOGRAFÍA POR PARTE DE ESTOS PSICÓPATAS EUGENISTAS #plandemia #circovid #genocidio #miedovirus #coronatimo #coronafake <https://t.co/aPXRfMRFXZ>».

A través de estos ejemplos, u otros que podrían citarse, incluyendo otras figuras retóricas, se aprecia el importante papel que las micronarrativas en línea pueden tener en la conformación de los estados de opinión ante temas sensibles.

#### 4.4.2. *Las etiquetas (#): Elementos clave en algunas narrativas que se difunden en Internet*

Los hashtags o etiquetas son quizás uno de los elementos por excelencia en las micronarrativas que se difunden en las plataformas de redes sociales. El símbolo almohadilla (#), que antecede a una palabra o combinación de caracteres, los identifica claramente. Una etiqueta puede ser utilizada para diferentes fines como para agregar contexto a una micronarrativa necesitando para ello muy pocos caracteres (#8M). También se puede usar con la idea de clasificar la temática sobre la que versa un tuit concreto (Congosto, 2016) e incluso crear una conversación organizada a través de esta etiqueta (por ejemplo, como estrategia para hacer el seguimiento de un evento, conferencia, concurso, etcétera). Los hashtags tienen la función de marcar la temática e identificarla, lo que facilita la búsqueda de tuits sobre ese tema por parte de otros usuarios. Se emplean también para poder hacer una búsqueda sobre las noticias más recientes que han sido clasificadas bajo esa temática. La etiqueta incluso puede crear algún tipo de efecto (estilístico) en la narrativa.

Un uso adecuado de las etiquetas puede conseguir un mayor alcance o propagación de los mensajes. Un aspecto importante también, con respecto al uso social de estas etiquetas, es que se emplean a veces de forma coordinada y colaborativa (por ejemplo, en campañas sociales, con el objetivo de promocionar a políticos, de impulsar empresas, o incluso como eslogan) con el objetivo de conseguir tendencias. Este etiquetado social (o *social tagging*) es un rasgo igualmente común de las micronarrativas en línea. Curiosamente, los hashtags se originan en 2007, a partir de la propuesta de Chris Messina (2007, [<https://chrismessina.me/>]) en Twitter cuando sugirió utilizar el símbolo de la almohadilla (#) para etiquetar los mensajes, de forma que se pudieran organizar y buscar fácilmente. La práctica de incorporar etiquetas en los mensajes fue adoptada con rapidez en Twitter y posteriormente en otras redes y se ha convertido en una práctica estándar. Algunos ejemplos de etiquetas relacionadas con micronarrativas conspirativas y de odio relativas al COVID-19 son #circovid, #plandemia, #repentinitis o #sideeffects.

#### 4.4.3. *Emojis, símbolos, memes, imágenes o vídeos: Buscando una mayor expresividad y emotividad de las micronarrativas*

Sin detenernos extensamente, no puede pasarse de largo otra pauta común en la comunicación actual que se produce en las redes sociales, a saber, la inclusión de emojis, imágenes, vídeos, memes, caricaturas, o similares que pueden ayudar a realzar la expresividad y emotividad de los mensajes, convirtiendo la comunicación en algo más estimulante y atractivo.

Aunque los resultados de la investigación no son totalmente conclusivos, se argumenta sobre “el poder de los memes” en diferentes escenarios, desde el político, social, cultural o incluso en el plano personal o psicológico. Así, en algunos trabajos que se detienen en cuestiones psicológicas, se muestra que los memes pueden ayudar a manejar y mitigar aspectos estresantes, como en el caso de la pandemia (Myrick, 2022), e incluso, ya desde una perspectiva cultural, se plantea cómo los memes pueden contribuir a reinterpretar situaciones (Flecha et al., 2021). El mayor atractivo visual y quizás la mayor facilidad para ser compartidos, incluyendo el hecho de que en mayor medida provocan respuestas emocionales (frecuentemente asociadas al humor) hacen que los memes sean un instrumento clave que se haya usado con frecuencia para la difusión de narrativas conspirativas y de odio, así como basadas en la desinformación (Rodríguez Ponce, 2022; Al-Rawi, 2021). En estos escenarios, además, igual que ocurre con los “emojis” u otras expresiones visuales, es importante dedicar una atención particular a los contextos en los que se emplean los memes (De Jong et al., 2022). Igualmente, las micronarrativas en línea pueden ser de utilidad para contrarrestar las narrativas conspirativas y de odio a través de la difusión

de “contra memes” (u otro tipo de imágenes) que pueden ayudar a combatir la desinformación existente (Astapova, 2022). A estos efectos, véase el capítulo de Santos, Ruiz y Ruiz en este mismo libro.

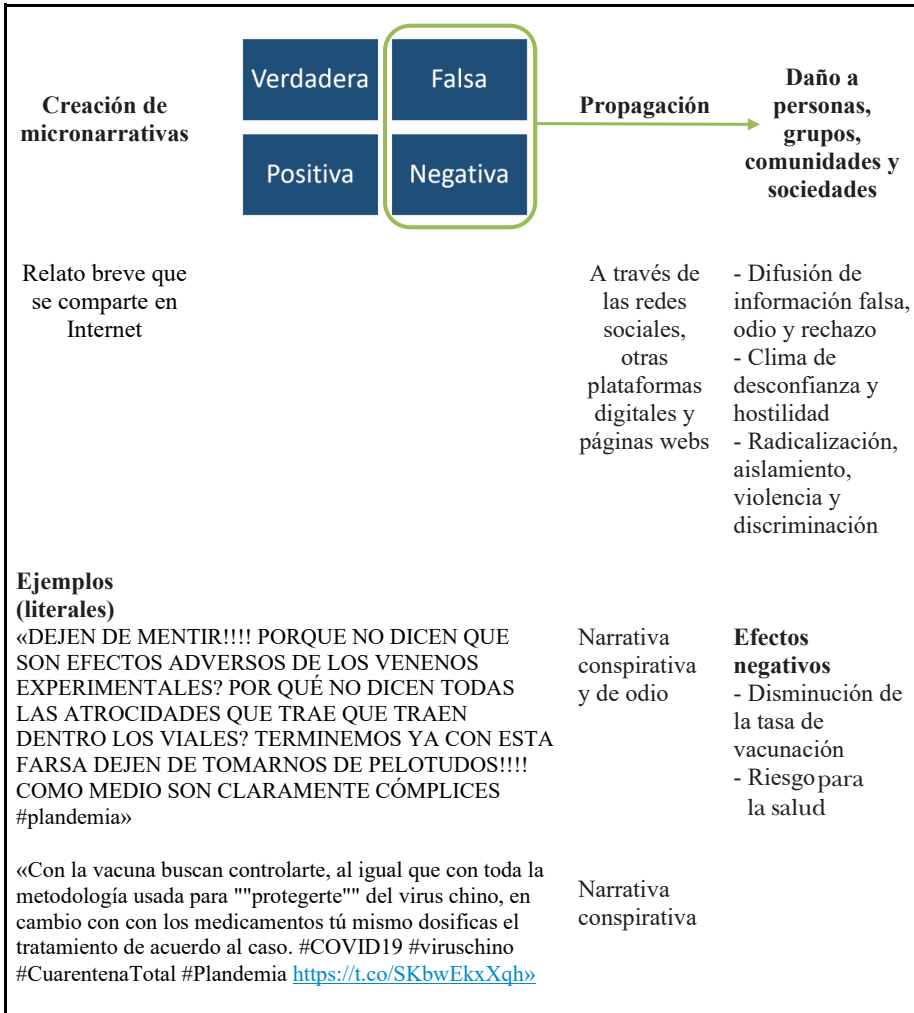
## 5. LAS MICRONARRATIVAS EN LÍNEA COMO VEHÍCULOS DE TEORÍAS DE LA CONSPIRACIÓN Y DISCURSOS DE ODIOS: EL CASO DE LA COVID-19, INMIGRANTES, REFUGIADOS Y PERSONAS LGBTIQ+

### 5.1. Micronarrativas en línea, conspirativas y de odio

Aunque en las páginas precedentes se han ido presentando algunos ejemplos concretos sobre micronarrativas en línea relacionadas con la pandemia de COVID-19, y en el apartado previo hemos abordado algunos aspectos estructurales referidos a la forma o el formato de las mismas (la forma), en este apartado nos referimos brevemente a su contenido (el fondo) y a las consecuencias negativas que generan en las sociedades actuales, para seguidamente, esbozar unas conclusiones y recomendaciones breves. Igualmente, visibilizamos en este apartado cómo conectan las micronarrativas en línea con las narrativas conspirativas y de odio que se encuentran en Internet (Figura 1).

Como se sugiere en la Figura 1, y se evidencia en los ejemplos tomados de Twitter, en el contexto de la sociedad de las plataformas, las micronarrativas en línea son un problema en crecimiento y una amenaza o riesgo en ocasiones, instrumentalizadas frecuentemente por ideologías radicales. Los impactos de micronarrativas conspirativas y de odio en línea pueden hacer daño tanto a nivel micro, como meso y macro, afectando a personas, grupos o comunidades y a sociedades. Esto se debe en gran medida a que estas micronarrativas son muy eficaces para difundir mensajes conspirativos, de odio, despectivos, discriminatorios o violentos hacia diferentes destinatarios o en diferentes circunstancias o contextos, utilizando a veces recursos muy eficaces para la propagación de sus mensajes, como se ha mencionado arriba. Igualmente, las ambigüedades y dobles sentidos, además de otros factores que se encuentran en los mensajes, hacen que a veces sea complicado que estos sean detectados automáticamente por los algoritmos, como da cuenta la bibliografía especializada (Alkomah & Ma, 2022; Mullah & Zainon, 2021; Shahsavari et al., 2020; MacAvaney et al., 2019). Si se encuentran dificultades para ser detectados, las habrá igualmente para bloquear, prohibir o denunciar estos mensajes.

Figura 1. Conexión entre las microrrelatos en línea, conspirativos y de odio



Fuente: elaboración propia. Los ejemplos proceden del conjunto de datos mencionado en la nota a pie 2.

### 5.2. Microrrelatos de odio en línea y teorías de la conspiración: COVID-19, inmigrantes y refugiados y personas LGBTIQ+

Con el desarrollo de Internet y la participación en redes sociales, es frecuente encontrar diversidad de narrativas de odio en línea, así como la difusión de teorías de la conspiración y desinformación relativas a inmigrantes y refugiados. Las teorías de la conspiración se han extendido intensamente en los últimos años, ayudadas por el altavoz que proporcionan diversos medios y redes sociales

como Twitter, Facebook, Instagram, TikTok y otros (Önnerfors y Krouwel, eds., 2021; Gualda, 2020) a través de los que circulan continuamente micronarrativas en línea que las sustentan, como se ha descrito en páginas precedentes.

Igualmente, instituciones internacionales y europeas han dado la voz de alarma sobre cómo se han desplegado expresiones de odio e intolerancia en estos años dirigidas a poblaciones como las de refugiados, inmigrantes, así como personas LGTBQ+ (UN Secretariat, 2020; European Commission against Racism and Intolerance, 2021; Consejo de Europa, 2021), si atendemos a la investigación reciente sobre los destinatarios de los discursos de odio y teorías de la conspiración que conectan con el proyecto que da pie a este capítulo [<https://eseis.es/investigacion/discursos-de-odio>]. Se trata de un odio que se difunde de forma rápida y frecuentemente de forma anónima e impune. Se ha documentado también que desde el inicio de la pandemia la población asiática, los refugiados, los inmigrantes, los gitanos, las mujeres, los judíos y las personas LGTBQ+, entre otros, fueron blanco de expresiones de odio, en ocasiones culpándoles del inicio de la pandemia y de facilitar su propagación (Comisión Europea contra el Racismo y la Intolerancia, 2021; ILGA-Europa, 2021, ILGA-Mundo, 2020; Naciones Unidas, 2021).

La retórica del odio en los medios sociales y, en ocasiones, la violencia o los incidentes y delitos de odio en la vida cotidiana, junto a la desinformación que circula sustentando algunas teorías de la conspiración, agudizan la vulnerabilidad y la privación de derechos de algunas minorías debido a los componentes de racismo, islamofobia, antisemitismo, homofobia y transfobia de muchos mensajes que circulan tanto en la calle (*offline*) como por las redes (*online*), reforzándose entre sí. Los mensajes negativos, a veces muy sutiles, propagan estereotipos, prejuicios, información falsa o distorsionada que contribuyen a la construcción y profundización en los procesos de estigmatización de personas y grupos vulnerables.

Igualmente, como suele ocurrir en otros periodos de crisis e incertidumbre, la pandemia de COVID-19 funcionó como catalizador para el florecimiento de teorías conspirativas que pueden hacer un importante daño a las sociedades (Douglas, 2021) y en las que se refleja, en gran medida, la orientación política de las personas que las suscriben (Imhoff et al., 2022). Las teorías conspirativas suelen caracterizarse por apelar a narrativas extremistas, difundiendo discursos de odio (por ejemplo, identificando a un grupo o persona como culpable) (Farinelli, 2021). También, se materializan a través de narrativas basadas en viejas narrativas y teorías conspirativas impulsadas por movimientos de extrema derecha y expresiones populistas en la actualidad (Bergman, 2018). Por otra parte, los grupos de extrema derecha han contribuido durante la pandemia a multiplicar el odio. El racismo y la xenofobia se han desarrollado a través de narrativas de odio y desinformación contra musulmanes, inmigrantes y refu-

giados, por ejemplo, desplegando la teoría de la conspiración de la invasión (Gualda, 2021). Las viejas narrativas se reformulan y actualizan en el escenario de Internet, propagándose en muchas ocasiones con apoyo de micronarrativas. La facilidad con la que en diferentes plataformas se encuentran estos relatos (vehiculizados a través de soportes como el texto, las fotos o los vídeos), hace necesario el estudio científico sistemático y específico de los mensajes desarrollados durante la pandemia (análisis retrospectivo) y en la actualidad para conocer en profundidad los mecanismos mediante los cuales se difunden las teorías de la conspiración y los discursos de odio.

## 6. A MODO DE CONCLUSIÓN: APUNTES PARA CONTRARRESTAR LAS MICRONARRATIVAS NOCIVAS Y FRAUDULENTAS EN UN CONTEXTO DE AUGE DE LA INTELIGENCIA ARTIFICIAL [IA]

Las micronarrativas en línea, comunes en la sociedad de las plataformas y una de las pautas clave de la comunicación en el siglo XXI, son usadas para difundir información errónea y dañina pudiendo alcanzar mayor (y más rápida) propagación que algunas narrativas clásicas y más extensas en palabras que circulaban fuera de línea. Estas narrativas se han convertido en un formato muy eficaz para difundir las teorías de la conspiración y los discursos de odio en diversidad de plataformas de redes sociales y páginas web. En este contexto, entender cómo funcionan y operan, así como saber cuáles son algunos de sus rasgos prevalentes en la difusión de las teorías de la conspiración y del odio en línea, es clave. Es importante también que, ante cualquier indicio, antes de creerlas o propagarlas, se hagan los esfuerzos pertinentes para identificarlas, adoptando una mirada crítica hacia la información que se recibe y se comparte en línea. Hay que acometer esfuerzos tales a revisar las fuentes de donde procede la información, valorar si los datos que se presentan son sólidos y verificables, o si los autores pueden contar con sesgos o prejuicios, entre otras estrategias.

Una vez que se identifica una narrativa conspirativa o de odio en línea, es importante no compartirla. Igualmente, se pueden denunciar a través de la plataforma o en otras instancias, como acto de responsabilidad que puede contribuir a construir espacios virtuales más seguros y respetuosos. Al hilo de esto, sería crucial contar con estrategias que permitan reducir las capacidades de los actores que distorsionan y radicalizan el discurso en línea.

Es clave hoy también no restar importancia a la inteligencia artificial (IA), que puede emplearse para crear micronarrativas en línea más virulentas que las que los propios humanos generan. Esto se debe a la capacidad de la IA para personalizar estas micronarrativas y que puedan ser, por una parte, más convincentes y, por otra, más difícil de ser detectadas o de verificar que son erróneas. La capacidad de la IA para simular textos o imágenes, como los que realizan

los humanos, dificulta la detección de las narrativas que no son auténticas. Además, la IA puede usarse para difundir estas micronarrativas de forma rápida y eficiente a través de las redes sociales, otras plataformas digitales y páginas web, por lo que pueden ser más difíciles de controlar.

Con estas características, parte de la agenda de investigación relativa a las micronarrativas en línea para los próximos años pasa necesariamente por estar atentos a cómo se pueda usar la IA para difundir micronarrativas conspirativas y de odio, fundamentándose en información manipulada, falsa o engañosa. Igualmente, será de interés estudiar si se usa la IA (y cómo) para manipular a las personas, o incluso si afecta de alguna forma a la generación de mensajes que contribuyan a crear un clima de odio o división que conduzca a la violencia o la discriminación. Sería, por tanto, clave que los avances en la inteligencia artificial en los próximos años puedan ponerse más al servicio de los retos sociales actuales, que a la profundización y desarrollo de sus efectos nocivos. Con la experiencia histórica de que los humanos usamos la tecnología a veces para hacer el bien, a veces para hacer el mal, se hace evidente también la necesidad por parte de las instituciones públicas de fortalecer los sistemas de regulación, supervisión y establecimiento de medidas disuasorias y sanciones destinadas a los actores responsables de la generación y propagación de micronarrativas conspirativas y de odio, actores que utilizan la desinformación en su propio interés, ocasionando daños a personas, grupos, comunidades y sociedades.

## REFERENCIAS

- Ackland, R. (2013). *Web Social Science. Concepts, Data and Tools for Social Scientists in the Digital Age*. Sage.
- Alkomah, F. & Ma, X. (2022). A literature review of textual hate speech detection methods and datasets. *Information*, 13(6), 273. <https://doi.org/10.3390/info13060273>
- Al-Rawi, A. (2021). Political Memes and Fake News Discourses on Instagram. *Media and Communication*, 9(1), 276-290. <https://doi.org/10.17645/mac.v9i1.3533>
- Astapova, A. (2022). Vaccine hesitancy counter memes. *Journal of American Folklore*, 135(538), 466-470. <https://www.muse.jhu.edu/article/866493>
- Bauman, Z. (2000). *Liquid Modernity*. Polity Press.
- Bell, D. (1976). *El advenimiento de la sociedad postindustrial*. Alianza.
- Bergmann, E. (2018). *Conspiracy & Populism: The Politics of Misinformation*. Palgrave MacMillan.
- Castells, M. (1997). *La Era de la Información: Economía, Sociedad y Cultura. Vol. I: La sociedad red*. Alianza.

- Congosto Martínez, M. (2016). *Caracterización de usuarios y propagación de mensajes en Twitter en el entorno de temas sociales*. Tesis Doctoral defendida en la Universidad Carlos III de Madrid. <http://hdl.handle.net/10016/22826>.
- Cuevas-Molano, E., Sánchez-Cid, M. & Gordo-Molina, V. (2022). Brand strategy and content management on Instagram: scheduling and message length as factors to improve engagement. *Communication & Society*, 35(2), 71-87.
- De Jong, S., Kowalchuk, S. & McKelvey, F. (2022). Who to blame this pandemic on: A qualitative study of the politicization of COVID-19 through political memes in Canada. *Global Media Journal*, 14(1), 5-27.  
[http://gmj-canadianedition.ca/wp-content/uploads/2023/01/02\\_DeJongKowalchuk-Fenwick-Volume-14-issue-1\\_Paper-FINAL.pdf](http://gmj-canadianedition.ca/wp-content/uploads/2023/01/02_DeJongKowalchuk-Fenwick-Volume-14-issue-1_Paper-FINAL.pdf)
- Douglas, K.M. (2021). Are Conspiracy Theories Harmless?. *The Spanish Journal of Psychology*, 24, E13. <http://dx.doi.org/10.1017/SJP.2021.10>
- European Commission against Racism and Intolerance [ECRI] (2021). *Annual Report on ECRI's Activities covering the period from 1 January to 31 December 2020*. <https://rm.coe.int/annual-report-on-ecri-s-activities-for-2020/1680a1cd59>
- Farinelli, F. (2021). *Conspiracy theories and right-wing extremism - Insights and recommendations for P/CVE*. Radicalisation Awareness Network. European Commission. [https://home-affairs.ec.europa.eu/system/files/2021-04/ran\\_conspiracy\\_theories\\_and\\_right-wing\\_2021\\_en.pdf](https://home-affairs.ec.europa.eu/system/files/2021-04/ran_conspiracy_theories_and_right-wing_2021_en.pdf)
- Flecha Ortiz, J.A., Santos Corrada, M.A., Lopez, E., Dones, V. (2021). Analysis of the use of memes as an exponent of collective coping during COVID-19 in Puerto Rico. *Media International Australia*, 178(1),168-181. <https://www.doi.org/10.1177/1329878X20966379>
- Gkikas, D., Tzafilkou, K., Theodoridis, P., Garmpis, A. & Gkikas, M. (2022). How do text characteristics impact user engagement in social media posts: Modeling content readability, length, and hashtags number in Facebook. *International Journal of Information Management*, 2(1), 100067. <https://www.doi.org/10.1016/j.ijime.2022.100067>
- Gualda, E. (2018). Medios de comunicación, medios sociales y análisis de redes sociales. En Iglesias, J.; Trinidad, A. y Soriano, R. (coord.). *La Sociedad desde la Sociología. Una introducción a la Sociología General*. Tecnos, 581-604.
- Gualda, E. (2021). Metaphors of invasion: Imagining Europe as endangered by Islamization. En Önnfors, A. & Krouwel, A. (eds.). *Europe: Continent of Conspiracies. Conspiracy Theories in and about Europe*. Routledge, 54-75.
- Gualda, E. (2020). Social network analysis, social big data and conspiracy theories. En Butter, M. & Knight, P. (ed.). *Handbook of Conspiracy Theories*. Routledge, 135-147.
- ILGA-Europe (2021). *COVID-19 and specific impact on LGBTI people and what authorities should be doing to mitigate impact*. [https://www.ilga-europe.org/sites/default/files/COVID19%20\\_Impact%20LGBTI%20people.pdf](https://www.ilga-europe.org/sites/default/files/COVID19%20_Impact%20LGBTI%20people.pdf)

- Imhoff, R., Zimmer, F., Klein, O. *et al.* (2022). Conspiracy mentality and political orientation across 26 countries. *Nature Human Behaviour*, 6, 392-403. <https://doi.org/10.1038/s41562-021-01258-7>
- Internet World Stats (2023). World Internet Users and 2023 Population Stats. *Internet World Stats. Usage and Population Statistics*. <https://www.internetworldstats.com/stats.htm>
- Lyon, D. (1994). *The electronic eye: The rise of surveillance society*. University of Minnesota Press.
- Lyon, D. (2018). *Liquid Surveillance: The Emergence of a New Paradigm*. Polity Press.
- Lyon, D. & Bauman, Z. (2019). *Vigilance Capitalism: The Fight for a Human Future in the Age of Big Data*. Polity Press.
- MacAvaney, S., *et al.* (2019). Hate speech detection: Challenges and solutions. *PLoS ONE*, 14 (8): e0221152. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0221152>
- Masuda, Y. (1980). *The Information Society as Post-Industrial Society*. World Future Society.
- Messina, C. (2007, 23 de agosto). How do you feel about using # for groups. As in #barcamp [mensaje de Twitter]. Twitter. <https://twitter.com/chrismessina/status/223115412>
- Mullah, N. S. & Zainon, W. M. N. W. (2021). Advances in Machine Learning Algorithms for Hate Speech Detection in Social Media: A Review. *IEEE Access*, vol. 9, 88364-88376. <https://www.doi.org/10.1109/ACCESS.2021.3089515>
- Myrick, J. G., Nabi, R. L. & Eng, N. J. (2022). Consuming memes during the COVID pandemic: Effects of memes and meme type on COVID-related stress and coping efficacy. *Psychology of Popular Media*, 11(3), 316-323. <https://doi.org/10.1037/ppm0000371>
- Naisbitt, J. (1982). *Megatrends: Ten New Directions Transforming Our Lives*. Warner Books.
- Narula, H. (2022). *Virtual Society: The Metaverse and the New Frontiers of Human Experience*. Penguin Business.
- Önnefors, A. & Krouwel, A. (eds.) (2021). *Europe: Continent of Conspiracies. Conspiracy Theories in and about Europe*. Routledge.
- Ritzer, G. (1996). The McDonaldization Thesis: Is expansion inevitable? *International Sociology*, 11(3), 291-308.
- Ritzer, G. (2015). Prosumer Capitalism. *The Sociological Quarterly*, 56, 413-445. <https://doi.org/10.1111/tsq.12105>
- Ritzer, G. & Jurgenson, N. (2010). Production, Consumption, Prosumption. The nature of capitalism in the age of the digital ‘prosumer’. *Journal of Consumer Culture*, 10(1), 13-36. <https://www.doi.org/10.1177/1469540509354673>
- Rodríguez González, F. (2022). Ironía y creatividad léxica durante la pandemia del coronavirus. *Lexis*, 46(2), 587-620. <https://dx.doi.org/10.18800/lexis.202202.005>

- Rodríguez Ponce, M. I. (2022). TERF: ¿Insulto o meme?: El sesgo ideológico en el análisis crítico del discurso. *Verba Hispanica*, 30(1), 173-193. <https://doi.org/10.4312/vh.30.1.173-193>
- Shahbaznezhad, H., Dolan, R. & Rashidirad, M. (2021). The Role of Social Media Content Format and Platform in Users' Engagement Behavior. *Journal of Interactive Marketing*, 53(1), 47-65. <https://doi.org/10.1016/j.intmar.2020.05.001>
- Shahsavari, S., Holur, P., Wang, T., Tangherlini, T. R. & Roychowdhury, V. (2020). Conspiracy in the time of corona: Automatic detection of emerging COVID-19 conspiracy theories in social media and the news. *Journal of Computational Social Science*, 3(2), 279-317. <https://doi.org/10.1007/s42001-020-00086-5>
- Soledad González, D. (2021). Metáforas y metonimias de la pandemia de COVID-19 y la cuarentena. *Semas*, 2(3), 89-112. <https://semas.uaq.mx/index.php/ojs/article/view/42>
- Toffler, A. (1980). *The Third Wave*. Bantam Books.
- UN Secretariat (2020). United Nations A/75/329 General Assembly. Seventy-fifth session Item 70 (a) of the provisional agenda. Elimination of racism, racial discrimination, xenophobia and related intolerance. <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N20/222/82/PDF/N2022282.pdf?OpenElement>
- Venditti, S., Piredda, F. y Mattana, W. (2017). Micronarratives as the form of contemporary communication. *The Design Journal*, 20, sup1, S273-S282. <https://www.doi.org/10.1080/14606925.2017.1352804>
- Van Dijck, J., Poell, T. & De Waal, M. (2018). *The Platform Society: Public Values in a Connective World*. Oxford Academic Press.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

# HATE SPEECH. A BIRD'S EYE VIEW OF A CONUNDRUM OF INTERNATIONAL EPIDEMIC PROPORTIONS

EMILIA E. MARTINEZ-BRAWLEY

*Professor. Watts College of Public Programs and Community Solutions  
Arizona State University, USA*

**DOI:** 10.14679/2751

## **Summary:**

1. Introduction.
  2. Cyberspace, the Internet and the broad network of electronic communication.
  3. Socio-cultural and historical considerations.
  4. The historical problem of definitions.
  5. The epidemic nature of hate speech via the Internet. Two salient facilitating factors.
    - 5.1. Manipulation.
    - 5.2. The contagion effect of on-line communication.
  6. Some concluding reflections.
- References.

## 1. INTRODUCTION

Many have asked whether hate speech like smallpox, or polio, or Ebola or Covid and even some mental illnesses, could be readily identified across different cultures, languages and historical situations. Hate speech has continued to be an international challenge since WWII even before the complexities of the internet made it epidemic. The concerns about combating hate speech, when many tools have failed, are real and continue to escalate. Consequently, it is not surprising that the academy today is interested in expanding the already existing sources of information to further the understanding of the phenomenon and solutions. In addition to expected entities such as the U.N., the European Union, the Council of Europe (through various commissions) and individual governments of member states, academic endeavors and NGO's, whether legal, political, religious or social services, have joined the quest for explanations. It is clear that the Internet constitutes a puzzle today and that anyone writing about "hate speech" and particularly "hate speech on the internet" must carefully assess the legitimacy of sources accessed, because even good judgement and caution- which this writing will strongly recommend (Martinez-Brawley, 2022;

Howard, 2019) cannot avoid the occasional bad choice. There is, in short, a plethora of choices for where to turn, but if one thing can be said in a bird's eye view of this conundrum is that different sources often contradict each other and differ in definitions and points of view. This chapter will offer a bird's eye view of the complexity of developing a narrative on the matter.

## 2. CYBERSPACE, THE INTERNET AND THE BROAD NETWORK OF ELECTRONIC COMMUNICATION

Let us begin with some basic clarifications. In a chapter entitled *Cyberspace and Social Media*, Cogburn (2016) offers some good explanations to begin our discussion:

«Cyberspace is a term that emerged in 1980s science fiction, and in popular culture in the 1990s. The term “cyberspace” generally refers to all of the spaces online that use the Internet, over fixed and mobile communication networks, including the World Wide Web and all the popular social media and other applications that use this infrastructure» (Cogburn, 2016, p.252).

The same author suggests that the internet can be described more narrowly because it operates as

«[...] a globally distributed “network of networks,” which allows social media and all of cyberspace to exist. Even this narrow focus is difficult [to define]. We get pulled between the highly technical and specific aspects of developing software and hardware protocols; and the broader policy-oriented aspects of managing [...] the Internet and its numerous applications and implications (Cogburn, 2016, p.252)».

With these basic distinctions, which are by no means “universal” but offer a start, this writer has approached the subject of hate speech in the internet, not as an expert but as a social-philosopher, if you will, someone who wanted to make sense of the incredible noise that surrounds the topic today. The timeframe for the preparation of this narrative was brief, therefore these reflections are based largely on material now in print and in the internet, which provided a functional start. Functionally, the internet and cyberspace can be, as the readers will know, very useful tools.

## 3. SOCIO-CULTURAL AND HISTORICAL CONSIDERATIONS

Before singling out the internet in particular and cyberspace in general, as the primary causes of the epidemic of “hate speech”, it is worth considering some historical, socio-cultural and psychological explanations, which seem to make

sense, at least in the U.S. and probably some other Western countries. David Brooks (2023), a respected cultural critic, wrote a series of editorials in the NY Times attempting to look into our current universal cultural crisis as a consequence of social developments. In general terms, his choice of cultural paradigms governing our social behavior for the past forty or so years has some validity. Brooks believes that past traditions that offered a connection to meaning and identity have been eroded. Writers have produced many works on the loss of community and the sporadic efforts to regain its meaning (Martinez-Brawley, 1990 & 2000; Nisbet, 1953; Roszak, 1978; Putnam, 1995 & 2000) and many others. Quoting Lasch (1979) and other philosophers, Brooks (2023) suggests that the current culture of narcissism and obsession with trauma, particularly among professionals, has generously added to a mental health crisis, with rising depression and suicide. Trauma, personal and collective, causes hopelessness, and in spite of the over “coddling” of the young against real and perceived enemies (Brooks, 2023), the result was often self-hate. According to Brooks, the young in particular have relied on a constant admiring audience and total reliance on others to validate their still developing sense of self. A popular author and storyteller on television and radio in the U.S., created an imaginary place called Lake Wobegon (Keillor, 1985&1997) where “all the women are beautiful and all the children above average”. For other reasons, he became at some point *persona non-grata* in the television world, but his humorous words and folk style had already developed a believable scenario and his observations spread through the media.

Another important historical dynamic that also points at the internet and social media is the expression of hate by members of majority and/or dominant groups, which reflect different contexts in various parts of the world. The assertive stances that all minority and immigrant, women, and LGBTQ groups almost unquestionably achieved throughout the world in the past 50 years or so, has unleashed not only awareness of rights, but also active opposition from the majority. For many decades, members of majority and other powerful groups had managed to suppress hatred and negativism towards those who challenged them. In the U.S., for example, the legacy of slavery and colonialism that had laid dormant underneath a cover of equality and opportunity came to the fore. After Civil Rights movements in the 1960, 70s and 80s, African Americans demanded equal treatment, social and political justice and the benefits of the democratic society. American Indians became involved in litigation about the lands that had been taken from them through the decades. Mexican Americans and different Indian tribes who had been the original inhabitants in vast portions of the territorial U.S. reactivated their claims. While not all majority groups had fostered messages of hate—which had become world-known through the Ku Klux Klan in decades past, in the U.S., fanaticism proliferated and such messages could in the new millennium be disseminated easily and become aggrandized in the process (SPLC, 2019, 2022 & 2023).

Similar events were occurring internationally. In the Netherlands, for another example, in 2004, Dutch film maker Theo Van Gogh “was murdered for supposedly having offended the prophet Mohammed in his movie, *Submission*, which dealt with the suppression of women in Islamic cultures.” (Afshin and Molier, eds., 2015). Another example in Paris in 2015, two jihadists ordered by Al-Qaeda, attacked the office of the satirical magazine Charlie Hebdo, where 10 editors and a policeman, who was of Islamic background, were killed by sympathizers with the Islamic state. The same year, a grocery store was attacked and four Jews were killed (Afshin and Molier, eds., 2015). These events highlighted the presence of fanatical groups in Europe and resulted in heated debates about what constitutes hate speech and what can be sanctioned under the Dutch penal code (p.1) or the measures existing in other countries. However, the contradictions of cyberspace were highlighted by the same authors, who suggested that the Arab Spring of 2011 “would have been nearly inconceivable” without the broad coverage provided by the networks broadcasting from Tahir Square in Cairo. Furthermore, they also concluded that

«The social media have been important in contributing to the protests in the Arab world and the realization of the Iranian Green Movement in the Summer of 2009. The tyrants realize all too well wherein the dangers of liberty lie: the exercise of power would be exposed to public control» (Afshin and Molier, eds., 2015, p. 3).

The duality of the internet in being helpful and yet destructive appeared by the early 2020’s to be a point of international agreement (Martinez-Brawley and Gualda, 2020 & Gualda and Martinez-Brawley, 2019). Utilizing more quantitative data, Castaño-Pulgarin, Suarez-Bentacur, Tilano Vega and Herrera-Lopez (2021) published a Systematic Review exploring how the internet and social media may or may not constitute an opportunity to online hate speech. The final paper was based on 67 studies eligible for their analysis, derived from 2,389 papers found in searches. Their methodology followed accepted principles of investigation and limitations were clearly stated. Their conclusions indicated that cyberhate is amplified by the use of the internet and social media and that it has real effects and consequences in the exacerbation of violence. Furthermore, in Europe, the latest Shadow Report by the European Network Against Racism (ENAR), for example, “pointed out the rise of racist discourse both on social media and the internet” (Baider, Assimakopoulos and Millar, 2017, p.2). The same authors foresee penalties “for those publicly inciting racial hatred...” (Baider, Assimakopoulos and Millar, 2017, p.2).

Looking again at the U.S., the Southern Poverty Law Center, which, as already mentioned, for decades had published research material (maps and articles, reports, etc.) documenting hate and extremist political groups involved

in this dissemination, redoubled its efforts on the internet (SPLC, 2022&2023). The type of groups mentioned by the SPLC are diverse and appear similar to those involved in Europe. The SPLC mentions over 15 types of hate groups active in the U.S. not including about 488 that are, in general, anti-government (SPLC, 2023). Additionally, the literature on hate groups in general found in searches conducted for these reflections further corroborates the lay-person's concerns that the network does have an effect on the dissemination of hate speech and on the possibilities of in-person violent or abusive consequences for the public and the democratic order.

#### 4. THE HISTORICAL PROBLEM OF DEFINITIONS

In spite of the commendable terms of the U.N. Universal Declaration of Human Rights (1948), which has successfully protected individuals against a number of oppressive governments and circumstances immediately following WWII, the more real, though often elusive damage of "hate speech" has continued to be felt. Anti-Semitism, some will arguably state, appeared to be subsiding but as we see today, it had not been eradicated. Furthermore, immigration crises in Europe and the U.S. coupled with totalitarian governments' decisions, have increased the flow of people across nations and continents and stirred dormant hatred against specific groups. Through manipulation and the internet, hate speech has continued to take hold of the minds of vulnerable people, unfortunately, sometimes of the very same people the Declaration is protecting, who might transfer their "hate" to another vulnerable group.

The Universal Declaration of Human Rights has protected categories of oppressed people but has also included freedom of expression. Article 19 of the UDHR states that

«Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers» (U.N. Declaration, *United Nations, 1948*).

Article 20, of the same UDHR states that «any advocacy of national, or religious hatred that constitutes excitement to discrimination, hostility or violence shall be prohibited by law» (Haraszti, 2012, xiv). The dilemma presented is that the different philosophies and the legal systems of the various members of the U.N. render the "right to freedom of opinion and expression" contained in Article 19 difficult to define let alone enforce in the various member states.

The U.N. Document acknowledges that their definition is not a legal definition. The U.N also points to related prohibitions in international human rights law. What appears clear is that the current U.N strategy launched in 2019 is broader than the 1948 definition. The 1948 international document was drawn as a result of the anti-Semitism, which had been exported out of Nazi Germany, but also applied to many other minority/ethnic groups such as the Romani and other itinerant people and to individuals because of sexual orientation and other differences.

But the actions of the U.N., regardless of the times, have been universally respected. They have been respected because as Haraszti, (2012) suggested, “it has always been inescapable that any universal standard reconciling Article 19... would have to tend toward the minimal intrusion principle” (in Herz and Molnar, eds, p. xiv.) So, Haraszti continues, “there we stand, without an official universal standard, while another sobering bit of reality is the proliferation of new local hate speech laws that strive to protect citizens from their own rows.” (p.xv).

The debate over definitions and other possible ways to curb hate speech was taken up by the European Commission. In a chapter on Regulating Hate Speech in the E.U, Alkiviadou (2017), comments that the Criminal Law Council of the E.U managed “to take a leap forward” publishing a Framework Decision on Combating Racism and Xenophobia through Criminal Law. As was observed before, this framework used the various member laws to analyze how countries have dealt with the phenomena underlying hate speech. In other words, various countries passed their own laws or used their criminal codes to deal with observable acts resulting from hate speech. In *Regulating Hate Speech in the E.U.*, Alkiviadou offers some excellent details of the laws or regulations passed by various countries, which clearly reveal the different understandings of what constitutes hate speech (2017, pp. 7, 8 & 9).

A couple of examples would be Denmark, where, as it happens in many other countries, a lot of presumed hate speech was not, according to the courts, punishable as such because it did not happen in public spaces or did not meet other requirements (e.g. “coloured people like you are not allowed in my parents’ apartment”, uttered in a nursing home was not considered hate speech because the home in question was not a public space) (Alkiviadou, 2017, p.8). Italy limits itself to the protected categories, which make it a crime “to propagate ideas based on racial superiority, racial or ethnic hatred. Or to instigate to commit or commit acts of discrimination for racial, ethnic or religious motives”. LGBTQ groups are not included (p. 8). A third case would be Spain, where Alkiviadou mentions an important Article 510 of the Criminal Code which punishes “those who provoke discrimination, hate or violence against groups or associations due to racist, anti-Semitic or any other reason

related to belief...”, 2017, p. 9-10). More broadly than in other countries, regulations in Spain include not only “race and national origin but also sexual preference, illness or handicap.” (Alkiviadou, 2017, p.10).

Finally, in a realistic observation based on much experience with international organizations, Haraszti (2012) laments,

«In the West, pursuit of a benign utopia of an offense free society takes its toll on freedom of expression. And when the model of ad hoc speech limitations is copied in new democracies, and even more so in new autocracies, it often becomes the “old utopia” of an offense-free state, a specter of the dictatorial past believed to have been buried» (pp. xv-xvi).

Overcoming such realistic pessimism by focusing on the gravity of current developments, U.N. Secretary General Gutierrez launched a universal call to eradicate hate speech particularly in the internet. Additionally, on 16 June 2023, the U.N. High Commissioner for Human Rights, Volker Türk, further called on all members to “work together to build a more respectful and civil world” ahead of the International Day for Countering Hate Speech. More importantly, he specifically indicted the internet, particularly social media, for its role in the dissemination of Hate Speech. He stated,

«We know that the spread of hate is used by those who want to sow divisions, to scapegoat and to distract from real issues. Social media is a remarkably fertile ground for hate speech, providing it with both unprecedented reach and speed. And hate breeds bigotry, discrimination and incitement to violence» (Türk, in *United Nations Human Rights. Office of the High Commissioner, 2023, par.2*).

Exhortations aside, in some countries such as the U.S., the legal and societal challenge or dilemma to curbe the dissemination of clearly corrosive hate speech lies at the junction of freedom of expression or the right to free speech (a constitutional right) and the ability of human beings to inflict pain and injury on individuals and groups that the various types of hate speech can reach. Unfortunately, people often disregard the consequences of hate speech, separating language from the external harm caused. (At the outset, we discussed the cultural crisis). Additionally, the tendency has been to disguise hate speech in codes or special words or symbols (intrinsic features); thus, at least in the U.S., its suppression may often have to wait to the determinations of a court. The diversity of definitions across the world makes the identification of hate speech difficult in itself, though it may appear clear to the lay person. The subtleties of language further complicate the intent which matters in law. Additionally, what is illegal vs. what can be denounced and criticized but is not considered criminal is affected by the lack of precision of the definitions,

linguistic issues and sometimes the spirit of the times, including the political climate and legal interpretations. Even proponents of a minimalist approach to legal interpretations of the basic breath of freedom of speech, such as the one that exists in the U.S, there is as Waldron suggested, a “weird artificiality about how we think about hate speech” (Liptak, in Herz and Molnar, 2012, xix).

The complications of definition are compounded by other complexities of context. Context has been used by national and international tribunals to interpret and clarify definitions that might not coincide. Mendel stated that “international and most constitutional guarantees of free speech provide for a balancing between freedom of expression and other key public and private interests, which inevitably depend, at least to some extent on context” (Mendel, 2012, p.423).

This chapter is not intended to be a legal analysis of hate speech, but in addressing it, any researcher must be aware of those complexities which seem to grow larger the moment one looks at them closely. But in spite of controversies tied to legal and political factors surrounding the interpretation of hate speech in the internet, the many activist groups show serious efforts to identify hate speech, assign it to the sources where it originates, and draw plans to combat it. Its epidemic nature has caused much harm and consternation and dissemination via the internet has often been connected to specific hate groups with a history of condoning criminal hate behavior and violence.

The Southern Poverty Law Center in the U.S., for example, has maintained a very up to date reference of documented cases of hate speech connected to existing hate groups, the methods used, their violent consequences, and the results of litigation primarily in cases of racist and anti-LGBTQ. Their publications are accessible in their website by the Southern Poverty Law Center (<https://www.splcenter.org/hate-map>) and via tweet @Hatewatch. This source permits the identification of hate speech by its origins, (Was it a known Hate Group? Was it and incidental occurrence? or other contextual questions). This information can frequently be helpful in litigations. The Center is also active in litigations regarding many other social injustices.

Other organizations with a long history of fighting hate speech, anti-Semitism, anti-Muslim and anti-Racist incidents litigate on behalf of its victims. Among them, salient are the American Civil Liberties Union; the Anti-Defamation League of B’nai Brith. Awareness of the consequences of hate speech in the internet is clearly growing in the U.S. and Western Europe, but unfortunately, so are the number of hate, anti-government and other extremist groups who find it hard to believe what is reported even by the best sources.

## 5. THE EPIDEMIC NATURE OF HATE SPEECH VIA THE INTERNET. TWO SALIENT FACILITATING FACTORS.

Looking at the many factors resulting in the epidemic nature of hate speech, some psychological and sociological and some technical factors appear more salient than others. I will mention here a couple of the more salient technical or psycho/technical ones.

### 5.1. Manipulation

There is a strong element of what can be called “manipulation”. In *The Philosophy of Manipulation*, Jongepier and Klenk (New York Routledge, 2022), have amply discussed the many factors and philosophic questions involved in the concept of manipulation. Applied to the internet “manipulative influences typically go against the self-interest of the persons being manipulated. That is, they lead to outcomes that are directly or indirectly unbeneficial or outright harmful [...] and may include beliefs, emotions or desires formed by the manipulated person” (p.24). These authors recognize the complexity of the concept of manipulation, which can be said to exist even in romantic love, or parental nudging, but those actions are seldom intended to cause harm to the manipulated. So good judgment becomes an important factor in distinguishing harm.

On a similar vein, Marianna Capasso has written in *Manipulation as Digital Invasion* (2022), that, for example,

«Political actors in the public sphere often manipulate others: they provide incentives and other means to purposely influence and alter individuals' behaviours and beliefs. In general, manipulation is deemed to be a kind of intentional disruption or imposition in the expected functioning of individuals' decision-making processes» (Capasso, 2022, p. 180).

However, there is no consensus on the definition of manipulation (Sunstein, 2016; Coons and Weber, 2014). Ethicists have been concerned with the manipulative aspects of technology because, unlike common mostly accepted and expected political behavior, the invasive and persistent nature of the messages confuses and undermine the users' autonomy to make rational choices in receiving them. One would often hear recipients upset by, let's say, repetitive text messages opened by mistake the first time. Furthermore, manipulation, unlike direct coercion, seems to operate in hidden ways. “It seems very plausible that, in order to succeed, manipulation must be hidden in the sense that the intentions of the manipulator... remain hidden” (Susser et al. 2019, in Jongepier and Klenk, 2022, p.27).

In reality, at the recipient's end, the damage done by hate speech becomes hidden from the recipient's good judgement. The recipient's good judgement, rationality and autonomy are curtailed and thus, even independent professional autonomy can be blurred (Martinez-Brawley, 2022; Howard, 2019). Like in the case of a disease, warnings can be issued as educational tools to avert the harmful results (calls for vaccination, or the use of masks, or isolation, etc.) but the persistence of contrary or just negative messages and their hidden intent can often overcome the good judgment and rationality of well-educated recipients. In my personal experience, I have been able to observe such situations in the US among professionals many of whom denied the validity of vaccination during Covid, considering it part of a conspiracy. In political circles, the U.S and other countries are experiencing the manipulative behavior of politicians, significant numbers of voters choosing to be convinced by messages not only of vain promises but of unfounded hate. Thus, the extent of the influence of propagandistic speech under the free speech clauses still remains an ethical and legal quandary.

## **5.2. The contagion effect of online communication**

Again, the metaphor of the disease brings us to speak of a certain contagion effect in the reception and dissemination of messages of hate. The ability of any recipient to spread the messages via the internet multiply the contagion. Unlike the dissemination of hate before and immediately after WWII, current dissemination does not require physical mobility. This is rather obvious and has been documented many times. An interesting additional factor that adds to the contagion effect is what Jongepier and Klenk (2022) have labeled targeting and personalization. Hate messages are targeted to specific vulnerable populations and often combine a level of personalization which enhances the recipient's vulnerability. Thus, hate messages that address LGBTQ groups, for example, have been identified as not only containing hate but also as making an individual recipient feel a pariah, outside the bounds of society as we know it.

This has proven to be particularly dangerous for teen-agers, often a target of hate messages leading to a pandemic of depression because given their age and stage of development, they have an acute need to belong to a group. That makes them already vulnerable. The natural identification they might have with idols of the time enhances the possibility of personalization of the message. The fact that teen-agers go to schools and clubs makes them identifiable and easier to personalize their characteristics. So, physical appearance as well as clothing, or domicile, etc. are often targeted. The same principles apply to girls and has resulted in dangerous epidemics of targeting them for characteristics such as weight, clothing, tattoos, and many others. Here, definitions of hate speech,

though essential to understand the problems, are not easy to enforce because intention and effect are influenced by context.

More recently, the *U.N. Strategy and Plan of Action on Hate Speech* (June 18, 2019) has added to the definition of hate speech while cautioning that it must be discriminatory and pejorative but that it can include real or perceived “identifying factors such as religion, ethnicity, nationality, race, colour, descent, gender, language, economic and social origin, disability, health status and sexual orientation.” (United Nations, 2019).

«A majority of developed, democratic nations have enacted hate speech legislation—with the contemporary United States being a notable outlier—and so implicitly maintain that it is coherent, and that its conceptual lines can be drawn distinctly enough. Nonetheless, the concept of hate speech does indeed raise many difficult questions: What does the ‘hate’ in hate speech refer to? Can hate speech be directed at dominant groups, or is it by definition targeted at oppressed or marginalized communities? Is hate speech always ‘speech’? What is the harm or harms of hate speech? And, perhaps most challenging of all, what can or should be done to counteract hate speech?» (Anderson and Barnes, 2023).

The distinction between “assaultive hate speech” and “propagandistic hate speech” is helpful when discussing these harms (Langton 2012; 2018; see also Gelber and McNamara, 2016) but as most authors suggest, it is always subtle and depends on judgement. These authors also discuss “face-to-face encounters” and “incidences of general circulation”. From their descriptions, we might say that an insult yelled in a bar by one misbehaving person on a corner to another at a distance, while it might still be an insult, and still be “an assaultive speech act”, may or may not be or legally qualify as hate speech. It would depend on the content, the language used and the intent. Hate speech must fulfil certain requirements. Lawrence (1993) and Delgado (1993) have focused on hate speech’s ability to produce “direct, immediate, and substantial injury” (Lawrence, 1993, 57), such as “immediate mental or emotional distress” (Delgado, 1993, 93–94). The problem of hate speech is its capacity to escalate through intent and dissemination and repetition through time.

Summarizing the results of their study *Evidencing the Harms of Hate Speech* which surveyed the experiences of the victims of hate speech, Katharine Gelber and Luke McNamara conclude that «the distinction between face-to-face encounters and general circulation hate speech is not always clear in the everyday experiences of racism endured by targets» (2016, p.326).

While it might not be possible to predict the various types of speech that can plausibly be harmful or psychologically problematic, and the debate over whether “hate speech” can or cannot be constitutionally regulated in various

countries continues, the reality is that even people outside the scope of having to make decisions (judgements) on this matter can readily identify the difference between a sign or cartoon or a shout of “No Muslims here”, or the old N.Y.C. example of “Irish need not apply” or “No Blacks allowed on these premises” (both regulated in the U.S. by labor and other codes) and the greater complexities of a court deciding whether the use of a particular word presents an infringement of people’s rights, or violates a group’s community dignity.

On a more hopeful tone, Alkiviadon writes:

«Notwithstanding the perplexities associated with defining “hate speech” [...] the E.U managed, after seven years of long negotiations (European Commission, 2014:1) to take a major leap forward in 2008 with its Framework Decision on Combating Racism and Xenophobia through Criminal Law (Council of the European Union, 2008)» (Alkiviadou, 2017).

Alkiviadon (2017) nevertheless makes clear that it was “hate speech” and the many divergences of understanding and legal systems across the various countries that kept the discussion going for so long. Personally, I must admit that researching the subtleties and problems associated with the interpretation – both public and legal – of what constitutes “hate speech” makes clear the need for ongoing research and continued international discussion.

## 6. SOME CONCLUDING REFLECTIONS

My intent in this chapter was to present, through a bird’s eye view or exploration, some clarifying comments on the conundrum of hate speech, its proliferation on the internet and any attempts at curbing or regulation. I believe that not only the casual readership but also students in different fields would benefit from the discussions. While the observations here add to the repertoire of knowledge, it would be presumptuous to assume definitive conclusions. However, my overview of literature, socio/political, historical and legal issues summarizes, in a sense, much material of general interest across fields of study.

**On Definitions:** It is clear that differences in histories, contexts and legal systems show divergent thinking. There is not one single way to define hate speech across the nations and even the agreed upon principles of the U.N. or the E.U or other organizations run against different road-blocks approaching remediation. There are different roads to possible solutions. There is no common “definition” of what constitutes hate speech that could be legally pursued, banned and prosecuted in different cultures, political systems, paradigms and languages. This road block continues to prevent substantial legal or negotiated results. Yet, the legal approach appears to have been and continues to offer respected guidelines albeit not “universal”. Categories of

speech, and types of behavior found offensive in one country are often found offensive in many.

A story told by Eleanor Roosevelt, a co-chair of the 1947 U.N. group drafting the Declaration on Human Rights about Peng Chung Chang of China, Vice Chair, bears repeating. Other members were Dr. Humphrey and Charles Malik from the U.N. Secretariat, who had the task of formulating a preliminary draft. Roosevelt writes in her memoirs:

«Dr. Chang was a pluralist and held forth in charming fashion on the proposition that there is more than one kind of ultimate reality. The Declaration, he said, should reflect more than simply Western ideas and Dr. Humphrey would have to be eclectic in his approach. His remark, though addressed to Dr. Humphrey, was really directed at Dr. Malik, from whom it drew a prompt retort as he expounded at some length the philosophy of Thomas Aquinas. Dr. Humphrey joined enthusiastically in the discussion, and I remember that at one point Dr. Chang suggested that the Secretariat might well spend a few months studying the fundamentals of Confucianism)» (Roosevelt, in United Nations, 2023, par.4).

The final draft was handed eventually to the Commission on Human Rights, which was being held in Geneva. The draft declaration sent out to all U.N. member States for comments became known as the Geneva draft. As we know, by its resolution 217 A (III) of 10 December 1948, the General Assembly, meeting in Paris, adopted the Universal Declaration, which though respected and abided by most countries is not a legal document. All international bodies continue to need Confucianism in their deliberations!

**On Different Countries:** Individual countries have regulated “hate speech” mostly through their criminal codes but in all cases, it appears that what is mostly regulated is not the speech itself but the consequences of the speech in acts of public offense, incitement to violence and public humiliation. Some countries have listed “protected categories” which differ from one to the other though the common denominators tend to be race, ethnicity, religious affiliations with categories such as gender and sexual orientation not always included. Litigations on these matters have to be adjudicated in courts because they butt against free speech rights or protections of political speech. These considerations are stated here in brief. Further details are offered in the discussion and the literature. Litigations in many cases have been successful and have become part of national and international jurisprudence. which tells about the success of court decisions.

**On Education:** Different tools have been quite successful in achieving public awareness of the goals of creating a more open and inclusive civil society, if not fundamentally curbing hate speech as such, least of all on the internet.

(Business pressures have done that in a few cases.) Individuals who often thought in terms of a single group to which they already belonged became aware that they lived in a very diverse society. In the U.S. government and non-government institutions were “mandated” or at least “encouraged” by various types of regulations, to be inclusive. Even though regulations around the use of labels and words and, at times, were controversial and many litigated, the public at large became conscious that offensive speech, whether or not classified as hate speech, was not acceptable. International organizations were re-enforcing the message using the “non-binding” but effective prohibitions of the U.N. and other international bodies.

Educational institutions, at least in higher education, have a tradition of governing themselves and permitting or prohibiting certain types of speech for the sake of their values and goals. Their messages have force as long as the community within those institutions support them. So, for example, employees in schools could be internally sanctioned for the use of hostile or unacceptable speech. However, for teachers and administrators in elementary and secondary public schools, correcting children and youth in the use of “hate speech” became a dilemma. Administrators are very constrained in what they can do when, for example, a 6 year old persists in using “hateful words” or “derogatory speech” in interactions with classmates or teachers if the parents do not support the correction. The educational process of change is slow and constraining of the change agent. This is very apparent in the U.S., where public schools are supported by local taxes and accountable to both local and central authorities. The role of the “public” becomes essential. This complex discussion is outside the scope of this chapter. It is important to note that similar dilemmas exist in different countries, even when schools are more centrally administered but, because of the nature of the task, must work in cooperation with parents and communities. Here, as legal experts would agree, context matters.

**On Public information:** What used to be called the “Press”, shares the same educational goals and is useful in remediation. The press, while very jealous in guarding its freedom, operates with a different business and partisan tradition. While large newspaper operations, for example, have tended to be often connected to specific political groups, they still keep a degree of autonomy that can prove very valuable in changing messages and minds and is also protected under “freedom of the press”. But unfortunately, in the measure in which public information confirms the harm that offensive and hateful speech does to society, fanatic groups entice members and non-members to join their causes. The criminal law, which is also slow and long term must be looked at carefully. There are many limitations inherent in law. Law must be complemented by common sense and good judgement. As Howard (2019) suggests, “Law protects our freedom in two ways” [...] but, Howard adds, “no

respected formulation of freedom prevents people from acting responsibly. The end of law... is not to restrain, but to preserve and enlarge freedom.” (p.78) “Law is not supposed to define what is reasonable... but to define and protect against what is unreasonable” (Howard, p.78). It is within those boundaries that good judgement, rationality and common sense must reign (Martinez-Brawley, 2022). Neither the law nor bureaucratic regulations can succeed if citizens view them only as a manual which will tell them how to act as responsible human beings under all situations. No law can do that, and no bureaucratic regulation should forbid us from behaving as our best judgement dictates.

Our world community has been disrupted by ambition and hate. The desire for power has always been a force in democracy, but the desire for absolute power has led societies away from democratic institutions and into autocracy and tyranny. But democracy requires good judgement, rationality and common sense, and pre-supposes autonomous beings. Thus, the non-ending chore of civil society is to be relentless in curbing the flow and use all instruments at its disposal within the boundaries of good judgement.

## REFERENCES

- Afshin, E., & Molier, G. (eds) (2015). The Various Aspects of Freedom of Speech (pp.1-13). *Freedom of Speech under attack*, Den Haag, The Netherlands, Eleven International Publishers.
- Alkiviadou, N. (2017). Regulating Hate Speech in the EU (pp-6-10). In Stavros, A.; Baider, F.H. and Millar, S. (2017). *Online, Hate Speech in the European Union A Discourse Analytic*. Springer Nature. <https://doi.org/10.1007/978-3-319-72604-5>
- Anderson, L., & Barnes, M. (2023). Hate Speech. In Zalta, E.N. & Nodelman, U. (eds.). *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2023/entries/hate-speech/>
- Baider, F.H., Assimakopoulos, S. and Millar, S. (2017). Hate Speech in the EU and the C.O.N.T.A.C.T. Project (pp.1-6). In Stavros, A.; Baider, F.H. and Millar, S. (2017). *Online, Hate Speech in the European Union A Discourse Analytic*. Springer Nature. <https://doi.org/10.1007/978-3-319-72604-5>
- Brooks, D. (2023, August 10). Hey America, Grow Up! *The New York Times*. <https://www.nytimes.com/2023/08/10/opinion/trauma-mental-health-culture-war.html>
- Castaño-Pulgarín, S.A., Suárez-Betancur, N., Tilano Vega, L.M., & Herrera López, H.M. (2021). Internet, social media and online hate speech. Systematic review. *Aggression and Violent Behavior*, vol.58, 101608, <https://doi.org/10.1016/j.avb.2021.101608>
- Capasso, M. (2022). Manipulation as digital invasion A neo-republican approach (pp.180-198). In Jongepier, F., & Klenk, M. (ed.). *The philosophy of online Manipulation*. Routledge. <https://www.taylorfrancis.com/chapters/oa>

- edit/10.4324/9781003205425-11/manipulation-digital-invasion-marianna-capasso
- Cogburn, D. (2016). Cyberspace and Social Media. In Acharya, A. (Ed.). *Why Govern?: Rethinking Demand and Progress in Global Governance* (pp. 252-271). Cambridge University Press. <https://doi.org/10.1017/CBO9781316756829.013>
- Coons, C., & Weber, M. (eds.) (2014). *Manipulation: Theory and Practice*. Oxford University.
- Delgado, H.L. (1993). *New immigrants, old unions: organizing undocumented workers in Los Angeles*. Philadelphia: Temple University Press.
- Gelber, K., & McNamara, L. (2016). Evidencing the harms of hate speech. *Social Identities*, 22(3), 324-341. <https://www.doi.org/10.1080/13504630.2015.1128810>
- Gualda, E., & Martinez-Brawley, E. (2009). Discursos que discriminan al Africano en España y al Latino en los Estados Unidos: Similitudes en el tratamiento en prensa (pp. 355-370). In Montaña Garcés, M., & Gualda, E. (eds.). *Africa-Occidente. Necesidades de Nuevas Relaciones*. Universidad de Huelva.
- Haraszti, M. (2012). Foreword: Hate Speech and the Coming Death of the International Standard before It was Born. (Complaints of a Watchdog) (pp. xiii-xxi). In Herz, M., & Molnar, P. (2012). *The Content and Context of Hate Speech. Rethinking Regulations and Responses*. Cambridge University Press.
- Herz, M., & Molnar, P. (2012). *The content and Context of Hate Speech. Rethinking Regulation and Responses*. Cambridge University Press. [http://assets.cambridge.org/97805211/91098/frontmatter/9780521191098\\_frontmatter.pdf](http://assets.cambridge.org/97805211/91098/frontmatter/9780521191098_frontmatter.pdf)
- Howard, P.K. (2019). *Try Common Sense: Replacing the Failed Ideologies of Right and Left*. W. W. Norton & Company.
- Jongepier, F., & Klenk, M. (eds.) (2022). *The Philosophy of Online Manipulation*. Routledge. <https://doi.org/10.4324/9781003205425>
- Keillor, G. (1985). *Lake Wobegon Days*. Viking.
- Keillor, G. (1997). *Wobegon Boy*. Penguin Books.
- Langton, L., Berzofsky, M., Krebs, C., & Smiley-McDonald, H. (2012). Special Report. National Crime Victimization Survey. Victimization not reported to the police, 2006-2010. United States. *Bureau of Justice Statistics*. Washington, D.C.: U.S. Department of Justice, Office of Justice Programs. <https://bjs.ojp.gov/content/pub/pdf/vnvp0610.pdf>
- Lasch, C. (1979). *The Culture of Narcissism. American Life in an Age of Diminishing Expectations*. W.W Norton and Company. [https://thezeitgeistmovement.se/files/Lasch\\_Christopher\\_The\\_Culture\\_of\\_Narcissism.pdf](https://thezeitgeistmovement.se/files/Lasch_Christopher_The_Culture_of_Narcissism.pdf)
- Liptak, A. (2012). Foreword: Hate Speech and Common Sense (pp. xix-xxii). *The content and Context of Hate Speech. Rethinking Regulation and Responses*. Cambridge University Press. [http://assets.cambridge.org/97805211/91098/frontmatter/9780521191098\\_frontmatter.pdf](http://assets.cambridge.org/97805211/91098/frontmatter/9780521191098_frontmatter.pdf)
- Martinez-Brawley, E. E. (1990). *Perspectives on the Small Community*. Silver Spring, MD: NASW Press.

- Martinez-Brawley, E., & Gualda, E. (2020). Transnational Social Implications of the 'War Metaphor' concerning Coronavirus: A Bird's Eye View. *Culture e/revi Studi del Sociale-CuSSoc*. Special Issue, June 2020, 259-272. <https://www.cussoc.it/journal/article/view/133/97>
- Martinez-Brawley, E. (2000). *Close to Home: Social Services and the Small Community*. Washington D.C., NASW Press.
- Martinez-Brawley, E. (2022). Judgement, Common Sense and Discernment: Contributions from Humanistic Thinking to the Social Professions. *Revista Centra de Ciencias Sociales*, 1(1), pp.51-66. <https://doi.org/10.54790/rccs.11>
- Mendel, T. (2012). Does International Law Provide for Consistent Rules on Hate Speech? In *The Content and Context of Hate Speech: Rethinking Regulation and Responses*. Cambridge University Press, p. 417-429.
- Nisbet, R. (1953). *The Quest for Community*. Oxford University Press.
- Roszak, T. (1978). *Person/Planet. The creative disintegration of industrial society*. Anchor Press/Doubledate
- Putnam, R.D. (2000). *Bowling Alone*. Simon & Schuster.
- Putnam, R.D. (1995). Bowling Alone: America's Declining Social Capital. *Journal of Democracy*, 6(1), 65-78. <https://doi.org/10.1353/jod.1995.0002>
- Southern Poverty Law Center, SPLC (2019). Combating Nativism. Protecting the Rights of Immigrants. [https://www.splcenter.org/sites/default/files/1901\\_splc\\_immigration\\_portfolio\\_web\\_0.pdf](https://www.splcenter.org/sites/default/files/1901_splc_immigration_portfolio_web_0.pdf)
- Southern Poverty Law Center, SPLC (2022) The Year of Hate and Extremism 2021. <https://www.splcenter.org/year-hate-extremism-2022>
- Southern Poverty Law Center, SPLC (2023) The Year of Hate and Extremism 2022. <https://www.splcenter.org/year-hate-extremism-2022>
- Sunstein, C. (2016). *The Ethics of Influence: Government in the Age of Behavioral Science*. Cambridge University Press.
- United Nations (1948). *Universal Declaration of Human Rights*. <https://www.un.org/en/about-us/universal-declaration-of-human-rights>
- United Nations (2019, June 18). United Nations Strategy and Plan of Action on Hate Speech. <https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml>
- United Nations Human Rights. Office of the High Commissioner (2023, June 16). *U.N. Human Rights Chief: Hate speech has no place in our world*. <https://www.ohchr.org/en/press-releases/2023/06/un-human-rights-chief-hate-speech-has-no-place-our-world>
- Waldron, J. (2012). Hate Speech and Political Legitimacy (pp.329-340). In Herz, M. and Molnar, P. (2012). *The content and Context of Hate Speech. Rethinking Regulation and Responses*. Cambridge University Press. [http://assets.cambridge.org/97805211/91098/frontmatter/9780521191098\\_frontmatter.pdf](http://assets.cambridge.org/97805211/91098/frontmatter/9780521191098_frontmatter.pdf)
- United Nations (2023). History of the Declaration (par. 4). <https://www.un.org/en/about-us/udhr/history-of-the-declaration>



# ONLINE HATE SPEECH AS AN UNINTENDED AND PARADOXICAL OUTCOME OF CANCEL CULTURE<sup>1</sup>

EMILIANA MANGONE

*Professor. University of Salerno, Italy*  
*Associate member. CNR-IRPPS of Rome, Italy*

STELLAMARINA DONATO

*Ph.D. LUISS. University Rome. Italy*

**DOI:** 10.14679/2752

## **Summary:**

1. Cancel culture and its relationship to hate speech.
2. Related concepts and unintended consequences of cancel culture.
3. From cancel culture to hate speech: the dynamics for the LGBTQIA+ community and migrants.

References.

## 1. CANCEL CULTURE AND ITS RELATIONSHIP TO HATE SPEECH

The phenomenon of cancel culture has experienced extensive proliferation since the latter part of the 20th century, primarily propelled by the influence of social media. This trend is defined by the public denunciation and isolation of individuals due to their perceived offensive or contentious statements. Intriguingly, this phenomenon has engendered a paradoxical outcome in the shape of amplified hate speech—an online manifestation of violence. This very form of virtual aggression finds its definition in the perspective of the Committee of the Council of Ministers of Europe, as laid out in Recommendation CM/Rec(2022)16<sup>2</sup>:

---

<sup>1</sup> This chapter is part of the Project entitled “Conspiracy theories and hate speech online: comparing patterns in narratives and social networks about COVID-19, immigrants and refugees, and LGBTI people [NON-CONSPIRA-HATE!]”, Project Grant PID2021-123983OB-I00, funded by MCIN/AEI/10.13039/501100011033/ and by “ERDF A way of making Europe”. We are also grateful for the support of the following research centers and groups: “Estudios Sociales E Intervención Social, ESEIS” and the center “Pensamiento Contemporáneo e Innovación para el Desarrollo Social, COIDESO” of the University of Huelva (Spain), and research group “POGES ‘Popolazione, Genere e Società” of the CNR-IRPPS of Rome (Italy).

<sup>2</sup> Recommendation CM/Rec(2022)16 of the Committee of Ministers to member States on combating hate speech (Adopted by the Committee of Ministers on 20 May 2022 at the 132 Session of the Committee of Ministers).

«as all types of expression that incite, promote, spread or justify violence, hatred or discrimination against a person or group of persons, or that denigrates them, by reason of their real or attributed personal characteristics or status such as “race”, colour, language, religion, nationality, national or ethnic origin, age, disability, sex, gender identity and sexual orientation».

While cancel culture is often seen as a means of holding people accountable for their actions, it has inadvertently created an environment that fosters hate speech. Cancel culture, with its collective action and public outcry, may seem like a powerful tool for social justice and accountability.

Nevertheless, this phenomenon can foster a fractured and deeply divided cultural landscape, providing fertile ground for the proliferation of hate speech, as evidenced by research (Lumsden & Harmer, 2019). This phenomenon finds a troubling breeding ground within digital platforms (Müller & Schwarz, 2018), consequently amplifying the prevalence of hate speech within social media realms (Bartlett et al., 2014). Regrettably, this surge in hateful discourse raises the specter of heightened aggression directed towards marginalized and susceptible communities, including but not limited to women, migrants, LGBTQIA+ individuals, and others who find themselves at a higher risk (Bartlett et al., 2014).

Cancel culture operates on the premise of punishing individuals for their past mistakes or offensive behavior by publicly shaming them, boycotting their work, and demanding their removal from platforms or positions of influence. This form of punishment, frequently motivated by the pursuit of social justice and the safeguarding of marginalized factions, has the unintended consequence of potentially stoking the flames of hate speech, as it inadvertently constructs individuals or entities as “scapegoats” (Girard, 1986). Paradoxically, this approach can at times trigger a reciprocal surge of animosity from those who find themselves cast in the role of these very “scapegoats”. When individuals face the prospect of being canceled, they may feel attacked and marginalized themselves, leading them to respond with even more hateful and offensive language. Furthermore, cancel culture can also create an environment of fear and self-censorship, where individuals are hesitant to express their opinions or engage in open dialogue.

The transformation of communication systems has not only altered how individuals interact but has also facilitated the migration of offline violence to online platforms, disproportionately affecting some groups inside the society like migrants and the LGBTQIA+ community. The belief that online manifestations of hate and violence are less real due to their digital nature is, indeed, erroneous. On the contrary, online forms of discrimination and violence often result in tangible effects, with real victims and perpetrators (Donato, Eslen-Ziya & Mangone, 2022), and even contribute to the commodification of online violence (Slater, 2005).

The inherent intricacies encapsulating acts of violence, regardless of their diverse manifestations, underscore the necessity to explore tools that more effectively facilitate the augmentation of understanding regarding this longstanding yet evolving phenomenon. Particularly with the proliferation of social media, which has facilitated the migration of offline violent behaviours into the multifaceted realm of the online world, there arises a compelling impetus to harness resources that can illuminate these complexities further. This is even more accurate in the current historical phase, during which the transition from the network society (Castells, 1996) occurs. This society is characterized not only by the consequences of technological innovation and changes in capitalist structures but also by cultural transformations based on individual freedoms and social autonomy, through which identity-based claims are expressed. This transition is towards the platform society (van Dijck, Poell & De Waal, 2018), where platforms are characterized as places for exchanging communicative practices, forms of togetherness, and participation in public life. These platforms also encompass technologies that enable both individuals (individual or collective) and institutions to connect and achieve their goals. In this manner, a new ecosystem for violence (online) is created, to the extent that it becomes necessary to define new perspectives of study that provide an additional interpretive framework for hate and violent phenomena. This framework should not be limited to a perspective centered solely on phenomena occurring in physical real-world environments, but should also extend to the study of online environments where the same, or similar, phenomena occur. These phenomena, in turn, should not only focus on the medium utilized but also consider the relationships between micro and macro aspects of interconnected social life, facilitated by digital environments.

Hitherto, it is of unprecedented importance to witness and debate about the complexity of hate speech and its violent undercurrents since they necessitate new perspectives to comprehensively understand the new society shaped by the interconnectedness of practices and realities. Indeed, as society shifts from a network society to a platform society, hate discourses adapt to the digital environment, requiring fresh approaches to studying online and offline violence dynamics.

Online violence encompasses various phenomena including cyber hate (Mortellozzo & Jene, 2017), and this chapter primarily focuses on one of them: hate speech, examining its emergence as a by-product of the digital political subculture known as cancel culture.

Prior to embarking on this endeavour, it becomes imperative to first elucidate certain evolutionary facets. This preliminary step aims to establish a more precise delineation of the theoretical constructs intertwined within the scope of the current contemplation. There is, for instance, much confusion

surrounding the concept of cancel culture, both in terms of its actual meaning – often considered synonymous with wokeness and call-out culture, or even politically correct – and concerning the timeframe when it actually emerged. However, its effects and its consequences are quite clear. Cancel culture, though often misunderstood, involves ostracizing individuals for violating social norms. It originated in the Black Twitter community and gained prominence through movements like #MeToo and Black Lives Matter. The definitions of cancel culture vary, but they share common threads of public shaming and withdrawing support to enforce consequences or penalties for objectionable actions.

The introduction of this terminology into common language is relatively recent. In fact, in 2019, the Australian Macquarie Dictionary named it the word of the year, and in 2020, the Oxford Dictionary did the same. Nevertheless, pinpointing an exact date for the term's entry into common usage is not possible. Despite this, we can argue that the term "cancel culture" originated (starting from 2015) from Black Twitter (a community of Black Twitter users capable of mobilizing a large number of people to address racist and misogynistic discourse). It used hashtags such as #cancelled or #x'isover, where 'x' represents a person or an organization (Roos, 2020). From there, debates about "cancel culture" emerged, a term that began to gain traction around 2017 with the #MeToo movement. This movement, following public revelations of sexual harassment and assault committed by Harvey Weinstein, fought against sexual abuse and violence against women through a strategy based on public denunciation campaigns and boycotts of well-known and powerful figures accused of harassment or violence.

In 2020, the term gained further prominence with the actions of the Black Lives Matter movement, which erupted following protests over the killing of African American George Floyd by the Minneapolis police. The impassioned demands for an end to systemic racism, police brutality, and the racial disparities that have persisted in societies have been the hallmark of the Black Lives Matter movement's response to this occurrence. The influence of social media and technology contributed to the movement's quick growth and resonance by enabling activists to mobilize, organize, and magnify their message on a never-before-seen scale. Using hashtags like #BlackLivesMatter allowed people to voice their support, share their own experiences, and participate in a wider conversation about racial injustice. These digital tools helped coordinate protests, rallies, and educational campaigns in addition to acting as a platform for spreading awareness, which resulted in a surge of support that cut across geographical lines. Finally, the heightened prominence of the Black Lives Matter movement in 2020 signified a turning point in societal consciousness, prompting a widespread reevaluation of systemic racial biases and discriminatory practices.

## 2. RELATED CONCEPTS AND UNINTENDED CONSEQUENCES OF CANCEL CULTURE

Prior to delving into the paradoxical ramifications of hate speech within the context of cancel culture, particularly within the LGBTQIA+ and migrant communities—two groups notably subjected to significant online animosity—it is imperative to establish a conceptual groundwork. This entails providing a clear theoretical elucidation of both the cancel culture phenomenon (which remains subject to substantial confusion) and its closely associated notions. It should be underscored that while only the practical manifestation of cancel culture gives rise to online animosity due to its distinct attributes, the same cannot be said for the erroneously conflated synonymous concepts.

To accomplish this, we shall commence by exploring the diverse definitions attributed to cancel culture, laying a foundation for a comprehensive understanding of the subject.

Cancel culture «can be defined broadly as attempts to ostracize someone for violating social norms» (Norris, 2020, p. 2), or in a narrower sense, as «the practice of withdrawing support for (or canceling) public figures and companies after they have done or said something considered objectionable or offensive» (Lizza, 2020), or even as:

«the withdrawal of any kind of support (viewership, social media follows, purchases of products endorsed by the person, etc.) for those who are assessed to have said or done something unacceptable or highly problematic, generally from a social justice perspective especially alert to sexism, heterosexism, homophobia, racism, bullying, and related issues» (Ng, 2020, p. 263).

or, in the final analysis, «“canceling” is an expression of agency, a choice to withdraw one’s attention from someone or something whose values, (in) action, or speech are so offensive, one no longer wishes to grace them with their presence, time, and money» (Clark, 2020, p. 88).

As clearly expressed, each definition is characterized by certain specifics, but as argued by Tandoc et al. (2022), there are elements that can be considered common and can be synthesized as follows:

«(a) the public shaming of unacceptable behaviour, and (b) withdrawal of support, which are (c) motivated by wanting to see the target persons experience some form of consequence or penalty due to their actions (e.g. losing employment, other revenue streams) or to ensure these persons are socially banished» (p. 3).

Hence, from the shared characteristics of the diverse definitions, it emerges that first and foremost, the act of “canceling” predominantly takes place through social media call-outs, made easier by forms of anonymity that certain platforms provide (Tandoc et al., 2022). These call-outs resemble public condemnations and punishments designed to enforce adherence to societal rules. Furthermore, the concrete actions that manifest range from the simpler action of “unfollowing” to more active measures like “boycotting,” which aims not to purchase products from a particular brand and also to persuade others to do the same.

While cancel culture, woke culture, call-out culture, and politically correct attitudes may appear synonymous, they operate along a continuum. The woke culture emphasizes awareness of systemic discrimination, call-out culture centers on public admonishment, and cancel culture encompasses a collective effort to withdraw support from public figures. These terms are not interchangeable; they represent distinct actions, targets, and outcomes.

In light of this attempt to clarify the concept of “cancel culture,” we can state without a shadow of doubt that, on the one hand, this practice employs strategies similar to consumer boycott logic to withdraw support from brands and companies, thereby damaging their reputation. On the other hand, it predominantly utilizes social media platforms to “shame” individuals or organizations deemed “guilty” – a phenomenon known as the “shitstorm” (Sdrigotti, 2018) – with the intention of exerting pressure to impose sanctions. These sanctions can range from limiting access to public platforms, damaging reputations, and consequently ending careers, to inciting legal actions.

To further elucidate, it’s necessary to clarify that cancel culture is not synonymous with woke culture, call-out culture, or political correctness. “Stay woke” became the rallying call of the Black Lives Matter movement since 2014 (the year of the killing of Michael Brown in Ferguson, Missouri, by the police) to raise awareness of police actions towards African American citizens. However, previously, woke culture referred to the awareness Black Americans had about the systemic and systematic violence perpetuated against them, although this term has since evolved in meaning (Madrid Gil, 2023). The second concept, call-out culture, unlike the other two previously discussed, has a more recent origin. It seems that the culture of “calling out” also originated on a social media platform (Tumblr) that allows the sharing of multimedia content. In 2010, a blog called “Your Fave is Problematic” was created, collecting behaviours or statements from public figures or celebrities deemed offensive, which then spread across the web. The blog functioned as a form of digital vigilance (Loveluck, 2020), involving online actions of surveillance, dissuasion, or targeted punishment (public denunciation or unsolicited attention). These actions are carried out in the name of justice, order, or safety.

Before presenting some reflections on the last of these concepts (political correctness), it’s important to draw a small conclusion. Despite often being

confused as synonyms, cancel culture, woke culture, and call-out culture are not interchangeable. They can be placed on a continuum, with woke culture leading to call-out culture and ultimately to cancel culture. This progression is rooted in the awareness of structural discrimination that leads to calling out and then cancellation. These last two actions not only differ in effects but also in terms of who takes action. Call-out culture can be driven by an individual, usually targeting a behaviour or statement. Cancel culture, on the other hand, requires a larger group of individuals and typically targets public figures or celebrity figures in entertainment.

The term “politically correct,” finally, signifies an ideological and cultural orientation of extreme respect for all, avoiding any potential offense towards specific groups of individuals. According to this perspective, expressed opinions must appear free from prejudice in both linguistic form and substance, based on various variables such as race, ethnicity, religion, gender, etc. From this definition, it’s evident that even the culture of political correctness cannot be considered synonymous with cancel culture.

Having clarified that these four political subcultures (politically correct, woke culture, call-out culture, and cancel culture) are not synonyms and differentiate in terms of their targets and effects, two additional dynamics stemming from their interaction warrant attention and careful consideration: “fake news” (Hughes, 2021) and “woke capitalism” (Lewis, 2020). Regarding fake news, it’s noteworthy that they have become a digital world “must,” yet cancel culture isn’t equipped to fact-check (Hughes, 2021). In many cases, public accusations (often in the form of infamy) can be excessive, simply serving as a way to judge and reject anyone with a socio-political viewpoint differing from one’s own.

As for “woke capitalism,” it was coined by journalist Ross Douthat (2018) to describe actions taken by large organizations or companies to pursue some form of social justice, but that ultimately prioritize profit. These strategies are akin to greenwashing or pinkwashing – advertising a company or organization’s commitment to a social cause based on how it will impact profit. Beyond the appearance of social justice, in many cases, these companies genuinely engage in awareness campaigns to promote the reduction of prejudiced attitudes. However, the ultimate goal remains profit, giving rise to woke capitalism.

At this juncture, having explored the characteristics of these political subcultures that, from different angles, are defined as such, it’s clear that while they give voice to segments of populations that may not otherwise be heard, they also introduce important debates that constitute the perverse effects of these practices. On one hand, there are processes of democratization tied to freedom of expression, potentially leading to a reduction in the plurality of opinions. On the other hand, the issue of fake news distorting phenomena significantly in terms of coverage dedicated to specific related events and the

uniformity of narrative frames (Mangone, 2022) can generate hate speech that accentuates certain themes, concepts, or mental categories while side-lining others. This manipulation of omission, trivialization, and condemnation not only defames identity communities of the discourse recipient but works to make these communities invisible through explicit lack of representation in all forms of storytelling or as scapegoats for societal issues.

### 3. FROM CANCEL CULTURE TO HATE SPEECH: THE DYNAMICS FOR THE LGBTQIA+ COMMUNITY AND MIGRANTS

The narrative approach embraced by participants in cancel culture within the realm of social media hinges on bestowing the “enemy” label upon individuals, reminiscent of the societal irrationality encapsulated in the concept of scapegoating (Girard, 1986). Within this intricate framework of interpretation, a perplexing phenomenon emerges—one that propels individuals to offer illogical societal justifications for their stances, thereby affording a rationalization for the propagation of hate speech. The act of assigning culpability, regardless of its veracity, to a subject or collective earmarked as ‘enemies,’ stands firmly rooted in this underlying logic. Illustrative instances abound, such as the discriminatory actions perpetrated against members of the LGBTQIA+ community or migrants.

This attribution of blame—whether grounded in verifiable truths or unfounded claims—unveils layers of the social contract upon which the bedrock of a community is built. It also brings to light the strategies enacted to shield this foundation from both internal and external threats. This holds true irrespective of whether the milieu is digital, where anonymity may be safeguarded or veiled through pseudonymous profiles, or physical. The ceremonial protocols adhered to serve as barometers, elucidating the societal structures and political ethos prevalent within the respective national contexts from which users of various social platforms hail. Within the context of hate speech, the paradoxical outcome stemming from cancel culture’s distinct attributes offers insight into the political subculture it engenders. This convoluted connection highlights the intricate interplay between cancel culture and hate speech, a testament to the idiosyncrasies intrinsic to the former’s character. We have argued that cancel culture, stemming from the digital age’s amplified voices and increased interconnectedness, finds its origins in the early 21st century as a social phenomenon driven by online activism and social media platforms. While it aims to hold people accountable for their words and actions, cancel culture has also sparked debates about its potential consequences. In the context of migrants, cancel culture repeatedly can lead to hate speech based on prejudice assumptions, stereotyping, and a lack of nuance in addressing complex issues related to immigration. Likewise, the LGBTQIA+

community has witnessed instances where cancel culture's swift judgment might inadvertently overshadow opportunities for education and dialogue, leading to negative polarization and stifling conversations necessary for progress. To support what is being said here, a survey carried out in Italy by the project Vox - Italian Observatory on Rights<sup>3</sup> (in collaboration with the State University of Milan, the University of Bari Aldo Moro, Sapienza - University of Rome and ITSTIME of the Catholic University of Milan), which has been drawing the 'Map of Intolerance' for seven years through a photograph of the hatred expressed in digital conversations, is brought to support this. In 2022, 629,151 tweets were extracted in the January-October period, of which 583,067 were negative (about 93% vs. 7% positive) which, based on the categories considered, are distributed as follows: Islamists (99.9% negative vs. 0.1% positive); Jews (97.7% negative vs. 2.3% positive); disability (98.8% negative vs. 1.2% positive); homosexuals (94.1% negative vs. 5.9% positive); women (89.9% negative vs. 10.1% positive); xenophobia (79.2% negative vs. 20.8% positive). As can be seen, hate speech, as mentioned at the beginning of this contribution, does not only affect the migrant and LGBTQIA+ community, but also other minorities. It should be noted that an important negative trend reversal emerges from this study: homosexual people are once again being targeted by online hate, something that has not happened since 2016, and this highlights a definite attack on personal rights.

The reasons why hate speech against migrants and LGBTQIA+ members emerges from the subculture of cancel culture are to be found in the construction of stereotypes and xenophobia as a consequence of hasty labelling and ostracism of individuals from a migrant background on the basis of assumptions or limited and often fake information, or belonging to minority communities. Associated with this are also conspiracy theories (Gualda & Rúas, 2019), which in the case of migrants translate into the metaphor of invasion (the invasion of the enemy or counter-nature in the case of members of the LGBTQIA+ community).

To further clarify what has been said so far, some possible paradoxical effects of cancel culture that can produce hate speech are presented. With concern to the LGBTQIA+ Community, cancel culture can sometimes misinterpret well-intentioned efforts and contribute to the chilling effect on open dialogue. For instance, a person seeking to learn about LGBTQIA+ experiences might inadvertently use incorrect terminology or express views that are outdated due to their lack of exposure. Instead of allowing room for education, cancelling them outright might discourage genuine efforts to understand and support the community. Cancel culture and related hate speech can also direct towards more divisiveness within the Community. For instance, cancel culture within

---

<sup>3</sup> The project's objectives and the extensive data accumulated over the course of seven years are comprehensively detailed here: <http://www.voxdiritti.it/>

the LGBTQIA+ community can lead to internal divisions, where differing opinions are silenced for fear of backlash. For example, if a member expresses reservations about certain strategies used by LGBTQIA+ advocacy groups, they might face public condemnation, hindering the community's ability to have constructive conversations about its own progress and direction.

For example, an online campaign to cancel a public figure for expressing concerns about immigration might overlook the nuanced conversation about economic, social, and political factors contributing to immigration. Moreover, cancel culture might thrive on quick judgments, which can overshadow the context behind statements or actions. In the case of migrants, this could mean ignoring the challenges and dangers they face in their home countries that compel them to seek better opportunities elsewhere. By hastily canceling discussions on immigration, we might encounter the problem of missing chances to advocate for comprehensive policies addressing the root causes of migration.

In both of these contexts, but also within the other categories that are targeted by hate speech, it's vital to strike a balance between holding individuals accountable for harmful actions or statements and providing space for dialogue, understanding, and education. Cancel cultures, in their everyday practice, are based on narratives and even if they are media-based - given the massive use of social media - they still have a significant political value (Ewick & Silbey, 1995). This stands valid because narratives, and by extension hate speech, possess a duality that is both potentially subversive, affording a 'voice' to those often marginalized, and hegemonic, expressing ideological influences. However, the prevalence and intricate conceptual ambiguities inherent to the political subculture of cancel culture, coupled with its paradoxical propensity to fuel hate speech, can be attributed in no small part to its propagation through the conduit of social media. The paradoxical impact of hate speech accentuates specific themes, concepts, or mental frameworks while obscuring others, thereby reshaping the societal value hierarchy. It reinforces the values it champions while rendering hollow the significance of those it overlooks. In this context, the concept of 'symbolic annihilation,' initially coined by Gerbner (Gerbner & Gross, 1976), gains clarity. This notion underscores how the absence or underrepresentation of certain events or groups in media becomes a mechanism for perpetuating social inequalities. In the age of digital connectivity and amplified voices, cancel culture's unintended consequences gives rise to a paradoxical phenomenon: the proliferation of hate speech. While cancel culture aims to hold individuals accountable for their actions, its execution has inadvertently created an online environment (which often also transposes to the physical environment) where hate speech can thrive. This contribution

has delved into the intricate interplay between cancel culture and hate speech, shedding light on how these intertwined phenomena impact society at large.

As society evolves within the platform society, where interconnectedness is paramount, the focus must shift toward understanding the dynamics of violence and hate speech both online and offline. Only by addressing the root causes of these issues can we hope to forge a more inclusive, empathetic, and understanding society that fosters change without perpetuating harm. The task at hand is to harness the potential of cancel culture for positive transformation while guarding against the unintended proliferation of hate speech in the digital age.

Striking a delicate equilibrium between upholding individual accountability for their actions and preserving the vibrant diversity of perspectives in the digital era is of paramount importance. This becomes especially crucial due to the fact that the mass media, especially within the realms of social media and online environments, do not merely reflect established societal values. Rather, while refraining from originating new values, they possess the capability to rearrange their hierarchy, strengthening some while draining meaning from others.

While it's acknowledged that there are no definitive frameworks to seamlessly explain the intricate interplay between cultural omissions, media representations, and hate speech, this realization should not lead us to conclude that no such relationship exists. Dismissing the possibility of a connection could potentially impede the advancement of research. The lack of concrete scientific proof shouldn't precipitate a denial of a potential correlation; instead, it should propel researchers towards uncovering the potential causal and interpretational links that might exist within the context.

The inquiry into the plausible paradoxical repercussions of hate speech as a by-product of cancel culture remains an open question, casting itself as a genuine theoretical and methodological challenge. Particularly, scholars of communication will need to squarely confront this challenge in the days to come.

## REFERENCES

- Bartlett, J., Reffin, J., Rumball, N., & Williamson, S. (2014). *Anti-social media*. Demos.
- Castells, M. (1996). *The Rise of the Network Society*. Blackwell.
- Clark, D. M. (2020). Drag Them: A Brief Etymology of So-called "Cancel Culture." *Communication and the Public*, 5(3-4), 88-92.
- Donato, S., Eslen-Ziya, H., & Mangone, E. (2022). From offline to online violence: new challenges for the contemporary society. *International Review of Sociology*, 32(3), 400-412. <https://doi.org/10.1080/03906701.2022.2133405>

- Douthat, R. (2018, February 28). The Rise of Woke Capital. *The New York Times*. <https://www.nytimes.com/2018/02/28/opinion/corporate-american-activism.html>
- Ewick, P., & Silbey, S. S. (1995). Subversive Stories and Hegemonic Tales: Toward a Sociology of Narrative. *Law & Society Review*, 29(2), 197-226.
- Gerbner, G., & Gross, L. (1976). Living with Television: The Violence Profile. *Journal of Communication*, 26(1), 172-199.
- Girard, R. (1986). *The scapegoat* (Y. Freccero, Trans.). Johns Hopkins University Press (Original work published 1982).
- Gualda, E. (2021). Metaphors of invasion: Imagining Europe as endangered by Islamization. In A. Önnfors & A. Krouwel (Eds.). *Europe: Continent of Conspiracies. Conspiracy Theories in and about Europe* (pp.54-75). Routledge.
- Gualda, E. & Rúas, J. (2019). Conspiracy theories, credibility and trust in information. *Communication and Society*, 32(1), 179-193.
- Hughes, E. (2021, April 27). #DragThem: The Toxicity and Problematic Art of Cancel Culture. *Debating Communities and Networks XII*. <https://networkconference.netstudies.org/2021/2021/04/27/dragthem-the-toxicity-and-problematic-art-of-cancel-culture/>
- Lewis, H. (2020, July 14). How capitalism drives cancel culture. *The Atlantic*. <https://www.theatlantic.com/international/archive/2020/07/cancel-culture-and-problem-woke-capitalism/614086/>
- Lizza, R. (2020, July 22). Americans tune in to 'cancel culture' - and don't like what they see. *Politico*. <https://www.politico.com/news/2020/07/22/americans-cancel-culture-377412>
- Loveluck, B. (2020). The Many Shade of Digital Vigilantism. A Typology of Online Self-Justice. *Global Crime*, 21(3-4), 213-241. <https://doi.org/10.1080/17440572.2019.1614444>
- Lumsden, K., & Harmer, E. (Eds.). (2019). *Online Othering. Exploring Digital Violence and Discrimination on the Web*. Palgrave Macmillan.
- Madrid Gil, S. (2023). Woke culture and the history of America: From colonisation to depersonalisation. *Church, Communication and Culture*, 8(1), 18-42. <http://doi.org/10.1080/23753234.2023.2174890>
- Mangone, E. (2022). *Narratives and Social Change. Social Reality in Contemporary Society*. Springer.
- Mortellozzo, E. & Jane, E. A. (Eds.). (2017). *Cybercrime and Its Victims*. Routledge.
- Muller, K., & Schwarz, C. (2018). Fanning the Flames of Hate: Social Media and Hate Crime. *SSRN Electronic Journal*. <https://dx.doi.org/10.2139/ssrn.3082972>
- Ng, E. (2020). No Grand Pronouncements Here...: Reflections on Cancel Culture and Social Media Participation. *Television & New Media*, 21(6), 621-627
- Norris, P. (2020). Closed Minds? Is a 'Cancel Culture' Stifling Academic Freedom and Intellectual Debate in Political Science?. *HKS Working Paper*, RWP20-025. <http://dx.doi.org/10.2139/ssrn.3671026>

- Roos, H. (2020). *With(Stan)ding Cancel Culture: Stan Twitter and Reactionary Fandoms*. Muhlenberg College. <https://jstor.org/stable/community.31638145>
- Sdrigotti, F. (2018). *Shitstorm*. Open pen.
- Slater, S. (2005). The Commodification of Violence on the Internet: An Analysis of 166 Websites Containing Commodified Violence. *Internet Journal of Criminology*. <https://www.internetjournalofcriminology.com/undergraduate-masters-dissertations>
- Tandoc, E. C., Tan Hui Ru, B., Lee Huei, G., Min Qi Charlyn, N., Chua, R. A., & Goh, Z. H. (2022). #CancelCulture: Examining definitions and motivations. *New Media & Society*. <https://doi.org/10.1177/14614448221077977>
- van Dijck J., Poell T., & de Waal M. (2018). *The Platform Society: Public Values in a Connective World*. Oxford University Press.



# AUTOMATIC DETECTION OF HATE SPEECH: RESEARCH CHALLENGES AND OPPORTUNITIES<sup>1</sup>

MARIELBA SILVA DE ZACARIAS

*Professor Auxiliar.*

*Universidade do Algarve. Portugal*

**DOI:** 10.14679/2753

## **Summary:**

1. Introduction.
  2. Challenges and research avenues.
    - 2.1. What is Hate Speech?
    - 2.2. The Contextual Nature of Hate Speech.
    - 2.3. Temporal and Dynamic Nature of Hate Speech.
    - 2.4. A Multilingual and Multicultural problem.
    - 2.5. Dataset and Annotation Challenges.
      - 2.5.1. Data Imbalance.
      - 2.5.2. Data Representativeness.
      - 2.5.3. Annotator Bias and the danger of “Ground Truth”.
  3. The larger picture.
  4. Conclusions.
- References.

With the widespread usage of social media and online communication, the automatic detection of hate speech and conspiracy theories has become a critical area of research. Machine learning (ML), and more recently, Deep Learning (DL) algorithms have been used in addressing this issue with some degree of success. However, numerous challenges and limitations remain to be solved, regarding several aspects of ML/DL such as dataset creation or selection, dataset annotation, feature extraction, and algorithm evaluation. This paper aims at showcasing the complexities of automated hate speech detection, the challenges faced by ML technology in addressing such complexity, and current approaches to overcome them.

---

<sup>1</sup> This chapter is part of the Research Project entitled “Conspiracy theories and hate speech online: comparing patterns in narratives and social networks about COVID-19, immigrants and refugees, and LGBTI people [NON-CONSPIRA-HATE!]”, Project Grant PID2021-123983OB-I00, funded by MCIN/AEI/10.13039/501100011033/ and by “ERDF A way of making Europe”. We are also grateful for the support of the Research Center for Tourism, Sustainability and Well-being (CINTURS) of the University of Algarve.

## 1. INTRODUCTION

Whereas the proliferation of online platforms and social media have enabled global communication and information sharing, it has also facilitated the dissemination of aggressive and harmful content (Matos et al., 2022) including hate speech and conspiracy theories, posing a significant risk to targeted individuals and groups of becoming hate crime victims.

To mitigate such risks, governments are legally binding the owners of social networks to provide content moderation mechanisms (Poletto et al., 2021; Sindhu et al., 2020). Nonetheless, content moderation has proved to be time consuming, labor intensive and an emotionally straining task (Sindhu et al., 2020; Battistelli et al., 2020; Fortuna et al., 2022; Parker and Ruths, 2023). Automated detection offers several advantages over manual detection; (a) it is much faster in detecting hate speech than humans, (b) it accomplishes the task with greater quantities of data, and (c) it alleviates content moderators from the emotional burden of having to read vast amounts of hateful contents (Parker and Ruths, 2023).

It is precisely the unprecedented amount of user-generated content provided by social media that has given researchers and practitioners the opportunity of exploring machine learning (ML) and in particular, Deep Learning (DL) algorithms for this task. These algorithms are supervised learning techniques generally involve the following processing pipeline: i) defining the problem, ii) collecting, sampling and annotating datasets with labels, iii) extracting features, iv) applying algorithms and v) evaluating according to given metrics and techniques (Fortuna et al., 2022; Jahan & Ossalah, 2023). Thereafter, the efficacy of these algorithms depends on the quality of the datasets, the techniques used for annotation, extracted features and the predictive capability of the algorithms themselves. Aiming at showcasing the complexity of automated hate speech detection, this paper highlights the limitations and challenges regarding different aspects of automated hate speech using ML/DL algorithms that are being discussed in the literature, as well as some current approaches seeking to overcome them.

## 2. CHALLENGES AND RESEARCH AVENUES

### 2.1. What is Hate Speech?

The existence of several definitions of hate speech and terms has been widely acknowledged by researchers (Sindhu et al., 2020; Schmidt & Wiegand, 2017; Paz et al., 2020; Martins et al., 2018; Wazeem & Hovy, 2016). The lack of a universal definition poses important challenges for fields studying the subject (Delgado and Stefancic, 1995, as cited in Paz et al. 2020, p.1). While some definitions merely refer to negative sentiments towards a group of people without specifying any particular type, such as the definition provided in

Delgado & Stefancid, (1995), where hate speech is defined as a “*willful public statement intended to denigrate a group of people*”, others specify the identifying characteristics of the targeted groups. As an example, Fortuna and Nunes (2018, p. 5) define hate speech as “*the content that promotes violence against individuals or groups based on race or ethnic origin, religion, disability, gender, age, veteran status, and sexual orientation/gender identity*”. This shows that different definitions lead to different characterizations of the term. Consequently, automatic detection techniques are also contingent to the definition employed that may result in different classifications of the same text, hindering the comparison of different approaches (Sindhu et al., 2020). Papcunová et al. (2023) seek to address this problem by operationalizing hate speech. The authors propose an empirical definition by providing a list of 10 hate speech indicators and the rationale behind them. The indicators are observable and measurable characteristics that offer a practical definition of hate speech and can be a valuable tool in developing comparable hate speech detection methodologies and techniques.

## 2.2. The Contextual Nature of Hate Speech

A fundamental challenge in hate speech detection is the ambiguity and context sensitivity of language. Since hate speech operates at the pragmatics level of communication (Parvarish, 2023), hate speech is manifested verbally, non-verbally and symbolically (Nielsen, 2002, as cited in Paz et al. 2020, p.1). Verbal expressions can be encoded in various ways, including subtle linguistic cues, sarcasm, satire, cultural references, and coded language, all of which can be used to mask hate speech. Since context plays a crucial role in determining whether a statement is genuinely harmful, developing models with a deeper and more comprehensive understanding of the contextual information embedded in online content is crucial. Natural Language Processing techniques (NLP), including contextual embeddings and transformer models able of capturing and leveraging context, show promise in capturing nuances and covert meanings and intentions. Some feature extraction methods play a crucial role in this regard (Schmidt & Wiegand, 2017). Among these are methods for extracting linguistic features that consider deeper syntactic information, features including domain and/or world knowledge, meta-information (e.g. user and group profiles), emotional features (Martins et al., 2018; Chiril et al., 2022), sociological features (Nagar et al., 2023), and multimodal information including images, audio and video (Chhabra & Vishwakarma, 2023).

## 2.3. Temporal and Dynamic Nature of Hate Speech

Online communities often create and adopt their own slang and terminology (Florio et al., 2020). Moreover, hate speech is not static; it evolves alongside

language and according current events and trends. ML algorithms struggle to keep up with these changes, leading to false negatives, thus forcing them to be subject of continuous retraining and adaptation to maintain their performance and accuracy. Annotations of hate speech may also evolve over time, potentially leading to outdated training data.

The temporality of datasets and annotations requires research on the temporal robustness of ML algorithms (Florio et al., 2020), temporal analysis to uncover the evolution of hate speech intensity and hateful users in time (Binny et al., 2020). Other potential lines of research may include the development of detection systems capable of identifying and responding to harmful content in real-time, and adaptive systems capable of quickly learning new forms of hate speech to reduce the lag time for model updates as the ones proposed by Agarwal and Chowdary (2021) and Ludwig et al. (2018).

#### **2.4. A Multilingual and Multicultural problem**

Hate speech is not limited to a single language or culture. Each culture has its own interpretation of hate speech, leading to different ways of reacting to it (Al-Hassan & Al-Dossari, 2019). Detecting such content in multilingual and multicultural contexts requires the development of annotation methods and models capable of covering a wide range of languages and cultural nuances. This diversity further complicates the task. While there is some research covering multiple languages (Al-Hassan & Al-Dossari, 2019 and Arango et al, 2022), many datasets used in research continue to be dominated by English-language content (Maronikolakis et al., 2023). Therefore, research on the creation multilingual and multicultural datasets is essential in order to train models that can effectively detect hate speech and conspiracy theories across different languages and cultural contexts.

Other example of multilingual and multicultural research is HATELEXICON (Maronikolakis et al., 2023), a lexicon of slurs and target group denotations developed as an aid to model training and interpretability for Brazil, Germany, India and Kenya. The NON-CONSPIRA-HATE! Project (Ref. PID2021-123983OB-I00) seeks to develop methodologies and techniques to detect hate speech and conspiracy theories in 4 languages (Spanish, Italian, Portuguese, and English). The project, integrated by researchers from Spain, Italy, Portugal, Argentina and Ecuador, is yet another example of multilingual and multicultural research on this topic.

## 2.5. Dataset and Annotation Challenges

The vast volume of online content, the disproportion between hate and innocuous content, and the subjective nature characterizing data collection and annotation methods poses important challenges and limits to the applicability and generalizability of the proposed solutions. The following sections summarize and discuss the issues of data imbalance, lack of data representativeness, data and annotation biases, as well as the challenges posed by the requirement of a “ground truth”.

### 2.5.1. *Data Imbalance*

Comparing to the vast amount of non-offensive content, a dataset for hate speech detection is that hate speech is relative rare, rendering random sampling of text to annotate highly inefficient and leading to content selection bias. This imbalance can affect the performance of ML algorithms, as they may become biased towards the majority class, resulting in high false-negative rates. Some authors address this issue filtering texts with predefined “hate” vocabulary as shown in (Maronikolakis et al., 2023). However, this approach restricts the dataset and may filter out relevant samples, also introducing biases in the data. Data augmentation techniques seek to improve the quality of existing datasets by generating synthetic samples that mimic the distribution of original samples (Madukwe et al., 2022; Rizos et al., 2019). Information Retrieval techniques have also been applied as a way building balanced datasets, as well as improving the consistency and value of annotations (Rahman et al., 2021).

### 2.5.2. *Data Representativeness*

Most datasets used to train and evaluate hate speech detection models are not fully representative of the online content landscape. Biases in data collection methods or the focus on specific topics or platforms can limit the generalizability of models when applied to other dataset collections. Research on cross-platform datasets may help in overcoming this issue. Mody et al. (2023) describe an extensive curated dataset comprising emojis, emoticons, and contractions bundled into two classes, hateful and not-hateful. The raw data was collected from 18 sources and contractions data from 8 sources. Exploring the use of transfer learning techniques to adapt models trained on one domain to other domains may improve their generalizability. Ludwig et al. (2022) approach the improvement of hate speech systems generalization capabilities to novel target groups, through domain adaptation.

### 2.5.3. *Annotator Bias and the danger of “Ground Truth”*

The task of annotating consists in labeling dataset items by human annotators. Hence, it is by nature, a subjective task where the background, age and experience of annotators may lead to different interpretations of the same content and inconsistencies in the labeled datasets (Fortuna et al., 2022). Thus, techniques for reducing annotation biases should be developed and integrated into the hate speech detection processing pipeline. Incorporating “fairness” in ML entail identifying and rectifying biases in hate speech detection models, collaborating with experts in algorithmic biases, assessing the real-world impact of detection algorithms on different demographic groups, and making necessary adjustments to minimize disparate effects.

Essentially, “ground truth” refers to facts assumed to be true by an ML system (Kang, 2023). As the complexity of tasks performed through ML increases, agreement of what constitutes adequately stable ground truths has become more complicated. The lack of a universal definition and variation of hate speech across regions and cultures means that there is no universal ground truth, hindering the establishment of a gold standard for annotation and the generalization capabilities of ML algorithms. Developing annotation frameworks allowing explaining the annotators’ decisions will certainly help users to understand why certain content is flagged. Binny et al. (2021) propose a benchmark dataset to enable model explainability by jointly learning the target classification and the rationale for the human decision. Fortuna et al (2019) proposes a hierarchical labeling framework built by a researcher trained in social psychology that allows to understand the labeling rationale, helps to reducing annotation biases, which improves significantly the agreement among annotators and foster a more stable ground truth. Another way of helping to understand labeling rationales, the ground truth subjacent to each model and to better interpret model results, is to develop models capable of informing users about potential prediction biases.

## 3. THE LARGER PICTURE

In spite of its benefits, “*automated detection is just a small component of the much larger challenge of online hate speech*” (Parker & Ruth, 2023, p.2). In order to devise more comprehensive solutions to the problem of online hate speech, computer scientists need to work with other stakeholders. Researchers need to collaborate with interdisciplinary teams to properly answer technical questions. Furthermore, researchers and practitioners need to address *social, financial, legal, and ethical* questions.

Regarding technical issues, seeking collaborations with linguistics and NLP experts will allow to gain a deeper comprehension of the linguistic nuances of hate speech. This collaboration can help in developing more linguistically-informed models and annotation guidelines. Social scientists and psychologists are essential in identifying and characterizing emotional and sociological features to design more effective annotation frameworks and detection algorithms. Annotation quality can be improved (1) through an adequate combination of ML technology and human expertise and (2) active learning approaches. As an example, Ahmed and Lin (2022) use active learning to expand lexicons with term synonyms, resulting in increasing labeled data and improving accuracy. Active learning can also enable ML model adaptation to evolving language and content. The participation of expert reviewers into the annotation process can be a valuable aid in improving dataset quality and reducing annotation biases. User participation fosters the accuracy and adaptiveness of detection systems through the implementation of reinforcement learning, which include feedback mechanisms where users are able to report false positives and false negatives.

It is also crucial to recognize the importance of linguistic cultural diversity in hate speech and the need to work with experts in various languages and cultures to create inclusive models and datasets, and to conduct cross-cultural studies to understand how hate speech is manifested in various regions and communities.

From a social perspective, collaborations with educators and online media experts are required to develop educational programs in topics such as critical thinking and information assessment for individuals and communities (United Nations, 2020; UNESCO, 2023). This type of collaboration is also necessary to promote online civics education programs to encourage responsible and respectful online behavior. Global collaborations entail creating international partnerships in order to mitigate hate speech that transcend borders. Sharing research findings, datasets, and experiences with the global research community is the way to the multilingual and multicultural aspect of hate speech. It is also important to work toward the development of international standards and guidelines for hate speech detection.

From a legal perspective, the development of legal frameworks such as *EU Code of Conduct on countering illegal hate speech online*, is essential to mitigate hate speech online (European Commission, 2016) and punish hate crimes. Both ongoing research and development and the aforementioned educational programs, require investment of financial resources. Finally, collecting and sharing potentially harmful content for research purposes raises ethical concerns and may violate the privacy of individuals whose content is included in datasets. In this regard, the involvement of ethicists is crucial to balance the need for data with ethical considerations and to ensure adherence to ethical guidelines and standards need to be ensured.

#### 4. CONCLUSIONS

Automatic detection of hate speech and conspiracy theories is a complex and evolving field with numerous challenges and research opportunities. All the components of the ML processing pipeline, including algorithms, feature engineering, annotation techniques, and dataset creation, play pivotal roles in addressing these challenges and exploring the research opportunities provided by ML. Nonetheless, the complexity of hate speech detection and mitigation means that technological solutions are not enough and require collaboration between researchers, policymakers, industry stakeholders, and civil social organizations. While there is no universal solution, building more effective and responsible hate speech detection systems, require holistic, comprehensive and adaptive approaches that cannot disregard the inherent complexity of the topic, the value of human participation in the process, the evolving nature of online content, and the ethical considerations surrounding this critical issue.

#### REFERENCES

- Agarwal, S., & Chowdary, C.R. (2021). Combating hate speech using an adaptive ensemble learning model with a case study on COVID-19. *Expert Systems with Applications*, 185, 115632. <https://doi.org/10.1016/j.eswa.2021.115632>
- Ahmed, U., & Lin, J.C-W. (2022). Deep Explainable Hate Speech Active Learning on Social-Media Data. *IEEE Transactions on Computational Social Systems*. <https://doi.org/10.1109/TCSS.2022.3165136>
- Al-Hassan, A., & Al-Dossari, H. (2019). Detection of hate speech in social networks: a survey on multilingual corpus. In *6th international conference on computer science and information technology*, 10, pp. 10-5121.
- Arango Monnar, A.; Perez, J., Poblete, B.; Saldaña, M., & Proust, V. (2022). Resources for Multilingual Hate Speech Detection. In *Proceedings of the Sixth Workshop on Online Abuse and Harms (WOAH)*, pages 122-130, Seattle, Washington (Hybrid). Association for Computational Linguistics. <https://aclanthology.org/2022.woah-1.12/>
- Battistelli, D.; Bruneau, C., & Dragos, V. (2020). Building a formal model for hate detection in French corpora. *Procedia Computer Science*, 176, pp. 2358-2365. <https://doi.org/10.1016/j.procs.2020.09.299>
- Chhabra, A., & Vishwakarma, D.K. (2023). A literature survey on multimodal and multilingual automatic hate speech identification. *Multimedia Systems*, 29, pp.1203-1230. <https://doi.org/10.1007/s00530-023-01051-8>
- Chiril, P., Pamungkas, E.W., Benamara, F. Moriceau, V., & Patti, V. (2022). Emotionally Informed Hate Speech Detection: A Multi-target Perspective. *Cognitive Computation*, 14, pp.322-352. <https://doi.org/10.1007/s12559-021-09862-5>

- Delgado, R., & Stefancid, J. (1995). Images of the outsider in American law and culture: Can free expression remedy systemic social ills? (pp. 217-227). In Delgado, R. (ed.). *Critical Race Theory: The cutting edge*. Temple University.
- European Commission (2016). *The EU Code of Conduct on Countering Hate Speech*. [https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online\\_en](https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en)
- Fortuna, P., & Nunes, S. (2018). A Survey on Automatic Detection of Hate Speech in Text. *ACM Comput. Surv.* 51(4), article 85, 30 pages. <https://doi.org/10.1145/3232676>
- Fortuna, P.; Rocha da Silva, J.; Soler-Company, J.; Wanner, L., & Nunes, S. (2019). A Hierarchically-Labeled Portuguese Hate Speech Dataset (pp.94-104). In *Proceedings of the Third Workshop on Abusive Language Online*. Florence, Italy. Association for Computational Linguistics. <https://aclanthology.org/W19-3510/>
- Fortuna, P.; Dominguez, M.; Wanner, L., & Talat, Z. (2022). Directions for NLP Practices Applied to Online Hate Speech Detection (pp.11794-11805). In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. <https://aclanthology.org/2022.emnlp-main.809/>
- Jahan, Md.S., & Oussalah, M. (2023). A systematic review of hate speech automatic detection using natural language processing, *Neurocomputing*, 546, 126232. <https://doi.org/10.1016/j.neucom.2023.126232>.
- Kang, E. B. (2023). Ground truth tracings (GTT): On the epistemic limits of machine learning. *Big data & society*, 10(1), 20539517221146122.
- Ludwig, F.; Dolos, K.; Zesch, T., & Hobley, E. (2022). Improving Generalization of Hate Speech Detection Systems to Novel Target Groups via Domain Adaptation. In *Proceedings of the Sixth Workshop on Online Abuse and Harms (WOAH)*, pages 29-39, Seattle, Washington (Hybrid). Association for Computational Linguistics. <https://aclanthology.org/2022.woah-1.4/>
- Madukwe, K.J.; Gao, X., & Xue, B. (2022). Token replacement-based data augmentation methods for hate speech detection. *World Wide Web*, 25, pp.1129-1150. <https://doi.org/10.1007/s11280-022-01025-2>
- Maronikolakis, A.; Koksai, A., & Schutze, H. (2023). Sociocultural knowledge is needed for selection of shots in hate speech detection tasks. <https://arxiv.org/abs/2304.01890>
- Martins, R.; Gomes, M.; Almeida, J.J.; Novais, P., & Henriques, P. (2018). Hate Speech Classification in Social Media Using Emotional Analysis (pp.61-66). *7th Brazilian Conference on Intelligent Systems (BRACIS)*, Sao Paulo, Brazil. <https://doi.org/10.1109/BRACIS.2018.00019>
- Mathew, B., Illendula, A., Saha, P., Sarkar, S., Goyal, P., & Mukherjee, A. (2020). Hate begets Hate: A Temporal Study of Hate Speech. *Proc. ACM Hum.-*

- Comput. Interact.* 4, CSCW2, Article 92 (October), 24 pages. <https://doi.org/10.1145/3415163>
- Mathew, B.; Saha, P.; Yimam, S. M.; Biemann, C.; Goyal, P., & Mukherjee, A. (2021). Hatexplain: A benchmark dataset for explainable hate speech detection. In *Proceedings of the AAAI conference on artificial intelligence*, 35(17), pp. 14867-14875.
- Matos, B.C.; Santos, R.B.; Carvalho, P.; Ribeiro, R., & Batista, F. (2022). Comparing Different Approaches for Detecting Hate Speech in Online Portuguese Comments. In Cordeiro, J.; Pereira, M.J.; Rodrigues, N.F., & Pais, S. *11th Symposium on Languages, Applications and Technologies (SLATE 2022)*. <https://doi.org/10.4230/OASiCS.SLATE.2022.10>
- Mody, D.; Huang, Y., & Alves de Oliveira, T. E. (2023). A curated dataset for hate speech detection on social media text. *Data in Brief*, 46, 108832. <https://doi.org/10.1016/j.dib.2022.108832>.
- Nagar, S.; Barbhuiya, F.A., & Dey, K. (2023). Towards more robust hate speech detection: using social context and user data. *Soc. Netw. Anal. Min*, 13, 47. <https://doi.org/10.1007/s13278-023-01051-6>
- Nielsen, L. (2002). Subtle, pervasive, harmful: Racist and sexist remarks in public as hate speech. *Journal of Social Sciences*, 58 (2), pp. 265-280.
- Papcunová, J.; Martončík, M.; Fedáková, D. *et al.* (2023). Hate speech operationalization: a preliminary examination of hate speech indicators and their structure. *Complex Intell. Syst.*, 9, 2827-2842. <https://doi.org/10.1007/s40747-021-00561-0>
- Parker, S., & Ruths, D. (2023). Is hate speech detection the solution the world wants?. *Proceeding of the National Academy of Sciences (PNAS)*, 120 (10), e2209384120, <https://doi.org/10.1073/pnas.2209384120>
- Parvaresh, V. (2023). Covertly communicated hate speech: A corpus-assisted pragmatic study. *Journal of Pragmatics*, 205, pp. 63-77. <https://doi.org/10.1016/j.pragma.2022.12.009>
- Paz, M.A.; Montero-Díaz, J., & Moreno-Delgado, A. (2020). Hate Speech: A Systematized Review. *SAGE Open*, 10(4). <https://doi.org/10.1177/2158244020973022>
- Poletto, F.; Basile, V.; Sanguinetti, M. *et al.* Resources and benchmark corpora for hate speech detection: a systematic review (2021). *Lang Resources & Evaluation*, 55, pp.477-523. <https://doi.org/10.1007/s10579-020-09502-8>
- Rahman, M.M.; Balakrishnam, D.; Murthy, D.; Kutlu, M., & Lease, M. (2021). An Information Retrieval Approach to Building Datasets for Hate Speech Detection. <https://arxiv.org/abs/2106.09775>
- Rizos, G.; Hemker, K., & Schuller, B. (2019). Augment to Prevent: Short-Text Data Augmentation in Deep Learning for Hate-Speech Classification. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM '19)*. Association for Computing Machinery, New York, NY, USA, 991-1000. <https://doi.org/10.1145/3357384.3358040>

- Sindhu Abro, Sarang Shaikh, Zahid Hussain Khand, Zafar Ali, Sajid Khan, & Ghulam Mujtaba (2020). Automatic Hate Speech Detection using Machine Learning: A Comparative Study. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 11(8). <http://dx.doi.org/10.14569/IJACSA.2020.0110861>
- Schmidt, A., & Wiegand, M. (2017). A Survey on Hate Speech Detection using Natural Language Processing (pp. 1-10). In *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*. Valencia, Spain. Association for Computational Linguistics. <https://aclanthology.org/W17-1101/>
- UNESCO (2023). *Countering Hate Speech*. <https://www.unesco.org/en/countering-hate-speech?hub=70287>
- United Nations. (2020). *United Nations Strategy and Plan of Action on Hate Speech*. [https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20PoA%20on%20Hate%20Speech\\_Guidance%20on%20Addressing%20in%20field.pdf](https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20PoA%20on%20Hate%20Speech_Guidance%20on%20Addressing%20in%20field.pdf)
- Waseem, Z., & Hovy, D. (2016). Hateful symbols or hateful people? Predictive features for hate speech detection on Twitter (pp.88-93). In *Proceedings of the NAACL student research workshop*.



**II.**  
**COVID-19, TEORÍAS DE LA**  
**CONSPIRACIÓN, INFODEMIA**  
**Y ODIO EN LÍNEA**



# MOVILIZACIONES DURANTE LA PANDEMIA DE COVID-19 EN ARGENTINA. UNA EXPLORACIÓN DE MENSAJES Y USUARIXS EN TWITTER DURANTE LOS “BANDERAZOS” (MAYO-OCTUBRE 2020)<sup>1</sup>

WANDA JUARES

*Becaria doctoral. INHUS - CONICET - CEHis  
Universidad Nacional de Mar del Plata. Argentina*

SILVANA FERREYRA

*Investigadora Adjunta. INHUS - CONICET - CEHis  
Universidad Nacional de Mar del Plata. Argentina*

DOI: 10.14679/2754

## Sumario:

1. Introducción.
  2. Antecedentes.
  3. Construcción del corpus.
  4. El contenido de las protestas.
  5. Tres miradas sobre los usuarixs.
  6. Reflexiones finales.
- Referencias.

## 1. INTRODUCCIÓN

La pandemia como evento crítico mundial tuvo aspectos compartidos por todos los habitantes del planeta, tales como el temor a la enfermedad, el confinamiento, la mortalidad, la crisis económica y la experiencia de atravesar un peligro sanitario de magnitud que, por primera vez en la historia, se afronta en una sociedad hiperconectada (Hao y Basu, 2020). Ante este escenario los gobiernos tuvieron distintas respuestas de gestión en relación a sus recursos y posicionamientos políticos. Aun así, expresiones de descontento social tomaron las calles con repertorios de protestas y manifestaciones en contra de la gestión de la pandemia, desafiando las medidas de confinamiento. Sobre esta caracterización, diversos autores indagan si las condiciones socioeconómicas de cada país, los

---

<sup>1</sup> Este capítulo forma parte del Proyecto I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGBTI [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDER A way of making Europe”. También agradecemos el apoyo del Instituto de Humanidades y Ciencias Sociales de CONICET y la Universidad Nacional de Mar del Plata.

fallecimientos ocasionados por la COVID-19 y las medidas sanitarias, han tenido relación o no con la intensidad de las acciones de protesta (Koen van der Zwet et al., 2022; Hanspeter y Ioana-Elena, 2023). En este contexto las plataformas de redes sociales cobraron aún mayor relevancia y configuraron el espacio público digital, escenario de posicionamientos ideológicos y, circulación de narrativas en torno a tópicos de distinta índole organización y activismo digital.

Asimismo, la proliferación de (des)información sobre el virus, el debate público sobre las medidas de política sanitaria y el confinamiento articularon sentidos que amalgamaron el descontento social. En esta coyuntura los procesos democráticos mostraron cierta fragilidad a partir de la reconfiguración en el espacio público de movimientos de derechas extremas, asociados a discursos negacionistas, teorías conspirativas y el incremento de la violencia política en el espacio público digital (Imhoff et al., 2022; Parodi et al., 2022).

En este capítulo nos proponemos realizar una primera aproximación a una serie de protestas que se articularon en Argentina a partir de Twitter, una plataforma conocida por su capacidad de amplificación de los problemas público-políticos (Van Dijck, 2016), durante los primeros meses de la pandemia de COVID-19. Nos ocuparemos de analizar, en una mirada conjunta, las movilizaciones del 25 de Mayo (día de la revolución de mayo), 20 de junio (día de la bandera), 9 de julio (día de la independencia), 17 de agosto (paso a la inmortalidad del General San Martín) y 12 de Octubre (día de la diversidad cultural), en tanto hechos masivos acontecidos durante el aislamiento social, preventivo y obligatorio decretado por el gobierno de Alberto Fernández<sup>2</sup>.

Trabajaremos en dos ejes para obtener una primera mirada cuantitativa de esta serie de manifestaciones. En primer lugar ¿cuáles fueron las principales narrativas que articularon estas movilizaciones? En esa línea, buscaremos dar cuenta de los principales hashtags que funcionaron como ensamblajes discursivos en estos eventos tecnosociales (Rambukanna, 2015). Este estudio y un ejercicio de modelado de tópicos nos permitirá aproximarnos a las consignas que defendieron y rechazaron estos eventos que se desarrollaron entre las redes y en las calles. En particular, buscaremos identificar en qué medida emergieron discursos de odio en los mensajes.

En segundo lugar, nos proponemos indagar en los usuarios que participaron de estos eventos. Proponemos tres vías de exploración. Una más tradicional,

---

<sup>2</sup> Alberto Fernández asumió como presidente en diciembre de 2019, como candidato de una coalición compuesta por distintos sectores del peronismo, entre los que ocupaba un lugar de relevancia el kirchnerismo, representado por la propia Cristina Fernández de Kirchner como candidata a la vicepresidencia. Su triunfo electoral marca la derrota de Mauricio Macri, presidente entre 2015 y 2019, quien se transformó en uno de los principales líderes de la oposición, articulada en Juntos por el Cambio. A su vez, el triunfo de Macri en 2015 había significado la primera derrota del kirchnerismo en elecciones presidenciales desde 2003.

guiada por aquellos que obtuvieron más retweets y favs. Otra que busca sacar el foco de los influencers, para identificar las características del grupo de cuentas que se mantienen estables a lo largo de los distintos eventos. Por último, ensayaremos un análisis de la localización de estas cuentas para indagar en torno al carácter federal de estas iniciativas y reflexionar sobre los territorios de co-presencia entre lo presencial y lo virtual.

## 2. ANTECEDENTES

Sobre la pregunta acerca de las manifestaciones durante la pandemia y su configuración en el espacio digital, mencionaremos solo algunos textos de una literatura que ha crecido de manera explosiva, y seleccionaremos aquellos más relevantes para los interrogantes que aquí buscaremos dilucidar. En ese sentido, un grupo de autores abordaron las protestas en E.E.U.U asociadas a los movimientos #Liberarte (Karami y Anderson, 2020; Haupt et al., 2021) y al atentado del Capitolio como efectos de desplazamientos dentro y fuera de la red (Vishnuprasad et al., 2023). Estos trabajos revelan estrategias del activismo de las derechas a partir de la coordinación de narrativas mediante hashtags, utilización de bots y la proliferación y amplificación coordinada de contenidos conspirativos. En Alemania analizan los movimientos de extrema derecha denominados Querdenken y Q-Anon (Jarynowski et al., 2020). Otras investigaciones sobre Australia (Graham et al., 2021) y Tailandia (Charoenthansakul y Natee, 2023) coinciden en que las protestas tienen consignas contra la corrupción y de corte antigubernamental, aunque sus posicionamientos político-partidarios no resultan tan marcados. Los distintos trabajos dan cuenta que las manifestaciones propiciaron la configuración de escenarios polarizados en relación con posiciones ideológicas-políticas sobre la crisis sanitaria, incluso en países como Brasil, donde el propio gobierno defendía posturas negacionistas respecto a la pandemia (Paulino y Waisbord, 2021).

Nos interesa pensar las redes como espacio global de conexión, que permiten la articulación de distintas comunidades, demandas, narrativas y repertorios de acción y adoptan distintas características de acuerdo al entrecruzamiento con disputas nacionales. El estudio de la red no se limita a una plataforma, ni a un contexto, sino que para aproximarnos a la sociabilidad en las redes adoptamos una perspectiva situada y comparada que de cuenta de las diferencias culturales y políticas que se imprimen en las tecnologías para cada región (Coleman, 2010). Para ello nos proponemos trazar una mirada local que recupere las características de la cultura política argentina en un espacio global como Twitter, observando los colectivos y las identidades políticas antinómicas en los polos kirchnerismo-antikirchnerismo, peronismo-antiperonismo, populismo-antipopulismo (Vommaro, 2017; Ferreyra, 2022; Canelo, 2019; Semán 2021; Calvo y Aruguete, 2020).

En Argentina el año 2020 fue escenario de manifestaciones ciudadanas denominadas en la escena mediática como *anticuarentena* (Coraita y Alvarez, 2021). En esa línea, algunos autores contabilizan que durante la pandemia acontecieron 33 hechos de protesta vinculadas con las demandas político-ideológicas de sectores de clase media y alta refractarios a las políticas del gobierno (Soto Pimentel, Gradín y Reiri, 2021). Morresi, Saferstein y Vicente (2020) las describen como expresiones de las derechas políticas del país, aunque sin dejar de mencionar la presencia de ciudadanos independientes que acudieron con reclamos particulares, como los comerciantes. Por su parte, Gamboa Troyano y Beccia (2021) sostienen que la pandemia abrió una ventana de oportunidad para las fuerzas políticas de derecha, encuadrando un conjunto de demandas heterogéneas bajo una supuesta defensa de la libertad.

Aunque la ocupación de las calles tuvo un lugar relevante en el marco del aislamiento preventivo, las redes sociales aparecen como escenarios de visibilización y coordinación de las manifestaciones. Esta articulación tiene como primeros antecedentes en la Argentina la serie de protestas opositoras al gobierno de Cristina Kirchner, denominadas “cacerolazos” entre los años 2012-2013. Gold (2017), quien analiza el uso inicial de redes sociales digitales como herramientas de difusión de estos acontecimientos, nota un quiebre en la clásica organización de la protesta, donde los medios digitales rompen la lógicas organizacionales y partidarias para derivar en una organización informal de los eventos. Ante la irrupción de las redes sociales en el activismo político, el autor menciona que los marcos de sentido de las protestas ya no están determinados sólo por largas luchas históricas, sino por problemáticas que involucran a públicos masivos, donde los hechos aparecen como estallidos fugaces y son las redes las que permiten la rápida difusión de los acontecimientos.

En este sentido, las conmemoraciones nacionales aparecen como un momento propicio para reaccionar en “defensa de la patria” y develar las fisuras irreparables de la grieta (Sigal, 2006). A diferencia de lo acontecido en otras latitudes, la elección de las fechas no parece responder a criterios epidemiológicos. Plumper et al. (2021) muestran para Alemania que los organizadores de la protesta promueven las marchas cuando las tasas de incidencia y mortalidad eran bajas, pero las medidas de contención seguían siendo relativamente estrictas. En Argentina, por el contrario, la curva de fallecimientos empieza a ascender a fines de junio y alcanza el pico máximo de la primera ola en octubre de 2020, mientras la movilización crece paralelamente<sup>3</sup>. De hecho, las marchas se suceden tras un primer momento de aceptación de las medidas de gobierno y baja circulación del virus, aunque ya en Twitter habían aparecido las primeras tendencias polarizantes (Nieto y Ferreyra, 2020).

---

<sup>3</sup> Para más información sobre la curva epidemiológica en Argentina puede consultarse <https://coronavirus.jhu.edu/region/argentina>

### 3. CONSTRUCCIÓN DEL CORPUS

Esta investigación toma a Twitter como objeto y fuente de estudio. En tanto objeto digital, consideramos la incorporación de métodos y herramientas computacionales como parte del giro digital para las ciencias sociales (Rogers, 2013). Estos procedimientos permiten trabajar con grandes volúmenes de datos extraídos de las redes sociales (Gualda, 2020), y aproximarnos a una comprensión del entramado sociotécnico que gobierna las plataformas. El enfoque de este estudio es realizado con técnicas cuantitativas, principalmente mediante el lenguaje de programación R y su interfaz Rstudio.

Nos preguntamos por la serie de manifestaciones acontecidas durante el año 2020 con el objetivo de explorar narrativas y discursos, los distintos tipos de internautas y la topología de la red para el conjunto de los eventos en Twitter. Nos enfocamos principalmente en la observación de interacciones digitales: tweets, rtweets, hashtags y comunidades de usuarios.

El trabajo con métodos digitales conlleva un proceso de etapas reconocidas como extracción, limpieza, procesamiento y visualización de los datos. Sobre el scraping, o extracción de los datos, utilizamos la API v2 académica de Twitter. Mediante este procedimiento obtenemos datos y metadatos semi-estructurados sobre tweets y usuarios. Mucho se ha discutido sobre la calidad de la información que proporcionan las API, al considerar que puede haber sesgos, restricciones e inconsistencias en los datos recabados (Bruns y Burgess, 2016). Para revisar estas limitaciones y pensar cómo sortearlas realizamos descargas en distintas “etapas”. En la primera realizamos búsquedas para identificar los hashtags centrales a partir de la prensa, el archivo histórico de trending topics<sup>4</sup>, y la observación de perfiles claves de autoridades en la red. A partir de esta lista extrajimos un conjunto de tweets y confeccionamos un primer dataset que contenía columnas con información sobre el texto de los mensajes, fecha de creación, id, usuario, ubicación, retweets, favs, entre las variables principales. Tras efectuar un ranking con los principales hashtags, elaboramos una segunda base en la que incluimos aquellas etiquetas que no habían sido consideradas en la primera oportunidad<sup>5</sup>. El corte temporal de la extracción de la muestra responde a la observación de “agitamiento” en la red que nos permite comparar la incidencia del evento cinco días antes y cinco días posteriores al mismo.

---

<sup>4</sup> Twitter Trending Archive, Argentina Twitter trends, <https://archive.twitter-trending.com/argentina>

<sup>5</sup> Cuando nos encontrábamos realizando el trabajo de reconstrucción del corpus Twitter modificó sus políticas de acceso a la información y anuló nuestra API v2 académica. Por esta razón, la base del evento del 25 de Mayo no pudo ser reconstruida como las demás y tiene un tamaño considerablemente menor.

Tabla 1. Listado de hashtags extraídos para cada movilización

BASE	HASHTAGS	INICIAL	FILTRADO
25M	25M OR plazademayo OR QueSeVayanTodos OR CacerolazoNacional OR QuedateenCasa	491771	20792
20J	20J OR BanderazoNacional OR BanderazoFederal OR el20todosalobelisco OR Caravana20J OR 20jyovoy OR banderazo20j OR 20jsalimosalascalles	245593	237064
9J	9J OR Banderazo OR 9JYoVoy OR 9JBanderazoPorLaRepublica	249872	246329
17A	17A OR 17AJuntosContraLaImpunidad OR 17ABanderazoPorLaLibertad OR CaravanaPorLaRepublica OR 17ASalimosTodos OR 17AYoFui OR YoNoMarcho OR LaMarchaDeLosHDP OR 17AContagiarse OR RompanlaCuarentena	780255	734127
12O	12O OR 12OctubreTodosALaCalle OR 12OSomosLibres OR 12OYoVoy OR 12OBanderazoPorLaRepublica	321839	298093

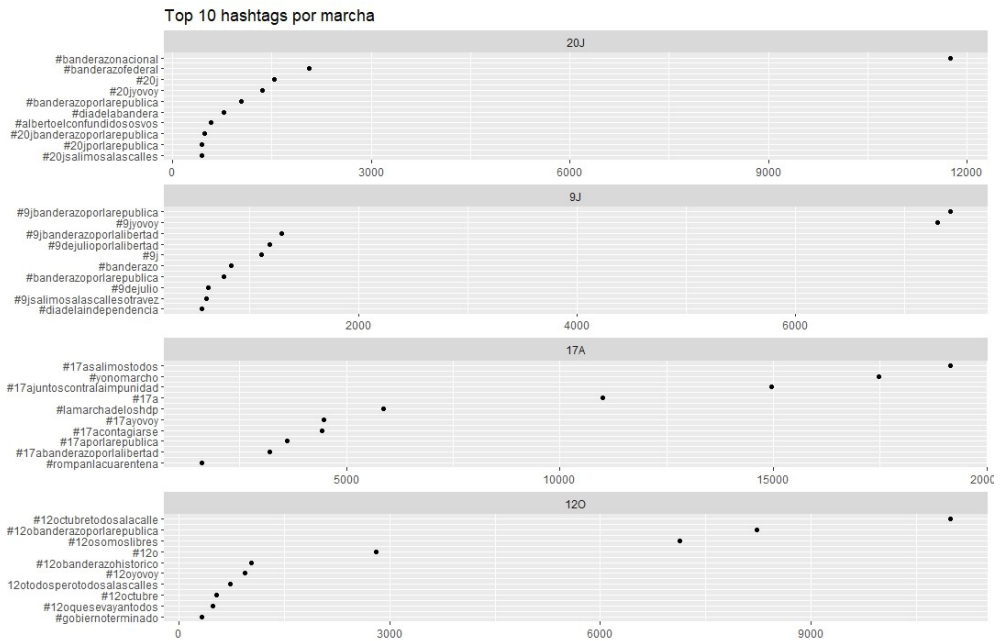
Fuente: elaboración propia. Para la descarga de tweets utilizamos el paquete *AcademicTwiiterR* (Barrie & Ho, 2021).

En relación a la limpieza o curaduría de los datos, convertimos el dataset y lo leímos en Excel, seleccionando sólo algunas columnas (Id, hashtags, user). Una mirada cercana nos permitió eliminar los tweets que no tienen relación con los casos analizados, tales como protestas de Ecuador o España que utilizaban los mismos hashtags; spam; mensajes conmemorativos sobre las efemérides, pero sin relación con las protestas y otros tweets residuales. El producto final es un dataset de 1.536.405 tweets y retweets, que se transforma en un corpus de 134.714 cuando eliminamos los retweets.

#### 4. EL CONTENIDO DE LAS PROTESTAS

Como señalamos en el apartado anterior, nuestro corpus está conformado a partir de una selección de hashtags, por lo cual un análisis de las principales etiquetas corre el riesgo de volverse tautológico. No obstante, podemos señalar que una o dos etiquetas son dominantes en términos de uso para cada marcha. En ellas se alude a la fecha en cuestión y se alternan convocatorias a las calles, lo que puede considerarse especialmente disruptivo en un contexto de distanciamiento social, con alusiones al banderazo (véase Imagen 1).

Imagen 1. Ranking de hashtags más utilizados en cada manifestación



Fuente: elaboración propia. Para el análisis de hashtags utilizamos el paquete *quanteda* (Benoit, Watanabi et al, 2018).

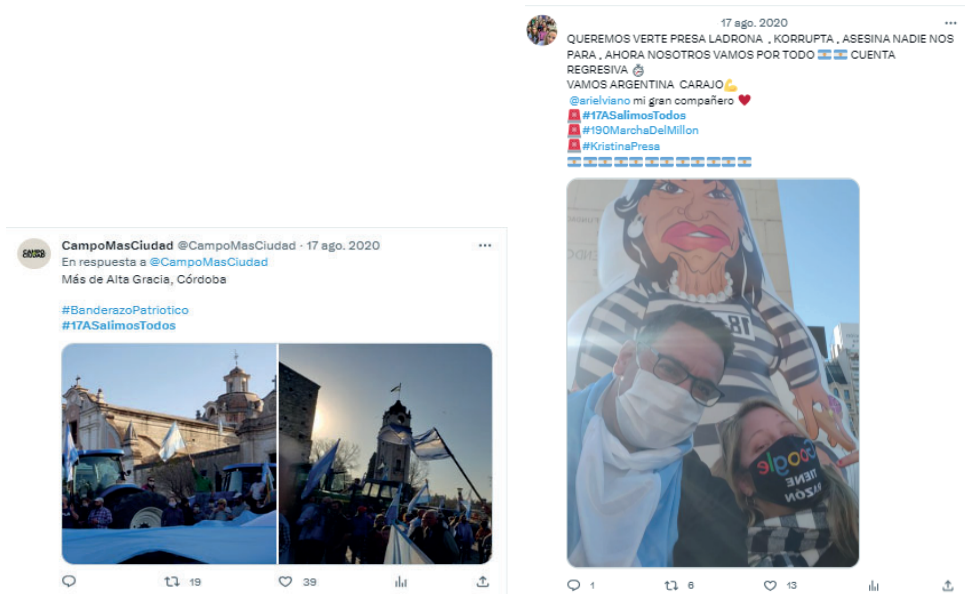
Si buscamos qué repertorios de movilización se repiten como hashtags en las distintas movilizaciones encontramos que, con contundencia, el *banderazo* se define como la forma predominante (ver tabla 2). Esta serie de manifestaciones exponen la simbología nacional y dan un carácter de movimiento autoconvocado y no partidario (ver imagen 2).

Tabla 2. Cantidad de menciones por fecha a hashtags que refieren repertorios de lucha

	20J	9J	17A	12O
<b>banderazo</b>	16081	11225	10135	4553
<b>cacerolazo</b>	92	31	36	4
<b>bocinazo</b>	48	8	8	4

Fuente: elaboración propia.

Imagen 2. Tweets con fotografías de los “banderazos”

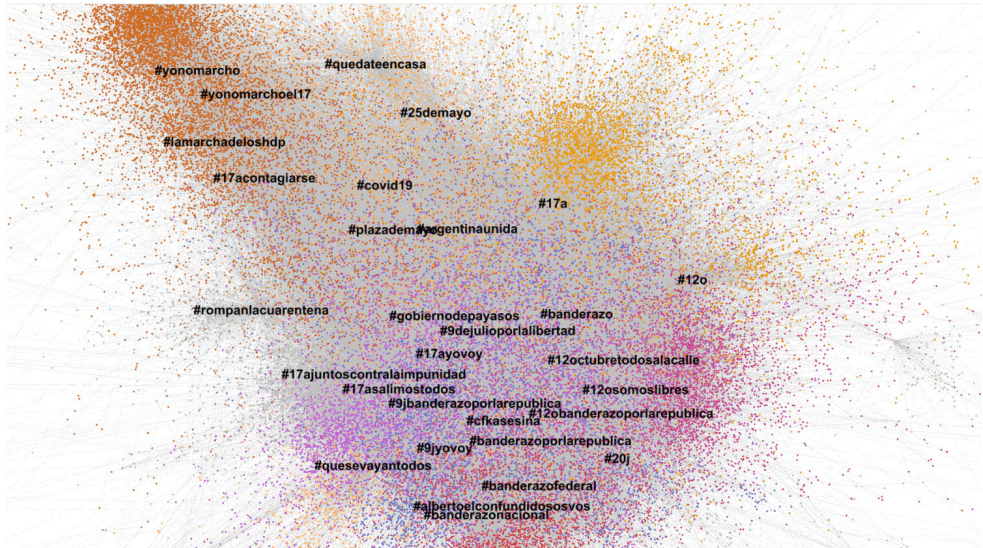


Fuente: extraído de <https://Twitter.com/Ivisinger1/status/1295515979422871554> y <https://Twitter.com/Ivisinger1/status/1295515979422871554>

El seguimiento de hashtags transversales puede ser útil a la hora de caracterizar otros rasgos de la movilización. Por un lado, la elección de fechas patrias para realizar la convocatoria, otorga cierta previsibilidad que permite ir articulando las convocatorias, lo que genera menciones cruzadas, especialmente entre 20J, 9J y 17A. Un conjunto de hashtags opositores (*#albertitere*, *#argentinasincristina*, *#bastadeimpunidad*, *#elpeorgobiernodelahistoria*, *#ladronadelanaciónargentina*, *#quesevayantodos*) habituales en el repertorio tuitero del arco opositor al gobierno, se repite en todos los eventos, aunque con pocas menciones relativas.

Por otro lado, si bien sólo durante el 17A destacan como principales los hashtags que se oponen a estas movilizaciones, una red hashtag-usuarix (imagen 3) nos permite advertir la presencia de una comunidad que emite mensajes en oposición a las movilizaciones.

Imagen 3. Red de vínculos usuario- hashtags para 25M, 20J, 9J, 17A y 12O en Twitter

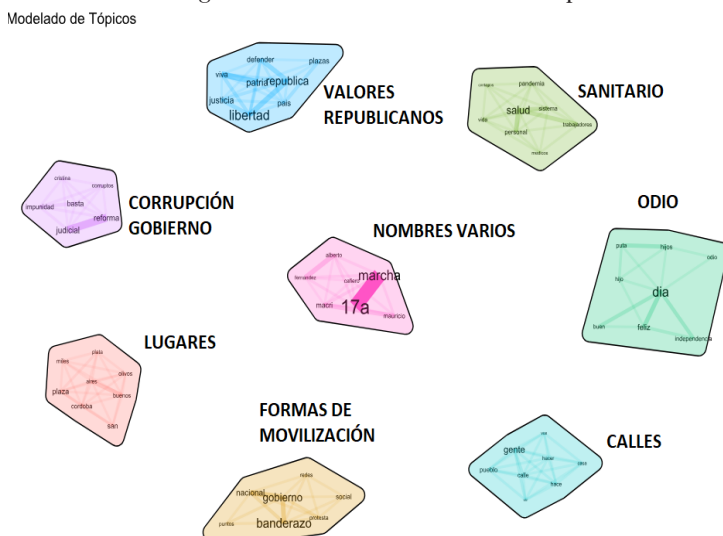


Fuente: elaboración propia en base a Gephi a partir del dataset sin RT. Distribución Force Atlas 2, Colores según algoritmo detección de comunidades. Etiquetas de hashtags con mayor grado de entrada.

Pero ¿cuál es el contenido de estos mensajes a favor y en contra de las movilizaciones? Un modelado de tópicos puede ayudarnos a describir con más claridad las razones que activan estas protestas. Hemos realizado un análisis con el modelo BTM (Xiaohui et al, 2014) diseñado para resolver el problema de la escasez de patrones de co-ocurrencia de palabras a nivel de documento que muestra LDA en textos breves.

Si lo comparamos con un modelado de tópicos para un corpus más amplio, en base a comentarios contra la cuarentena en Twitter en EE. UU., resalta la ausencia de temas como el de las mascarillas o barbijos; la cuestión de las ayudas a comercios por los cierres o las noticias falsas. No obstante, sí hay coincidencias en temas como orden inconstitucional, protestas y cuestiones sanitarias (Karami y Anderson, 2020).

Imagen 4. Gráfico de modelado de tópicos



Fuente: elaboración propia en base a paquete BTM y ggplot.<sup>6</sup> Dataset sin RT y sin hashtags.

En concreto, parece existir una continuidad con el ciclo de movilización opositora al kirchnerismo (Gold, 2017), donde los marcos de sentido de los actores movilizados apuntaban mayoritariamente contra el gobierno por temáticas ligadas a la representación política y/o a características del régimen. Esteban Iglesias (2020) destaca la publicación de un documento titulado “La Democracia está en peligro” que fue firmado por intelectuales, investigadores del Conicet, personalidades, etc. y que plantea el polémico término “infectadura”, aludiendo a cómo la pandemia había convertido a la democracia argentina en una dictadura. Si bien “infectadura” e “infectaduras” aparecen 600 veces entre las etiquetas relevadas, también las denuncias de corrupción, los pedidos de justicia, la defensa de valores como patria, república y libertad, pueden leerse mejor en ese contexto.

La bolsa de términos que hemos denominado como odio, pero que incluye además insultos y referencias positivas en torno al día de la independencia, no nos permite definir un tópico con claridad, pero nos sugiere que resulta relevante estudiar la propagación de mensajes de odio en estas protestas.

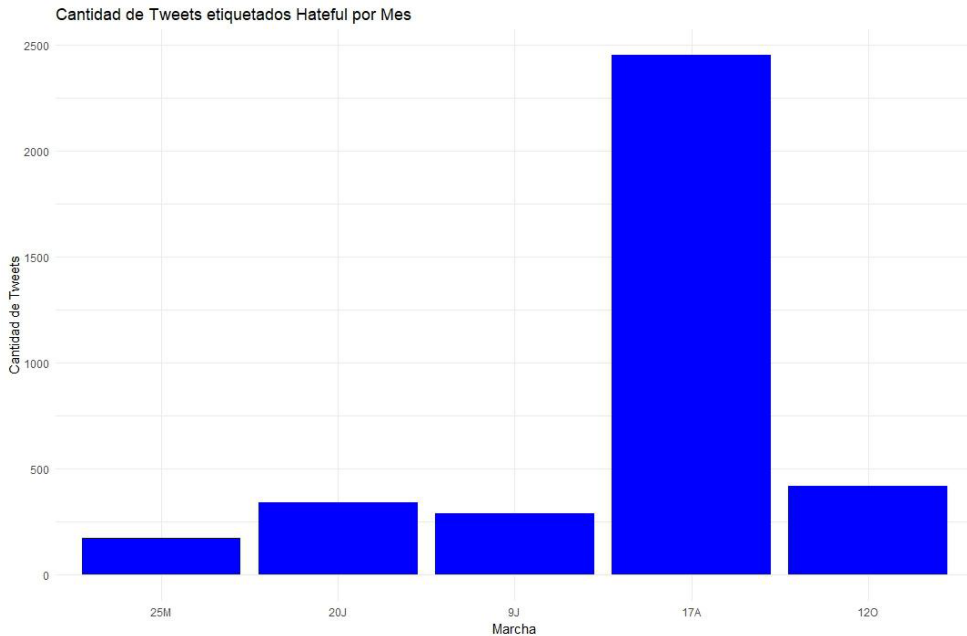
Para clasificar si nuestros tuits contienen o no “discursos de odio” utilizamos la librería `pysentimiento`<sup>7</sup>, una biblioteca de Python que provee modelos pre-entrenados y herramientas de preprocesado para textos sociales en español e

<sup>6</sup> El código está basado en <http://www.bnosc.be/index.php/blog/98-biterm-topic-modelling-for-short-texts>

<sup>7</sup> Véase <https://github.com/pysentimiento/pysentimiento>.

inglés. Una primera mirada nos permite contar 3506 tweets etiquetados como portadores de discurso de odio y distribuirlos en cada marcha.

Imagen 5. Cantidad de tweets que contienen discursos de odio en cada manifestación



Fuente: elaboración propia con pysentimiento.

El gráfico muestra que la marcha del 17A es aquella donde los tweets pueden considerarse más agresivos, algo que podría estar empujado por la batalla virtual que disputan los defensores y detractores de la movilización. A nivel general, no obstante, si ajustamos el umbral de predicción en valores superiores a 0,9, sólo obtenemos 322 tweets, un valor que no llega al 0,25 % de los tweets relevados. Este número indicaría bajos niveles de agresividad si lo comparamos con el 9 % que, con la misma metodología, encontraron al revisar comentarios a noticias sobre COVID-19 publicadas en Twitter por diarios de Argentina (Cotik, et al. 2020).

Si miramos el 10 % más retuiteado, encontramos que 279 fueron etiquetados como “discurso de odio”, lo que significa una proporción prácticamente idéntica a la que obteníamos para el corpus global.

## 5. TRES MIRADAS SOBRE LOS USUARIXS

En los estudios sobre activismo digital, los análisis de usuarioxs son claves. Para Argentina, algunos trabajos han avanzado en la caracterización de los perfiles que se consideran más influyentes. Por un lado, Gamboa Troyano y Beccia (2021) parten de un análisis de contenido de perfiles de los principales representantes políticos de la derecha radical en *Twitter* de Argentina (liberal-conservadora representada por el PRO; “libertaria” en la coalición La Libertad Avanza; y nacionalista-reaccionaria del partido NOS). Los autores se proponen dar cuenta de la inscripción de discursos claves en estos perfiles, tales como la vulneración de las libertades ante el confinamiento y el carácter polémico y reaccionario de los mensajes en redes. Por otro lado, Kessler et al. (2022) abordan la caracterización de internautas conservadores como “influencers de derecha populista” y sus estrategias de intervención en el espacio público digital en *Twitter*. Partiendo de estos antecedentes nos interesa diseñar estrategias para caracterizar a todos los que participan de estas movilizaciones. A continuación, ensayamos tres abordajes.

En primer lugar, efectuamos una aproximación clásica, a partir del análisis de la distribución de favs y RT para cada usuariox (tabla 3) y el análisis de redes de retweets (imagen 7). Entre lxs internautas que fueron más faveados y retuiteados tenemos cuentas de políticos del bloque de Juntos por el Cambio, como la presidenta del PRO Patricia Bullrich (@PatoBullrich) o los diputados nacionales por el PRO Fernando Iglesias (@FerIglesias) y Waldo Wolff (@WolffWaldo). Las cuentas de medios opositores al gobierno como los diarios @LANACION y @Clarín y el programa de televisión @PPTenel13. En la red también se destacan activistas digitales con perfiles anónimos como @ChauOperetaK @ElCoya1977 y @CampoMasCiudad.

Tabla 3. Las 10 cuentas con más interacciones recibidas (favs y retweets)

	Favs recibidos	RT recibidos
@chauoperetaK	146631	18776
@elcoya1977	86240	11305
@patobullrich	80381	16345
@feriglesias	73028	10219
@maquialifraco	59247	11735
@lanacion	44913	7125
@clarincom	43291	6655
@pptenel13	38586	8006
@wolffwaldo	34846	7162
@campomasciudad	32006	5592

Fuente: elaboración propia.

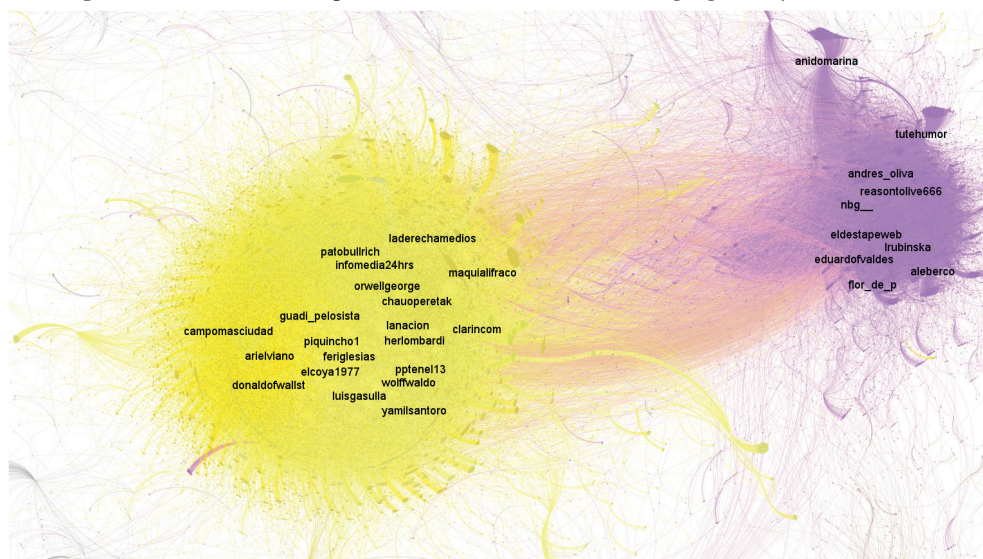
Imagen 6. Tweets de cuentas que figuran en el top 10



Fuente: Izquierda: Patricia Bullrich @PatoBullrich Twitter.  
<https://twitter.com/PatoBullrich/status/1295344731602456580>  
 Derecha: Fernando Iglesias @FerIglesias Twitter.  
<https://twitter.com/FerIglesias/status/1280886122307112965>

El análisis de redes de retweets (imagen 7) da cuenta de clusters cercanos, coordinados, de gran peso, que nuclean a los sectores movilizadores durante estos eventos. Una mirada más amplia de los usuarios más retuiteados agrega nuevas cuentas, pero mantiene un esquema similar en cuanto a la intervención de medios y periodistas (@clarincom @infomedia24hs @laderechamedios @LuisGasulla), políticos (@HerLombardi @yamilsantor) y activistas anónimos (@DonaldOfWallist @OrgellGeorge @Piquincho1 @Guadi\_Pelosista). Agrupados en color amarillo, esta comunidad de usuarios, es reconocida por su aparición en diferentes eventos opositores (Calvo y Aruguete, 2020). En oposición, y con un tamaño considerablemente menor, visualizamos la comunidad que aparece en magenta en el mapa. Al vincularlo con la imagen 3, podemos asociarla con los contra-hashtags como #YoNoMarcho, #17AContagiarse o #Quedateencasa. En esta contra-ofensiva aparecen cuentas menos relevantes y ausentes en los rankings. En su mayor parte se vinculan con periodistas (@aleberco, @eldestapeweb, @nbg\_, @lrubinska, @andres\_oliva), activistas feministas (@anidomarina) o humoristas (@tutehumor). Solo podemos mencionar a un político de segunda línea, el diputado del bloque gubernamental (@eduardofvaldes), y una activista digital kirchnerista (@reasontolive666). En ese sentido, se refuerza la hipótesis de la alineación de medios a posicionamientos ideológicos-partidarios como una característica relevante en la construcción de disputa de los problemas públicos en Argentina (Zunino 2021). También Freelon et al. (2020), resaltan para la experiencia estadounidense las articulaciones en las manifestaciones de derechas entre activistas y medios opositores partidarios.

Imagen 7. Red de retweets para las movilizaciones 25M, 20J, 9J, 17A y 12O en Twitter



Fuente: elaboración propia en base a Gephi. Distribución Force Atlas 2. Colores según algoritmo de detección de comunidades. Hashtags con mayor grado de entrada.

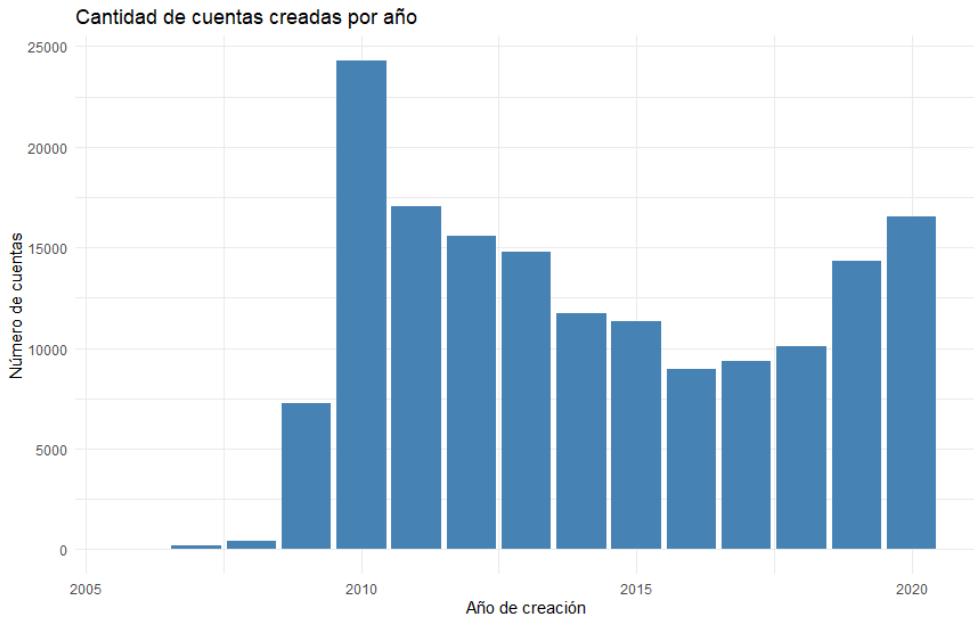
Estas observaciones nos permiten encontrar varios puntos de contacto con la caracterización que Graham et al. (2021) realizan sobre las cuentas claves en las manifestaciones anticuarentena en Australia. Los autores observan una retroalimentación entre medios conservadores, figuras políticas, y una comunidad poco coordinada de lo que denominan líderes de opinión hiperpartidistas.

En segundo lugar, nos preguntamos por aquellos internautas “no conocidos” que participan en los flujos de tweets, y son quienes por debajo configuran la circulación de las narrativas. En este corpus trabajamos sobre 161.985 cuentas que participaron, alentando o repudiando, en alguna o varias de las cinco manifestaciones. Aunque no nos concentramos en esta oportunidad en la diferenciación entre intervenciones humanas y maquinicas (Ferrara, 2020), exploramos el año de creación de las cuentas como un modo de aproximarnos a la existencia de operaciones orquestadas artificialmente detrás de estas movilizaciones. En esa línea, podemos subrayar que en 2010 se crearon el 15% de las cuentas, lo que resulta coincidente con el año de “despegue” de Twitter en la Argentina, tal como lo evidencia una mirada a la creación de las cuentas de políticos profesionales<sup>8</sup>. No obstante, en 2020 se crearon el 10% de las cuentas que participaron en el evento, número que no parece asombroso si consideramos que 2019 devuelve un porcentaje similar y el inicio de la pandemia implicó

<sup>8</sup> Véase el capítulo de Emmanuel Borthiry en este mismo libro.

un impulso importante para las plataformas de redes sociales. En Alemania (Jarynowski et al., 2020) tras la hipótesis que las derechas son más proclives al uso de bots, los autores también analizan el año de creación de las cuentas de los activistas que intervienen en estas manifestaciones y tampoco detectan ninguna relación significativa en estas dos variables. Aunque la información no aparece desglosada del mismo modo, el porcentaje parece más significativo si lo comparamos con el 6% de cuentas que fueron creadas entre julio y octubre de 2020, según las estimaciones que Graham et al. (2021) realizaron en el corpus de tweets de hashtags opositores que recolectaron para la misma etapa durante las manifestaciones antigubernamentales que tuvieron lugar en la “tuitósfera” de Australia.

Imagen 8. Cantidad de cuentas creadas por año



Fuente: elaboración propia (n=161985).

El carácter difuso de estos resultados nos invitó a explorar otras dimensiones. En esa línea, nos propusimos identificar el grupo de usuarios que participaron en todas las manifestaciones, bajo el supuesto que este ejercicio nos permitiría aislar y caracterizar mejor a un conjunto de internautas opositores que promovieron los diversos movimientos. La tabla 4 desglosa la cantidad de usuarios para cada marcha y muestra, al igual que confirmaron Penteadó et al. (2022) para el análisis de conflictos políticos en Twitter en Brasil, que la esfera pública digital se presenta muy dinámica. Cabe resaltar entonces que

la presencia de nuevos usuarios para cada evento aparece como el rasgo más sobresaliente.

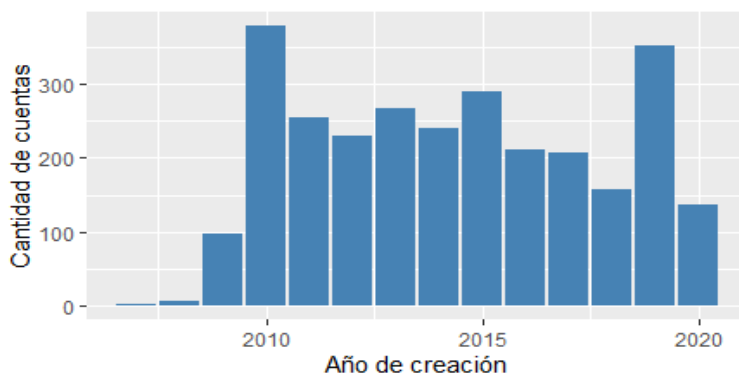
Tabla 4. Cantidad de usuarios recurrentes en cada evento

Usuarios en 2 o más marchas	25182
Usuarios en 3 o más marchas	12529
Usuarios en 4 o más marchas	10838
Usuarios en 5 marchas	2869

Fuente: elaboración propia.

Para conocer mejor a estos activistas recurrentes tomamos el recorte de los usuarios que tuvieron intervención en todos los eventos y proponemos –o al menos esbozamos– una reconstrucción prosopográfica<sup>9</sup> de sus trayectorias que nos permita trazar un camino alternativo para profundizar en el análisis de usuarios. Sobre esta nueva base de datos de 2.869 usuarios observamos el año de creación de las cuentas y verificamos que los picos se dan entre 2010 y 2019, lo que desestima en buena medida la tesis de una creación masiva de cuentas para impulsar las movilizaciones (véase imagen 9)

Imagen 9. Cantidad de cuentas creadas por año para usuarios recurrentes



Fuente: elaboración propia (n=2869).

<sup>9</sup> Para el campo de la historiografía Ferrari (2010) recupera el concepto de *prosopografía*, como el estudio de biografías colectivas de sujetos, agrupaciones y movimientos políticos. Es posible describir los perfiles emergentes del conjunto y analizar las relaciones entre los individuos del mismo o de diferentes campos, o entre los individuos y otros actores colectivos para, finalmente, contribuir a explicar al actor colectivo como una configuración social siempre cambiante y de fronteras lábiles, que actúa dentro de una sociedad en un tiempo determinado.



Imagen 11. Perfiles de usuarios recurrentes



Fuente: Twitter. Capturas tomadas en julio de 2023.

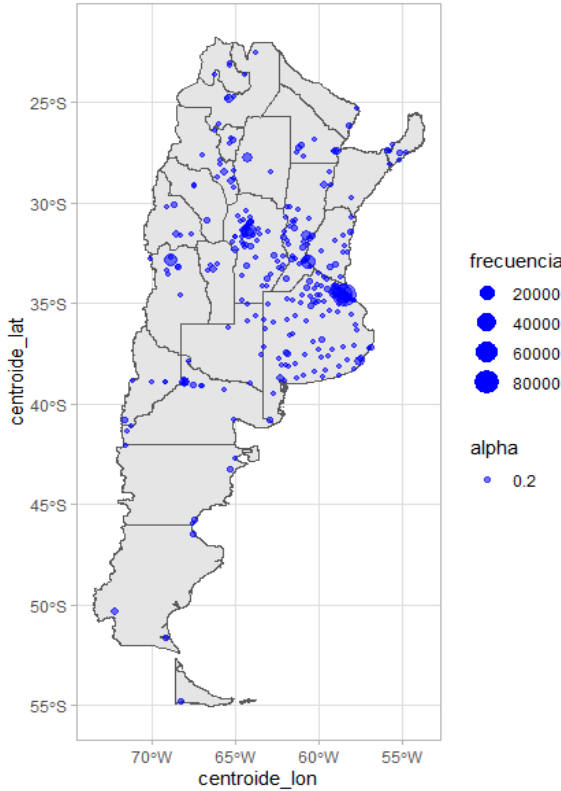
Como tercer alternativa, nos propusimos relevar la ubicación de los usuarios, una dimensión que la literatura ha señalado como problemática para su identificación (Veen et al., 2015). En la revisión de hashtags y contenidos aparecían múltiples menciones a localidades de Argentina, lo que nos sugirió la relevancia de mapearlas. Para este fin, cruzamos un listado de localidades censales<sup>11</sup> con el texto de nuestros tweets, para obtener una lista de aquellos mensajes que contienen menciones de ciudades. El resultado puede visualizarse en la imagen 10.

Tal como podíamos suponer por la distribución de usuarios, hay una mayor cantidad de menciones en la capital del país. No obstante, resulta muy llamativa la existencia de tweets que marcan la presencia de estos “banderazos” en todo el territorio nacional, mostrando un perfil más “federal” para la “tuitósfera” en Argentina.

<sup>11</sup> Véase [https://datos.gob.ar/dataset/jgm-servicio-normalizacion-direcciones-unidades-territoriales-argentina/archivo/jgm\\_8.14](https://datos.gob.ar/dataset/jgm-servicio-normalizacion-direcciones-unidades-territoriales-argentina/archivo/jgm_8.14)

Eliminamos de esta lista las localidades cuyos nombres se repetían, utilizando como criterio la eliminación de las menos pobladas. Por otra parte, eliminamos aquellas localidades cuyos nombres podían confundirse con palabras comunes en los tweets. Asimismo, incorporamos referencias a algunos puntos claves de la protesta, tales como: Obelisco, Plaza de Mayo y Patio Olmos.

Imagen 12. Georreferenciación de localidades mencionadas en los tweets



Fuente: elaboración propia a partir de tidygeocoder<sup>12</sup>.

Sin embargo, desconocemos si las menciones a distintas localidades del territorio, incluyendo algunas muy pequeñas como Las Breñas en Chaco o Las Armas en Provincia de Buenos Aires, provienen de usuarios de la zona. En ese sentido, como un modo de advertir territorios de co-presencia entre las redes y las calles en estas movilizaciones, cruzamos las localizaciones designadas por las cuentas que emitieron los tuits con estas menciones. En 1031 cuentas la ciudad que figuraba en el campo “location” que devuelve la API de Twitter coincidía con la ciudad mencionada en el tweet, comprobando que no se trataba de un fenómeno eminentemente capitalino.

Aunque no podemos verificarlo directamente, consideramos como un antecedente posible de estas expresiones opositoras que utilizan las redes como forma de articulación con las manifestaciones callejeras, las marchas del #SiSePuede. Estas movilizaciones, también conocidas como la marcha del medio millón, que impulsó el entonces presidente Mauricio Macri en la segunda

<sup>12</sup> Véase <https://jessecambon.github.io/tidygeocoder/>

mitad del 2015 con la intención de dar vuelta el resultado electoral desfavorable que había sufrido en las elecciones primarias, muestran una dinámica similar que invita a la participación tanto desde las redes como desde el lugar físico (Annunziata, Ariza y March, 2020).

## 6. REFLEXIONES FINALES

En este trabajo hemos obtenido una mirada panorámica a las manifestaciones que se desplegaron en Twitter antes, durante y después de las movilizaciones que tuvieron lugar en las calles en Argentina durante la segunda mitad de 2020, en plena vigencia de las medidas de distanciamiento social que el gobierno adoptó para evitar la proliferación del coronavirus. Al tratarse de un análisis sobre activismo digital, buena parte del trabajo consistió en recolectar y filtrar un dataset basado en hashtags, donde decidimos incluir las expresiones que impulsaron estas movilizaciones, pero también las que lo rechazaban. Una primera mirada del ranking de hashtags y la red de relaciones entre usuarios y etiquetas verifica que las resistencias movilizaron un volumen significativamente inferior de internautas y solo alcanzaron cierta relevancia en la marcha del 17 de agosto. El análisis de los hashtags también nos permitió mostrar que sus protagonistas identificaron estas protestas como “banderazos”, sentido que se reforzó con la elección de fechas patrias para su realización.

El modelado de tópicos sugiere que las consignas antigubernamentales (denuncias de corrupción, defensa de instituciones republicanas) fueron las centrales en Twitter. Estas etiquetas nos permiten establecer continuidades con manifestaciones antikirchneristas que también se desplegaron en esta red entre 2012 y 2015; pero también en 2019, cuando Mauricio Macri inició una fuerte campaña para procurar dar vuelta su derrota en las elecciones de renovación de mandato. Estas continuidades muestran como también en las plataformas los actores echan mano a repertorios conocidos, a la vez que sus prácticas probablemente resulten condicionadas por la visibilidad que los algoritmos otorgan o no hashtags recurrentes en el pasado.

Por otro lado, si bien aparece un tópico asociado a las medidas sanitarias y el sistema de salud, no hay un lugar central para temas vinculados con teorías de la conspiración asociadas al coronavirus. Sobre este punto, es importante resaltar que prácticamente no se registraron críticas al uso del barbijo. Aunque tampoco se observa como un tópico en el modelado que ensayamos, las consignas contra la cuarenta aparecen en una serie de hashtags que se reiteran en todas las manifestaciones y articulan con las consignas antipopulistas, en slogans como #infectadura.

En cuanto a la presencia de discursos de odio, una primera medición revela que hay índice bajos en los mensajes, con la excepción del evento de 17A, donde

el enfrentamiento entre agitadores y opositores a las marchas modifica el tenor de los mensajes.

El análisis de los usuarios centrales refuerza el perfil opositor que advertíamos en el análisis de los tweets, recibiendo el mayor número de interacciones cuentas de políticos opositores, medios de comunicación que habitualmente critican al gobierno y usuarios anónimos a los que hemos denominado hiperpartidistas de derechas. Aunque iniciamos este apartado con el estudio de los usuarios principales para la conversación, nos interesa resaltar la importancia de trazar estrategias que nos permitan explorar también a la totalidad de las cuentas que participan del evento. En esa línea, sin dejar de resaltar que existe un importante recambio de cuentas en los distintos eventos, proponemos una prosopografía –aún en ciernes– del grupo de usuarios que participan en todas las manifestaciones. Advertimos que se trata de un grupo pequeño de usuarios, integrado por cuentas que no han sido recientemente creadas, cuyas biografías insisten en sus identidades antiperonistas y/o antikirchneristas.

Por último, para registrar los cruces entre lo virtual y lo real propusimos un ejercicio que mostró la presencia de usuarios que acompañaron las movilizaciones en línea a lo largo y a lo ancho del territorio argentino. Al respecto, los hashtags que mencionaban ciudades permitieron articular redes de militancia federal que incluso superaron el sesgo urbano y “capitalinocéntrico” que muestran las estadísticas nacionales sobre usos de esta plataforma.

## REFERENCIAS

- Annunziata, R., Ariza, A. y March, V. (2021). *Promesa y negatividad: el mensaje electoral en las redes sociales*. Facultad de Ciencias Sociales Universidad de Buenos Aires.
- Barrie, Ch. y Ho, J.C. (2021). *AcademictwitterR: an R package to access the Twitter Academic Research Product Track v2 API endpoint*. *Journal of Open Source Software*, 6(62), 3272. <https://doi.org/10.21105/joss.03272>
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S. y Matsuo, A. (2018). “quanteda: An R package for the quantitative analysis of textual data.” *Journal of Open Source Software*, 3(30), 774. doi:10.21105/joss.00774, <https://quanteda.io>.
- Bruns, A. y Burgess, J. (2016). *Methodological Innovation in Precarious Spaces: The Case of Twitter*. En Watson, H, & Roberts, S (Eds.) *Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation*. Palgrave Macmillan.
- Calvo, E. y Aruguete, N. (2020). *Fake news, trolls y otros encantos. Cómo funcionan (para bien y para mal) las redes sociales*. Siglo XXI.
- Canelo, P. (2019). *¿Cambiamos?: La batalla cultural por el sentido común de los argentinos*. Siglo XXI.

- Charoenthansakul, T. y Natee, W. (2023). Twitter and the protest movement in Thailand: A thematic analysis of highly retweeted tweets during the pro-democracy protests. *First Monday*, 28(6). <https://doi.org/10.5210/fm.v28i6.12666>
- Coleman, E. G. (2010). Ethnographic Approaches to Digital Media. *Annual Review of Anthropology*, 39, 487-498. <https://techstyle.lmc.gatech.edu/wp-content/uploads/2011/11/0000005936-annurev.anthro.012809.104945.pdf>
- Coraita, M. y Álvarez, R. (2021). El “banderazo” del 9 de julio: ¿manifestación ciudadana o expresión de odio? *Anáfora Publicación de estudiantes de lingüística*. 4, 42-55. <http://letras.filo.uba.ar/sites/letras.filo.uba.ar/files/documentos/03%20Ana%CC%81fora%20V4%20-%20E1%20banderazo%20del%209%20de%20julio.pdf>
- Cotik, V., Debandi, N., Luque, F. M., Miguel, P., Moro, A., Pérez, J. M., Serrati, P., Zajac, J. y Zayat, D. (2020). A study of Hate Speech in Social Media during the COVID-19 outbreak, *Repositorio Digital UNC*. [https://rdu.unc.edu.ar/bitstream/handle/11086/16293/a\\_study\\_of\\_hate\\_speech\\_in\\_social\\_media\\_during\\_the\\_covid\\_19\\_outbreak.pdf?sequence=2&isAllowed=y](https://rdu.unc.edu.ar/bitstream/handle/11086/16293/a_study_of_hate_speech_in_social_media_during_the_covid_19_outbreak.pdf?sequence=2&isAllowed=y)
- Ferrara, E. (2020). #COVID-19 on Twitter: Bots, Conspiracies, and Social Media Activism. *arXiv*, v1 ([cs.SI]). <https://arxiv.org/abs/2004.09531v1>
- Ferrari, M., (2010). Prosopografía e historia política Algunas aproximaciones. *Antíteses*, 3(5), 529-550. <https://dialnet.unirioja.es/servlet/articulo?codigo=5026731>
- Ferreyra, S. (2022). “Perón Pedófilo”: usos políticos del pasado en Twitter. *Anuario De La Escuela De Historia*, (37). <https://anuariodehistoria.unr.edu.ar/index.php/Anuario/article/view/371>
- Ferreyra, S., Juarez, W., Borthiry E., Dolabani, M. y Calomarde, E. (2021). Procesos de (auto) aprendizaje de herramientas digitales en la investigación social: nubes de palabras en Twitter. *XXXV Jornadas de Investigación y XVII Encuentro Regional de la FAUD-UBA*. <https://www.youtube.com/watch?v=j8nP057sUno>
- Freelon, D., Marwick, A. y Kreiss, D. (2020). False equivalencies: Online activism from left to right. *Science (New York, N.Y.)*, 369(6508), 1197-1201. <https://doi.org/10.1126/science.abb2428>
- Gamboa Troyano, S. y Beccia, B. (2021). Derecha radical, Twitter y pandemia. Las convergencias discursivas entre Vox en España y las derechas argentinas. *Avatares de la Comunicación y la Cultura*, 0(22). <https://publicaciones.sociales.uba.ar/index.php/avatares/article/view/6623>
- Gold, T. (2017). *De redes y cacerolas. El ciclo de movilización anti-gubernamental en Argentina (2012-2013)*. Tesis para optar por el título de Magíster en Ciencia Política. Instituto de Altos Estudios Sociales, Universidad Nacional de San Martín.
- Graham, T., Bruns, A., Angus, D., Hurcombe, E. y Hames, S. (2021). #IStandWithDan versus #DictatorDan: The polarised dynamics of Twitter discussions about Victoria’s COVID-19 restrictions. *Media International Australia*, 179(1), 127-148. <https://doi.org/10.1177/1329878x20981780>

- Gualda, E. (2020). Social network analysis, social big data and conspiracy theories. En M. Butter & P. Knight (Eds.), *Handbook of Conspiracy Theories* (pp. 135-147). Routledge.
- Hao, K., & Basu, T. (2020, 12 de febrero). The coronavirus is the first true social-media “infodemic.”. MIT Technology Review. <https://www.technologyreview.com/2020/02/12/844851/the-coronavirus-is-the-first-true-social-media-infodemic/>
- Haupt, M.R., Jinich-Diamant, A., Li, J., Nali, M. y Mackey, T.K. (2021). Characterizing Twitter user topics and communication network dynamics of the “Liberate” movement during COVID-19 using unsupervised machine learning and social network analysis. *Online Social Networks and Media*, 21. <https://doi.org/10.1016/j.osnem.2020.100114>.
- Iglesias, E. (2020) Protestas sociales en Brasil y Argentina a los 100 días del COVID 19: la persistencia de la dinámica de la polarización. *Espacio Abierto*, 29(4), 70-84. <https://www.redalyc.org/journal/122/12265803004/movil/>
- Imhoff, R.; Zimmer, F.; Klein, O. et al. (2022). Conspiracy mentality and political orientation across 26 countries. *Nature Human Behaviour*, 6, 392-403. <https://doi.org/10.1038/s41562-021-01258-7>
- Jarynowski, A.; Semenov, A. y Belik, V. (2020). Protest Perspective Against COVID-19 Risk Mitigation Strategies on the German Internet. En S. Chellappan, K. K. R. Choo y N. Phan (eds.). *Computational Data and Social Networks. CSoNet 2020* (pp. 524-535). Springer, Cham. [https://doi.org/10.1007/978-3-030-66046-8\\_43](https://doi.org/10.1007/978-3-030-66046-8_43)
- Karami, A. y Anderson, M. (2020). Social media and COVID-19: Characterizing anti-quarantine comments on Twitter. *Proceedings of the Association for Information Science and Technology*, 57, e349. <https://doi.org/10.1002/praz.2.349>
- Kessler, G., Vommaro, G., y Paladino, M. (2022). Antipopulistas reaccionarios en el espacio público digital. *Estudios Sociológicos De El Colegio De México*, 40(120), 651-692. <https://doi.org/10.24201/es.2022v40n120.2213>
- Morresi, S., Saferstein, E. y Vicente, M. (2021). Ganar la calle. Repertorios, memorias y convergencias de las manifestaciones derechistas argentinas. *Clepsidra. Revista Interdisciplinaria de Estudios sobre Memoria*, 8(15), 134-151. <https://ojs.ides.org.ar/index.php/Clepsidra/article/view/252>
- Nieto, A. y Ferreyra, S. (Abril 2020) *Twitter en tiempos de pandemia: ¿unidxs o polarizadx?* Observatorio Político Ciudadano Electoral. <https://www.observatoriopolitico.com.ar/Twitter-en-tiempos-de-pandemia-unidxs-o-polarizadx/>
- Parodi, R., Cuesta, M. y Wegelin, L. (2022). Problematizar los discursos de odio: Democracia, redes sociales y esfera pública. *Trampas De La Comunicación Y La Cultura*, (87), e061. <https://doi.org/10.24215/2314274xe061>
- Paulino, F. O. y Waisbord, S. (2021). Las narrativas del populismo reaccionario. *Mediapolis - Revista de Comunicação, Jornalismo E Espaço Público*, 12, 33-48. [https://doi.org/10.14195/2183-6019\\_12\\_2](https://doi.org/10.14195/2183-6019_12_2)

- Penteado, C.L. de C., Goya, D.H., dos Santos, P.D. y Jardim, L. (2022). Populismo, desinformação e COVID-19: comunicação de Jair Bolsonaro no Twitter. *Media & Journalism*, 22(40), 239-260. [https://doi.org/10.14195/2183-5462\\_40\\_12](https://doi.org/10.14195/2183-5462_40_12).
- Plümper, T.; Neumayer, E. y Pfaff, K.G. (2021). The strategy of protest against COVID-19 containment policies in Germany. *Social Science Quarterly*, 102, 2236-2250. <https://doi.org/10.1111/ssqu.13066>
- Rambukkana, N. (2015). *Hashtag Publics: The Power and Politics of Discursive Networks*. Peter Lang. <https://www.peterlang.com/document/1118297>
- Rogers, R. (2013). *Digital methods*. Boston: MIT press.
- Semán, E. (2021). *Breve historia del antipopulismo*. Siglo XXI.
- Sigal, S. (2006). *La Plaza de Mayo. Una crónica*. Siglo Veintiuno editores.
- Soto Pimentel, V.; Gradin, A. y Reiri, M. (2021). *La conflictividad social en Argentina durante el 2020. Conflictividad social de Argentina en el año 2020. Un análisis de los hechos de protesta durante la pandemia del COVID-19*. Informe N°20 del Observatorio sobre Políticas Públicas y Reforma Estructural, FLACSO: Buenos Aires.
- van der Zwet, K., Barros, A.I., van Engers, T.M., et al. (2022). Emergence of protests during the COVID-19 pandemic: quantitative models to explore the contributions of societal conditions. *Humanit Soc Sci Commun*, 9, 68. <https://doi.org/10.1057/s41599-022-01082-y>
- Van Dijck, J. (2016). *La cultura de la conectividad: una historia crítica de las redes sociales*. Siglo XXI Editores.
- Veen, H.; Hiemstra, D.; van den Broek, T.; Ehrenhard, M. y Need, A. (2015). Determine the User Country of a Tweet. ([cs.SI]). <https://arxiv.org/abs/1508.02483v1>
- Vishnuprasad, P. S.; Nogara, G.; Cardoso, F.; Cresci, S.; Giordano, S. y Luceri, L. (2023). Tracking Fringe and Coordinated Activity on Twitter Leading Up To the US Capitol Attack. ([cs.SI]). <https://arxiv.org/abs/2302.04450v2>
- Vommaro G. (2017). *La larga marcha de cambiemos. La construcción silenciosa de un proyecto de poder*. Siglo XXI.
- Xiaohui, Jiafen, Yanyan y Xuequi (2014). *A Biterm Topic Model for Short Text*. <https://github.com/xiaohuiyan/xiaohuiyan.github.io/blob/master/paper/BTM-WWW13.pdf>
- Zunino, E. (2021). Medios digitales y COVID-19: sobreinformación, polarización y desinformación. *Universitas-XXI, Revista de Ciencias Sociales y Humanas*, (34), 133-154. <https://doi.org/10.17163/uni.n34.2021.06>

# MICRONARRATIVAS NEGACIONISTAS Y TEORÍAS DE LA CONSPIRACIÓN SOBRE LA “PLANDEMIA” EN TWITTER<sup>1</sup>

ESTRELLA GUALDA

*Catedrática de Sociología*

*Universidad de Huelva, ESEIS/COIDESO, España*

DOI: 10.14679/2755

## Sumario:

1. Teorías de la conspiración y “plandemia” de COVID-19.
  2. Objetivos, hipótesis y líneas de exploración.
  3. Métodos, técnicas y fuentes de la investigación.
    - 3.1. Diseño metodológico.
    - 3.2. Producción de datos: Colección y fuente de datos.
    - 3.3. Preprocesamiento y análisis.
  4. Evolución de los tuits sobre la plandemia y rasgos básicos.
    - 4.1. Rasgos básicos.
    - 4.2. Evolución temporal y comparativa de las etiquetas #plandemia y #plandemic.
  - 5.5. Plandemic & plandemia: Principales tokens en el dataset.
  6. Micronarrativas de hashtags en torno a la plandemia.
  7. Discusión y conclusiones.
- Referencias.

## 1. TEORÍAS DE LA CONSPIRACIÓN Y “PLANDEMIA” DE COVID-19

Es conocido que la pandemia se acompañó del crecimiento de la desinformación, popularizándose el término de “infodemia” para referirse a ella (Pan American Health Organization, 2020). La rápida difusión de teorías conspirati-

---

<sup>1</sup> Este trabajo forma parte del Proyecto I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGBTI [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDER A way of making Europe” (<http://www.uhu.es/nonconspirahate/>). También agradecemos el apoyo de los centros de investigación de la Universidad de Huelva, el grupo de investigación “Estudios Sociales E Intervención Social, ESEIS” y el centro “Pensamiento Contemporáneo e Innovación para el Desarrollo Social, COIDESO”. El capítulo se ha beneficiado igualmente de mi participación en el proyecto de excelencia andaluz: “Impacto de la desinformación en Andalucía: análisis transversal de las audiencias y las rutinas y agendas periodísticas. DesinfoAnd” (ProyExcel 00143, <https://www.disinformationresearch.uma.es>).

vas sobre la COVID-19, que aparece desde el inicio de la pandemia, se benefició del altavoz proporcionado por las redes sociales. Los efectos negativos de la creencia en teorías de la conspiración relativas a la pandemia de COVID-19 son evidentes para las instituciones públicas, habiéndose constatado que “Durante la pandemia de COVID-19 se ha visto un incremento de teorías de la conspiración dañinas y engañosas” (Comisión Europea, 2022).

Un ejemplo del tipo de desinformación que se difunde durante la pandemia es precisamente la teoría de la conspiración que alude a la existencia de la “plandemia” de COVID-19 (Çalışkan, 2022). La idea de “plandemia” se refiere a que la pandemia de COVID-19 obedece a un plan deliberado de una élite global, los poderosos, destinada a controlar a la población y a privarla de libertad. La palabra “plandemic”, deriva de la combinación de las palabras “pandemia” y “plan” y aparece en un documental llamado “Plandemic: The Hidden Agenda Behind COVID-19”, que fue lanzado en Facebook en mayo de 2020. El documental es dirigido por Mikki Willis y producido por Judy Mikovits, vinculada al movimiento antivacunas. Este documental alcanzó una amplia difusión en diferentes plataformas (Nazar y Pieters, 2021; Jaworsky, 2023; Neil y Campbell, 2020) y fue fuertemente criticado por la comunidad científica. El acceso público al mismo fue retirado pronto (Memetic.Media, 2020). No obstante, no es el único documental que se hizo popular durante la pandemia con el fin de difundir desinformación y teorías conspirativas, así como de desacreditar a científicos y expertos sanitarios. Otros como, por ejemplo, “The Truth About COVID-19” (2020), “The Great Reset” (2020) o “The COVID-19 Hoax” (2021), presentaron una visión negativa de la pandemia y contribuyeron a reforzar la teoría de la conspiración sobre la plandemia de COVID-19, así como a dar apoyo a otras teorías antivacunas o negacionistas.

La teoría de la conspiración sobre la “plandemia” se desarrolla en un contexto donde se difunden otras teorías que logran mucha difusión y que se encuentran conectadas entre sí habitualmente en redes sociales como Twitter. La creencia en estas teorías y los comportamientos de riesgo que las acompañan, como no seguir las recomendaciones sanitarias de los gobiernos (Gualda et al., 2021), pueden provocar efectos negativos para la salud pública. En los últimos años se ha desarrollado una diversidad de teorías de la conspiración relacionadas con el coronavirus (Pavela Banai et al., 2022) que, como veremos a lo largo de este capítulo, están enlazadas con la llamada “plandemia”. Algunas de estas teorías, muy descabelladas a veces, suelen jugar a crear la sospecha y la desconfianza de los ciudadanos. Habitualmente han venido atacando frontalmente el discurso oficial de los gobiernos, o deslegitimando a oponentes políticos, desacreditando a las autoridades sanitarias y al conocimiento experto, buscando su desprestigio, lo que cuenta con efectos negativos para la democracia (Muirhead y Rosenblum, 2022).

Entre las teorías que más se han popularizado, se encuentran algunas que tienen que ver con el origen del coronavirus y su propagación, como por ejemplo que: “El coronavirus se fabricó intencionalmente en un laboratorio”, o que “El coronavirus fue creado como arma biológica en la guerra entre los países poderosos (como China, Estados Unidos, Rusia)”, que “La propagación del coronavirus está relacionada con la tecnología del 5G” o incluso que “Bill Gates está usando la pandemia para el control de la población” (Ahmed, 2023).

Otro tipo de teorías de la conspiración sobre la pandemia apuntan hacia las amenazas del coronavirus y vienen a expresar que el coronavirus no es muy peligroso en realidad, que los vídeos y fotografías de los hospitales vacíos son una prueba de que la pandemia es un engaño, que tanto la infección como las tasas de mortalidad proporcionadas por la OMS son falsas, que las vacunas tienen innumerables efectos secundarios o que las vacunas provocan muertes repentinas. Este tipo de teorías, aparte de tener un componente político de fondo, suponen también un fuerte descrédito a la ciencia y la investigación. A veces han atacado directamente a los laboratorios encargados de producir las vacunas, como ha ocurrido con Pfizer y se expresa sintéticamente en el hashtag #Pfizergate. Con el desarrollo de la pandemia, por otra parte, se ha producido una revitalización del viejo movimiento antivacunas (Douglas, 2021), mezclado con movimientos negacionistas, anticuarentena y antimascarillas.

Encontramos también la teoría de la conspiración estadounidense “QAnon” que ha propagado diversidad de bulos durante la pandemia (Zhang, et al., 2022) a raíz de que un funcionario del gobierno conocido como “Q” declaró en 2017 en el foro del sitio web *Achan* (durante la presidencia de Trump) que había un grupo en la élite que controlaba los Estados Unidos y usaba su poder para abusar encubiertamente de niños (el conocido como #pizzagate) (Xu y Sasahara, 2022). Partidarios de Trump, negacionistas o antivacunas se han vinculado a QAnon, que ha difundido teorías que culpan a China de encubrir el coronavirus, relativas a la cuarentena e incluso ataca al movimiento “Black Lives Matter”. El “pizzagate”, como teoría conspirativa que se hace viral en las elecciones presidenciales de los Estados Unidos en 2016, ataca al partido demócrata de los Estados Unidos conectando a restaurantes y funcionarios de alto rango del partido con una red de tráfico de personas y abuso sexual infantil. Entre los establecimientos a los que se alude se encuentra la pizzería Comet Ping Pong en Washington. La difusión de la etiqueta #Pfizergate por parte del movimiento antivacunas, como teoría de la conspiración que desarrolla la idea, se fundamenta en mensajes que difunden desinformación sobre las vacunas, e incluso sobre las empresas farmacéuticas (Jones-Janh y Chung, 2022).

En este contexto, se consolidan tanto el argumento de la “plandemia”, como la sospecha sobre la veracidad del virus, así como la resistencia a los procesos de vacunación o la falta de cumplimiento de las medidas recomendadas por los go-

biernos. Con el tiempo, los términos y etiquetas “plandemia” o “plandemic” se han seguido usando durante la pandemia en las redes sociales, así como en otros idiomas. En este capítulo estudiamos, en perspectiva comparada, las narrativas sobre la “plandemia” que se encuentran en Twitter en estos años en español e inglés con el objetivo de identificar, entre otros aspectos, los elementos clave que las caracterizan, así como si existen pautas internacionales comunes en los discursos que se refieren a la plandemia. Este conocimiento, además de ser clave para comprender los ejes en los que se articulan las narrativas de negación de la pandemia o las acciones de resistencia a las medidas dispuestas para combatirla, es de utilidad para prepararnos ante futuras pandemias o situaciones similares que presenten riesgos para la salud pública, en un contexto internacional caracterizado por la circulación de grandes dosis de desinformación en las plataformas de redes sociales.

## 2. OBJETIVOS, HIPÓTESIS Y LÍNEAS DE EXPLORACIÓN

Son varias las preguntas de investigación y objetivos sobre los que descansa este trabajo, siendo clave conocer cuáles son las principales micronarrativas conspirativas en relación con las teorías de la conspiración negacionistas de la pandemia, encuadrándose aquí contestar a la pregunta de cuáles son los principales temas o cuestiones que se difunden en Twitter cuando los tuits introducen hashtags como los de “plandemic” (en inglés) o “plandemia” (en español). Nos interesa igualmente comparar las pautas de comunicación entre los datos de Twitter en inglés y en español, sobre la hipótesis a explorar de que existen conexiones claras en las narrativas del movimiento negacionista que se expresa a través de los hashtags #plandemia y #plandemic en dos lenguas diferentes. A los efectos de aproximarnos a las narrativas de la plandemia, enfocamos nuestro interés en identificar las narrativas negacionistas y sus conexiones con otros movimientos sociales a través del lenguaje que emplean y al tratarse de datos de Twitter, particularmente de palabras, hashtags y emojis, desde una perspectiva tanto descriptiva como relacional.

Igualmente, nos interesa conocer si la teoría conspirativa de la plandemia se vincula a otras teorías conspirativas en relación con la pandemia, así como a cuáles. Nos interesa explorar las relaciones que se perciben a través de los mensajes en Twitter entre los movimientos negacionistas, anticuarentena, antivacunación y antimascarillas. Entre las hipótesis, se prevé que los discursos pongan en conexión diferentes narrativas correspondientes a los movimientos citados arriba, y generando la idea de que están conectados entre sí. Igualmente, se prevé encontrar narrativas similares de rechazo a los gobiernos, líderes gubernamentales o fuentes de autoridad en el ámbito sanitario, como la bibliografía ha venido documentando (Ahmed, 2023; Muirhead y Rosenblum, 2022). Otro

aspecto que prevemos encontrar, ligado al descrédito de fuentes de autoridad, es que tanto en inglés como en español se encuentren mensajes que, a través de las palabras, hashtags, emojis u otros elementos empleados puedan suponer, directa o indirectamente, una elevación del nivel de crispación o ansiedad ante la pandemia, así como comportamientos incívicos.

### 3. MÉTODOS, TÉCNICAS Y FUENTES DE LA INVESTIGACIÓN

#### 3.1. Diseño metodológico

Esta investigación identifica y compara micronarrativas en línea sobre la “#pandemia” canalizadas y difundidas al público desde posturas negacionistas de la pandemia que suelen conectar con movimientos antivacunas, anticuarentena y antimascarillas en las redes sociales entre 2020 y 2023. Se comparan mensajes en Twitter en español e inglés con la idea de avanzar en la comprensión de las pautas de difusión del movimiento conspirativo negacionista de la pandemia y su conexión con otras teorías de la conspiración. Dadas las preguntas de investigación formuladas, y encuadrado en un enfoque computacional, el diseño metodológico elegido se inscribe en los métodos mixtos, siguiendo la línea de otras investigaciones en este campo, donde se complementan y triangulan métodos cuantitativos y cualitativos computacionales (Jikeli y Soemer, 2023; Nassauer y Legewie, 2020).

#### 3.2. Producción de datos: Colección y fuente de datos

Nuestra fuente de datos para este capítulo ha sido Twitter. Los datos que analizamos han sido extraídos conectándonos directamente con la API v2 de Twitter, con Academic Access, a través de *twarc2*, bajo Python<sup>2</sup>. *twarc2* es una herramienta de línea de comandos y una biblioteca Python que permite recopilar y archivar datos de Twitter a través de su API. Este capítulo obtiene una parte de los datos que se extrajeron a partir de un proyecto más amplio [véase la nota a pie 1] interesado en el estudio de las teorías de la conspiración que han circulado durante la pandemia de COVID-19. El proyecto NON-CONSPIRA-HATE generó un dataset global [*Conspiracy Theories Dataset, 2020-2023*<sup>3</sup>], compuesto

<sup>2</sup> Véase en <https://github.com/DocNow/twarc>, en <https://melaniewalsh.github.io/Intro-Cultural-Analytics/Data-Collection/Twitter-Data-Collection.html> y en <https://twarc-project.readthedocs.io/en/latest/>

<sup>3</sup> Para la obtención de esta colección global se emplearon 35 hashtags o palabras clave (en español e inglés) que representaban áreas relevantes de las teorías de la conspiración durante la pandemia. Previamente se había realizado un análisis para detectar palabras clave o hashtags populares o significativos que representaran el tema de estudio desde el inicio de la pandemia, esto es, teorías de la conspiración relativas a la pandemia y expresiones conexas de movimientos sociales que se resisten a las indicaciones de los gobiernos, entre las que estaban las palabras

por 5.509.549 tweets orgánicos (sólo tweets originales, sin RTs), a partir del que se han filtrado dos submuestras. Una submuestra en español, con tuits que contienen la etiqueta #plandemia y otra en inglés, con la etiqueta #plandemic. La colección abarca desde enero de 2020 a febrero de 2023, recogiendo aspectos clave de las manifestaciones en Twitter relacionadas con la difusión de teorías de la conspiración relativas a la pandemia.

Uno de los hashtags clave asociados internacionalmente al negacionismo de la pandemia de COVID-19 ha sido el de #plandemia o #plandemic (en español e inglés), con múltiples adaptaciones o traducciones a otras lenguas. El objetivo de llevar a cabo una submuestra con tuits originales que específicamente contuvieran dicha etiqueta era acotar un conjunto de datos con mensajes que, en principio, por el uso de esta etiqueta, nos podía aproximar al conocimiento de las formas de expresión del movimiento negacionista en las redes sociales, incorporando la dimensión internacional.

A partir del dataset global citado arriba, se obtuvieron dos submuestras, que son las que analizamos en este capítulo. Para la obtención de las submuestras relativas a la “plandemia” se aplicó al dataset principal una expresión regular con *grep* en R para seleccionar todos los tuits que contuvieran la cadena “plandem” (en mayúsculas o minúsculas). Por ejemplo, tuits que contuvieran cadenas de caracteres como: “plandemic”, “plandemia”, “#plandemia”, “plandemic”, “#coronaplandemic”, “#plandemichoax”, “#farsaplandemia”, “#plandemic2020hoax”, “#esunaplandemiadelnom”, u otras expresiones que contaran con esta cadena de caracteres que en una exploración previa se había detectado que siempre se referían a la versión en español o en inglés de la palabra “plandemia”, representativa del movimiento negacionista.

En el proceso de filtrado del dataset global se obtuvo una submuestra inicial con 388.636 tuits en los que aparecían tuits con la cadena “plandem” no solo en español (153.327 tuits) e inglés (131.629 tuits), sino también en otras lenguas (103.680 tuits). A partir de este dataset, se generaron dos submuestras a partir de la columna “lang” (lengua), seleccionando los tuits que siendo clasificados en Twitter como en español e inglés contuvieran específicamente los hashtags que nos interesaban. Se resume el proceso en el Cuadro 1.

---

y hashtags “plandemia” y “plandemic”, además de otras como #covididiots, #dictadurasanitaria, #dittadurasanitaria, #plandemia, #novax, #vaccinesideeffects, #Pfizergate, #nolockdown, #yonomepongoelbozal, #sin bozal, #5G, u otras, citando algunas referidas a movimientos conexos como el negacionista, antimascarillas, antivacunas y anticuarentena.

Cuadro 1. Proceso de selección de las submuestras #plandemia y #plandemic

<i>Conspiracy Theories Dataset, 2020-2023</i>		
5.509.549 tuits originales (sin RTs)		
Selección de tuits que contienen la cadena “plandem” (grepl)		
388.636 tuits		
Selección de tuits según la etiqueta en cada lengua		
#plandemia: español	#plandemic: inglés	Otras lenguas
153.327	131.629	103.680

Fuente: elaboración propia. *Conspiracy Theories Dataset, 2020-2023*.

### 3.3. Preprocesamiento y análisis

Antes de iniciar los análisis se llevó a cabo un primer análisis exploratorio de los datos, seguido de una etapa de preprocesamiento y limpieza de datos común en el análisis de texto basado en estrategias computacionales, que conlleva operaciones tales a la conversión a minúsculas, eliminación de puntuaciones, de palabras muy frecuentes, eliminación de palabras irrelevantes (*stopwords*), derivación, lematización y otras (Welbers, Van Atteveldt y Benoit, 2017; Silge y Robinson, 2022). Posteriormente, los datos fueron procesados y analizados con apoyo de R y Python, con el fin de identificar las diferentes narrativas de la plandemia en Twitter, como expresión del movimiento negacionista. Para este capítulo, se emplean varias estrategias de análisis que pasan, según el aspecto considerado, por llevar a cabo un análisis descriptivo y evolutivo del conjunto de datos, de corte más cuantitativo, un análisis textual y visual de contenido, con perspectiva crítica, así como el estudio de las redes semánticas y las micronarrativas que forman elementos claves en Twitter como son los hashtags. Se han empleado diferentes paquetes en R con este objeto, destacándose algunos como *quanteda* (Benoit et al., 2018), *tidyverse* (*dplyr*, *ggplot2*, *forcats*, *lubridate*, *readr*, etc., Wickham et al., 2023; Silge y Robinson, 2022), *plotly* (Sievert, 2020), u otros con funciones más accesorias.

## 4. EVOLUCIÓN DE LOS TUIITS SOBRE LA PLANDEMIA Y RASGOS BÁSICOS

### 4.1. Rasgos básicos

La serie de tuits que analizamos en este capítulo fue seleccionada (Cuadro 1) si se daban dos condiciones. Por una parte, que se tratara de tuits en español e inglés (a partir de la variable recogida en Twitter sobre la lengua) y, por otra, que los tuits contuvieran las etiquetas #plandemia (en el dataset en español) y #plandemic (en el dataset en inglés). Los tuits que analizamos son tuits originales, sin retuits, al objeto de captar la diversidad de las narrativas, obviando las repeticiones que tendrían lugar en el caso que se consideraran

los retuits. La Tabla 1 recoge algunas de las características básicas de las dos submuestras. Cabe decir que, aunque en las secciones que siguen nos vamos a referir a los principales tokens, hashtags o emojis mencionados en los tuits, nuestro análisis se concentrará en los elementos que se repiten de forma más sistemática, dejando fuera otros.

Tabla 1. Descripción básica de las submuestras

	#plandemia	#plandemic
Número de tuits	153.327	131.629
Número de tokens (features) diferentes	272.131	219.867
Número de hashtags diferentes	30.015	40.638
Número de usuarios diferentes	33.005	42.046
Número de páginas web diferentes	99.561	69.618
Número de emojis	1.300	18.392

Fuente: submuestras a partir del *Conspiracy Theories Dataset*, 2020-2023.

#### 4.2. Evolución temporal y comparativa de las etiquetas #plandemia y #plandemic

La Figura 1 permite observar comparativamente el número de tuits que se publican sobre #plandemia y #plandemic, en español e inglés. Como se aprecia en el gráfico, existe una evolución relativamente paralela en español e inglés en cuanto a la conversación en Twitter que contiene estos hashtags, en tuits originales. No obstante, no se trata de evoluciones idénticas, apreciándose algunas diferencias en las discusiones, siendo la más notable la que se produce en el pico que corresponde al 20 de julio de 2022 en la serie de datos en español donde, entre otros mensajes, se encuentra un número importante de ellos sobre Chile, conectando ese día con expresiones de protestas hacia las medidas de salud del gobierno y antivacunas, así como para expresar el deseo de un cambio de gobierno en Chile (por ejemplo, con el uso de palabras o hashtags como #gobiernoterminado o #chaopase), o que se oponen, vía #RechazoXAmorACHile, a la nueva constitución.

Algunas de las palabras y hashtags más importantes que se vinculan con #plandemia el 20 de julio de 2022 (#jsct, #aur, aur, #revolucionhumanista, #chaopase, vacunas, bozal, presidente, #noalpasdemovilidad, #noalnom, #dictadurasasanitaria, #rechazoxamorachile, #agenda2030, #gobiernoterminado, #poto y otros) de una forma u otra desarrollan la idea de que la pandemia es una conspiración dirigida por un gobierno o una élite global, del mismo modo que hace la etiqueta #plandemia. Por ejemplo, algunas palabras como “vacunas” y “bozal” se emplean para criticar medidas de salud pública, como la vacunación y el uso de mascarillas. Otros como “presidente” y #dictadurasasanitaria critican al gobierno por su manejo de la pandemia. Se suelen emplear en un contexto que argumenta que la pandemia es una mentira. El 20 de julio de 2022 estas palabras

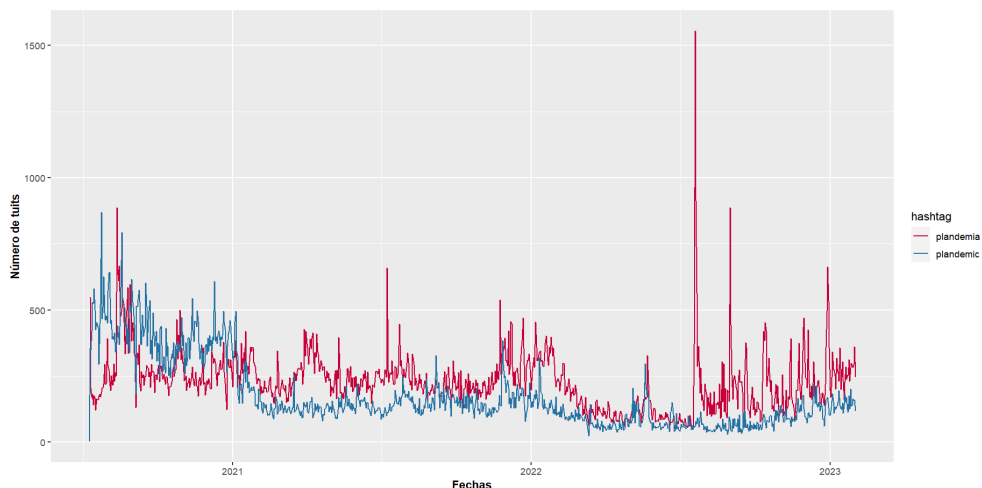
y hashtags se emplearon en el marco de protestas antivacunas y anti-COVID-19 que se llevaron a cabo en Chile en la Plaza de la Dignidad, en Santiago de Chile.

En inglés, por ejemplo, el principal pico de la palabra #plandemic lo encontramos el veinticinco de julio de 2020, día en el que destaca la difusión de un hashtag como #scamdemic, usado para argumentar que la pandemia es un engaño o una conspiración. Ese día son prevalentes otros elementos como “Sinclair” o “@wearesinclair”, en referencia a Sinclair Broadcast Group, que difundió noticias, alineadas con el movimiento antivacunas, que cuestionaban la seguridad de la vacuna contra la COVID-19. Igualmente, este día circuló mucho #wwg1wga o #qanon, hashtags utilizados a menudo por los seguidores de QAnon, teoría de la conspiración que alega que el mundo está dirigido por una secta satánica de pedófilos.

Otro pico importante en inglés lo encontramos, por ejemplo, el 19 de agosto de 2020, donde se mantienen prevalentes una etiqueta empleada durante todo el periodo como #scamdemic, junto a otros como #maga, de apoyo a Trump y que significa “Make America Great Again”, u otros como #Agenda21 y #covid1984 en alusión a la idea de que la pandemia se está empleando para controlar a la población.

En cualquier caso, como era de esperar, lo que se encuentra en la evolución comparativa que se sintetiza en la Figura 1 es que, sobre una base común de rechazo a las medidas gubernamentales y sanitarias, en diferentes expresiones de los movimientos anticuarentena, antivacunas, antimascarillas y negacionista, en diferentes momentos, en ambas series, destacan algunos debates o protestas de corte nacional, como las citadas a modo de ejemplo arriba.

Figura 1. Evolución de los hashtags #plandemic y #plandemia (2020-2023)



Fuente: *Conspiracy Theories Dataset: submuestras #plandemic y #plandemia, 2020-2023.*

## 5. PLANDEMIC & PLANDEMIA: PRINCIPALES TOKENS

Se han extraído los principales tokens que se encuentran en ambas submuestras, teniendo en cuenta los diferentes elementos que se podrían encontrar en un tuit (por ejemplo, palabras, hashtags o emojis), con el objeto de observar cuáles son los elementos principales que aparecen en las micronarrativas en Twitter. Siguiendo a Silge y Robinson (2022:1), un token es una “unidad significativa de texto, como una palabra, que nos interesa utilizar para el análisis, y la tokenización es el proceso de dividir el texto en tokens” [traducción propia]. En este caso, entre los principales elementos de la submuestra #plandemia, nos encontramos tanto con palabras como con etiquetas y emojis, mientras que los emojis aparecen igualmente en el dataset en inglés, pero no en las primeras posiciones. En el caso de las páginas web recomendadas, tanto en la submuestra en español como en inglés, son importantes, pero, igualmente, no en la misma medida que otros elementos del discurso como los hashtags.

Por otra parte, respecto al conjunto de datos sobre #plandemic, prevalecen como era esperable, por una parte, palabras y hashtags claramente asociados con la pandemia (#covid19, #coronavirus, covid, #covid) (Figura 2). Destacan también algunos hashtags que son muy representativos de diversas teorías de la conspiración que se difunden durante la pandemia. Es el caso de: #agenda21, #scamdemic, #id2020 y #greatreset (Figura 2), que se encuentran entre los principales tokens de este conjunto de datos, aunque no son los únicos que hemos encontrado. Como veremos en una sección posterior, muchos de estos tokens aparecen ligados entre sí, configurando una narrativa que destaca aspectos negativos y críticos asociados a la gestión de la pandemia, lo que se ha documentado ya para otras redes diferentes a Twitter (Baines, Ittefaq y Abwao, 2021). Se trata de una narrativa que también desarrolla discursos de miedo y alarmistas, con una fuerte carga emocional, así como grandes dosis de incivildad, fácilmente rastreable a través de palabras claves como insultos:

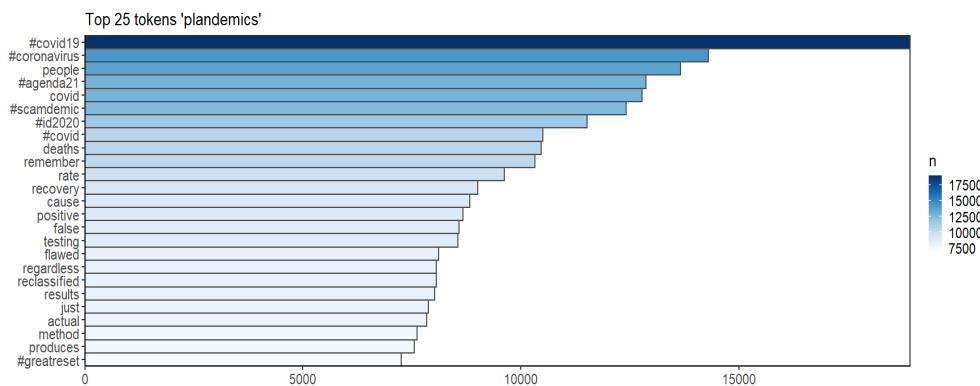
«This sums up why the sleepless 🦉 are behaving so dumb #Scamdemic #plandemic <https://t.co/6k7WzI2YjJ> [Fuente: *Conspiracy Theories Dataset: submuestra #plandemic, 2020-2023*, 21 de marzo de 2021]».

En cambio, en la submuestra #plandemia, parecen encontrarse con mayor frecuencia hashtags asociados al movimiento antivacuna o antimascarillas siendo, por ejemplo, el hashtag más destacado el de #yomevacuno, o destacando como tercer token más numeroso el emoji que representa una jeringuilla (💉). También sobresalen las palabras “vacunas” y “vacuna” entre las más frecuentes de los datos relativos a #plandemia. Destacan, igualmente, las etiquetas “#dictaduras sanitaria” y “#agenda2030” entre los tokens más frecuentes en la submuestra #plandemia, asociados, como veremos en un apartado posterior, a la argumentación de que existe un plan o una conspiración del gobierno para controlar a la población. Se trata de dos términos comúnmente usados junto con otros representativos de otras teorías

de la conspiración y que rechazan las medidas sanitarias adoptadas para combatir la pandemia de COVID-19 (el uso de mascarillas, el distanciamiento social y las vacunas). Estos elementos de la narrativa se encuentran igualmente junto a discursos que defienden la libertad individual (Figura 3) y comúnmente van acompañados de la difusión de información errónea sobre la pandemia de COVID-19.

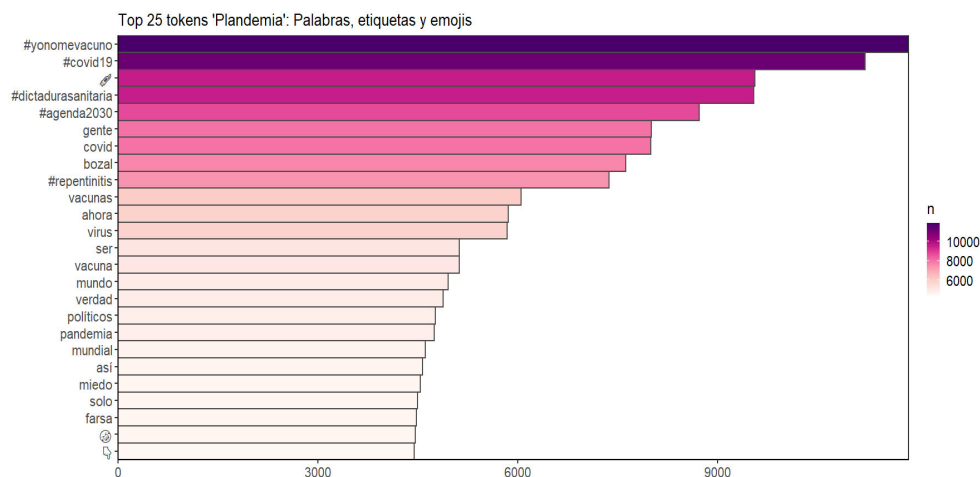
«@dangarcar #plandemia #pandemia #despierta #dictadurasanitaria la excusa para un #nuevoordenmundial globalista orweliano dictadura. Usan esta farsa mediática y el miedo para imponer agenda política. Ingeniería social contra la humanidad. La gente muere por otras causas. Están manipulándote! <https://t.co/sEgMAAsulQ> [Fuente: *Conspiracy Theories Dataset: submuestra #plandemic, 2020-2023*, 3 de marzo de 2021]».

Figura 2. Top 25 tokens “Plandemics” (palabras y hashtags)




Fuente: *Conspiracy Theories Dataset: submuestra #plandemic, 2020-2023*.

Figura 3. Top 25 tokens “Plandemia” (etiquetas, palabras y emojis)











Fuente: *Conspiracy Theories Dataset: submuestra #plandemia, 2020-2023*.

Si entre los principales tokens del conjunto de datos #plandemic no aparecía ningún emoji (el primer emoji lo encontramos en la posición 86), en el caso del conjunto español, en tercera posición se encuentra el emoji que representa la jeringuilla, lo cual nos remite a la coherencia interna de la argumentación, pues normalmente es un emoji que acompaña a textos relativos a la negación de la vacuna, como el ejemplo que sigue:

«@taitireme no somos ignorantes, solo que hemos visto los efectos secundarios terribles que han tenido algunas personas inoculadas, incluso algunos han muerto. sencillamente creemos que no es necesario una  para un virus con la letalidad menor al 1.8% #yonomevacuno <https://t.co/lxvaz9u1d7> [Fuente: *Conspiracy Theories Dataset: submuestra #plandemia, 2020-2023*, 7 de abril de 2021]».

Esta mayor orientación de los discursos en español hacia la negación de la vacunación en conexión con la etiqueta #plandemia se ve reflejada en otros tokens que se encuentran entre los principales, como la palabra “bozal”. Discursos que también con frecuencia son reforzados, en este caso, con el emoji enfundado con mascarilla:

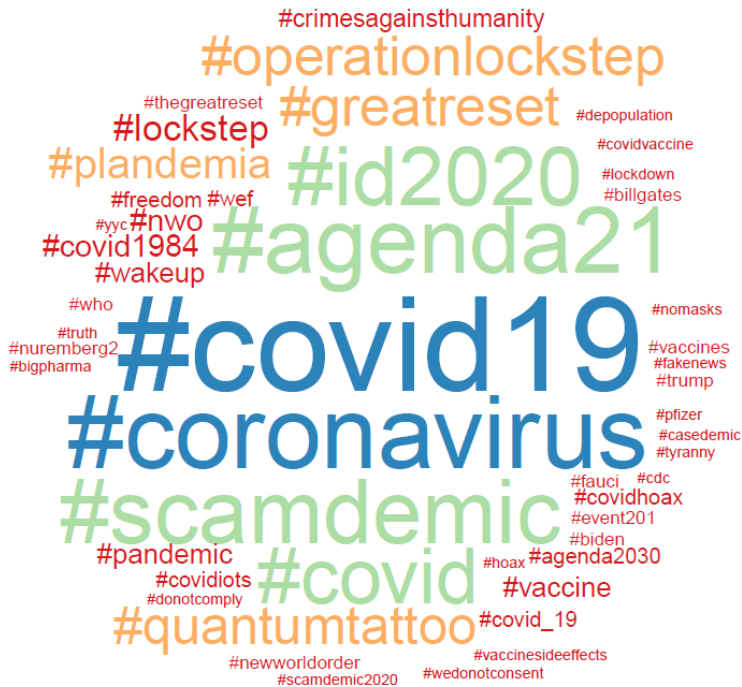
«@isaraosrnv eso que ni que.sin duda me imagino que te has de sentir incómoda estando embozalada  por mucho tiempo con este bozal  llamado tapabocas  pelando solamente los ojos  pero no te queda de otra que usar este bozal  llamado tapabocas  venga @ sonreír con la mirada  ánimo  [Fuente: *Conspiracy Theories Dataset: submuestra #plandemia, 2020-2023*, 7 de abril de 2021]».

## 6. MICRONARRATIVAS DE HASHTAGS EN TORNO A LA PLANDEMIA

Estudiamos ahora, a partir de los hashtags que contienen los tuits, qué tipo de narrativas se han construido en español e inglés durante la pandemia en torno a la teoría de la conspiración que se refiere a la “plandemia”. Aunque el número de hashtags en cada dataset es muy elevado (Cuadro 1), hay un conjunto de ellos más ampliamente difundidos y articulan el discurso que predominantemente se produce en torno a la llamada plandemia. Conocer estos hashtags y cómo conectan entre sí, establecer cuáles son las principales co-ocurrencias en cada dataset (de forma similar a Wei et al., 2023 o Hemsley, et al. 2020), permite observar si estas micronarrativas parecen estar operando de la misma forma en ambas lenguas. Su identificación posibilita avanzar en el conocimiento de la desinformación que se ha propagado sobre la pandemia, así como de las teorías de la conspiración asociadas a esta.



Figura 5. Principales hashtags en el conjunto de datos de #plandemic



Fuente: *Conspiracy Theories Dataset: submuestra #plandemic, 2020-2023*. Elaboradas en R con ayuda de los paquetes *quanteda* y *ggplot2*.

Un conjunto de narrativas explícitamente niega la pandemia de COVID-19 y argumentan que esta es falsa. Indican que hay una supuesta manipulación de los datos sobre la pandemia, lo que hace que sea una farsa, un engaño o una estafa. Los hashtags articulan las micronarrativas que difunden información errónea y desinformación sobre la pandemia de COVID-19. En inglés, por ejemplo, uno de los que aparecen más frecuentemente es el de #scamdemic, que hace clara alusión a la idea del engaño o fraude de la pandemia. Se emplea junto a otros para apoyar esta teoría de la conspiración. Con #covidhoax de nuevo aparece la idea de fraude y estafa. En español #falsapandemia se usa para difundir la idea de que la pandemia de COVID-19 no existe, es falsa, es un engaño, o no es real, y con #coronatimo se denuncia de nuevo que la pandemia es una estafa. Estos hashtags frecuentemente comparten espacio con otros que van insistiendo en esta idea de la falsedad (como es el caso de #fakenews o #hoax), o de la manipulación de los datos.

A veces argumentan que el COVID-19 es una enfermedad menor o que los síntomas se han exagerado o son inexistentes. Otra etiqueta frecuente es #circovid, que juega con la idea del “circo del COVID” y se utiliza a veces para

negar la pandemia y otras veces para ridiculizar las medidas de salud pública para prevenir la propagación del COVID-19, rechazando las medidas sanitarias que se toman, o subrayando que son innecesarias o incluso contraproducentes por lo estrictas que son o porque violan los derechos individuales.

Otro eje de los discursos, común en el movimiento antivacunas, se expresa a través de hashtags que claramente evidencian la resistencia y rechazo a la vacunación contra la COVID-19. En algunos casos, se alude a la peligrosidad de las vacunas (#nomasafp), a veces de forma muy extrema porque se argumenta que pueden provocar muertes repentinas (#repentinitis), mientras que en otras ocasiones se destaca el derecho a la libertad individual.

Un conjunto de etiquetas es utilizado para expresar oposición a la vacunación (#yonomevacuno), para negar los pasaportes de vacunación o los pases de movilidad (#noalpasesanitario, #noalpasedemovilidad) o para oponerse a los mandatos de vacunación. Los pasaportes de vacunación registran el estado de vacunación de una persona y se suelen usar para permitir que las personas viajen o asistan a eventos. Hashtags como #noalpasesanitario y #noalpasedemovilidad son similares en cuanto a que ambos expresan oposición al uso de documentos que certifican el estado de vacunación de una persona, pero se diferencian en que #noalpasesanitario se refiere específicamente a los pasaportes de vacunación, mientras que #noalpasedemovilidad se refiere a los pases de movilidad, que son un registro digital del estado de vacunación de una persona o evidencian una prueba de COVID-19 negativa y se usan para permitir que las personas viajen en transporte público. En inglés, #novaccinepassport y #novaccine mandate son similares en que ambos expresan oposición a los requisitos de vacunación, en el primer caso, para oponerse al uso de pasaportes de vacunación, mientras que en el segundo para oponerse a los mandatos de vacunación. No solo se rechaza la idea de tener que usar certificados o pasaportes de vacunación, igualmente con algunos como #mividamioxigeno se argumenta que las personas en realidad no necesitan vacunas para protegerse del COVID-19.

La lista es extensa. Hay otros hashtags relevantes en esta serie de datos que igualmente se utilizan para oponerse a la vacunación. Por ejemplo, #noalnom, rechaza el “nom” (la norma oficial mexicana), que es una regulación que exige que ciertos trabajadores se vacunen contra el COVID-19. En la misma línea de rechazo a la vacunación se encuentran otras etiquetas como: #novaccine, que se utiliza por personas que se oponen a la vacunación en general (o #novax, como versión abreviada). Encontramos también el hashtag #aur, Acción Humanista Revolucionaria, que es el nombre de un movimiento político en Bolivia que se opone a la vacunación, así como a otras medidas de salud pública relacionadas con la pandemia, como el uso de mascarillas y el distanciamiento social,

alegando que estas medidas son una violación de los derechos civiles<sup>4</sup>. Otra etiqueta de uso frecuente es #losniñosnosetocan, a través de la que se denuncia la vacunación obligatoria de los niños.

Por otra parte, encontramos hashtags asociados a teorías de la conspiración relacionadas con las vacunas. Destaca #Pfizergate, como referencia a una teoría de la conspiración que alega que Pfizer, la empresa farmacéutica que desarrolló una de las vacunas COVID-19, participó en prácticas fraudulentas o poco éticas para obtener la aprobación de su vacuna, en la línea de acusar a los gobiernos o a las farmacéuticas de estar detrás de la pandemia. Se trata de una etiqueta que se utiliza para denunciar los supuestos escándalos de la farmacéutica Pfizer. De forma más extrema, se acusa a los gobiernos de los efectos negativos de las vacunas, entre otros, se les acusa de las muertes repentinas o efectos secundarios de las vacunas. Es el caso de la etiqueta #repentinitis, que se emplea para denunciar los efectos negativos de la vacunación. Otros ejemplos en esta línea discursiva en el que se encuentra un conjunto amplio de etiquetas es #vaccine-deaths, en la serie en inglés:

«“safe and effective” 🤖🌐 #coronavirus #covid\_19 #covidvaccine #vaccineswork #vaccinated #depopulation #oxfordastrazeneca #mrna #biontech #eugenics #moderna #idonotconsent #transhumanism #technocracy #nwo #covidvaccineispoison #killedbycovidvax #vaccinedeaths <https://t.co/upscew4bu4>» [Fuente: *Conspiracy Theories Dataset: submuestra #plandemia, 2020-2023*, 11 de marzo de 2021].

En otro orden de cosas, pero asociado al movimiento antivacunas, se encuentran también, conversando sobre la “plandemia”, expresiones del movimiento antimascarillas, que conectan con los anteriores y se expresan a través de etiquetas como: #nomasafp, #nomask, #bozaltóxico o #noalbozal que se utilizan para denunciar el uso de las mascarillas, una de las medidas que se imponen durante la pandemia.

Mostrando una línea discursiva algo diferente, hallamos un conjunto de etiquetas que defienden otras teorías conspirativas relacionadas con la COVID-19. Hay varias relevantes, expresadas a través de los hashtags #nuevoordenmundial, #nom y #nwo, que se refieren a una teoría de la conspiración sobre un supuesto plan de una élite global para controlar el mundo. Sin embargo, hay algunas diferencias clave entre los tres hashtags. Mientras que #nuevoordenmundial (o su abreviatura #nom) se refiere a cualquier teoría de la conspiración sobre la supuesta implantación de un “Nuevo Orden Mundial”, del mismo modo que #nwo, “new world order”, en su vertiente en inglés, #nom también se usa en

<sup>4</sup> <https://boliviaverifica.bo/aur-el-movimiento-detras-de-teorias-conspirativas-sobre-barbijos-y-la-cuarentenas-que-circulan-en-cochabamba>.

español para referirse específicamente a una regulación mexicana que exige que ciertos trabajadores se vacunen contra la COVID-19.

Otras etiquetas que aparecen en el mismo contexto discursivo que #plandemia o #nwo son las de “#agenda2030” (o #agenda21 en inglés), que se emplean junto a la teoría de la conspiración de la “plandemia” para resaltar que estos planes globales (Agenda 21 y Agenda 2030) forman parte de un complot global para controlar el mundo, incluyéndose aquí el control de la pandemia y todo lo que esto supone.

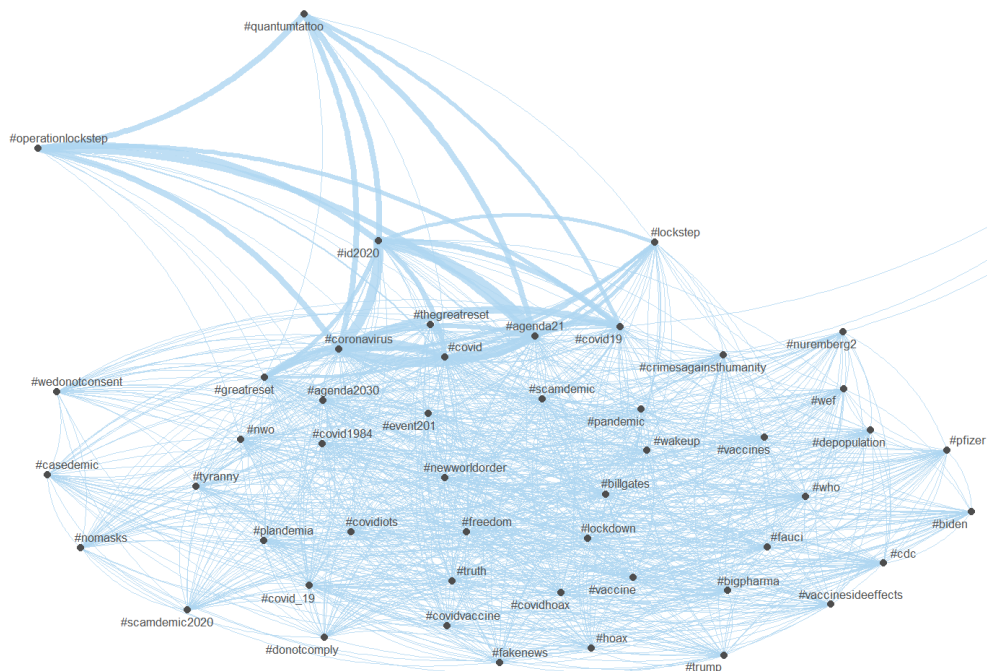
Además de #agenda21 o #agenda2030, encontramos otras etiquetas como #wef (en alusión al World Economic Forum). Estas son representativas de algunos grupos de teóricos de la conspiración que creen que estas agendas, impulsadas por las Naciones Unidas para promover el desarrollo sostenible, son una vía de control de los ciudadanos. #agenda2030 se utiliza para denunciar la supuesta agenda globalista que se está implementando a través de la COVID-19 y que todas las medidas dispuestas en la pandemia son un plan para controlar el mundo. Igualmente, por ejemplo, #rechazoagenda2030 se utiliza para denunciar las disposiciones de la agenda 2030. Del mismo modo, destacando esta idea de que la pandemia es una conspiración orquestada por una élite global para controlar a la población y establecer un nuevo orden mundial, circulan otros hashtags que las teorías de la conspiración conectan con reformas económicas como #id2020 (en referencia a la identificación digital), #Lockstep u #OperationLockstep (en relación con un informe del Foro Económico Mundial que imagina un futuro en el que los gobiernos utilizan la tecnología para controlar a la población) o #GreatReset, que alude a una iniciativa de reforma económica propuesta por el Foro Económico Mundial, que igualmente es percibida como un intento de control a la población (Christensen & Au, 2023).

Complementariamente, se lanza la sospecha de que #laverdadesunasola como etiqueta que se utiliza para expresar la creencia de que hay una verdad única sobre la pandemia, que está siendo ocultada por los gobiernos. Otro aspecto que genera desconfianza va unido a la etiqueta #depopulation para describir la creencia de que la pandemia es un intento deliberado de reducir la población mundial. Con mayor extremismo encontramos #genocidio, que se emplea para describir la pandemia como un intento deliberado de exterminar a un grupo de personas. O, de forma similar, #crimesagainsthumanity denunciando las violaciones de los derechos humanos que se han producido durante la pandemia. Como se aprecia a continuación, en muchos tuits se articula una narrativa de hashtags que comparten espacio con los que identifican la “plandemia” (en inglés o español):

«#stoptheshots #endthemandates #Health #Healthcare crisis #plandemic #Agenda2030 #depopulation #Genocide #crimesagainsthumanity



Figura 7. Red semántica de los principales hashtags en el conjunto de datos de #plandemic



Fuente: *Conspiracy Theories Dataset: submuestras #plandemic y #plandemia, 2020-2023*. Elaboradas en R con *quanteda*.

Otra etiqueta significativa en español referida al control es la de #covid1984, que compara la pandemia de COVID-19 con la novela distópica 1984 de George Orwell en la que el gobierno totalitario controla a la población mediante la vigilancia, la propaganda y la represión, lo que a través de este hashtag se hace equivalente a que las medidas de salud pública para prevenir la propagación del COVID-19 sean similares a las medidas de control que se describen en la novela, incluyendo la vacunación obligatoria.

Esta misma línea de resaltar asociar la “plandemia” con el control la encontramos en otras etiquetas como #infektadura, neologismo que combina las palabras “infección” y “dictadura” para describir la creencia de que los gobiernos están utilizando la pandemia para establecer una dictadura, o #controlavirus que expresar la creencia de que los gobiernos están utilizando la pandemia para controlar a la población o, entre otras, #tyranny que se utiliza para difundir la idea de que los gobiernos están utilizando la pandemia para establecer una tiranía. Los gobiernos, vienen a argumentar, están utilizando la pandemia para controlar a los ciudadanos, las medidas de salud pública, como los mandatos de mascarillas y las vacunas, son una forma de opresión. En otros casos se desa-

rolla la idea de que la pandemia es una conspiración para reducir la población mundial o para crear un nuevo orden mundial. Se encuentra también la teoría de la conspiración que acusa a Bill Gates de ser un terrorista biológico que está detrás de la pandemia (#billgatesbioterrorist o #billgates), o incluso a causa de la tecnología #5G. En esta línea, encontramos también #quantumtattoo, como teoría de la conspiración que afirma que las vacunas contienen microchips que pueden ser utilizados para rastrear a las personas.

Por último, además de otras etiquetas con argumentos similares que sería muy prolijo describir aquí, hay que mencionar que se introducen en la conversación insultos frecuentes como #covidiot o #covidota (palabra creada en 2020). Estos términos se emplean de forma crítica contra las personas que no siguen las reglas sanitarias establecidas en la pandemia, poniendo en riesgo a los demás. Se trata de un insulto dirigido precisamente a personas que difunden teorías de la conspiración sobre la plandemia, pertenecientes a los movimientos anticuarentena, antivacunas o antimascarillas. Esta misma práctica de insultar la encontramos también por parte de los defensores de la “plandemia” expresada a través de múltiples palabras, hashtags o emojis: “borrego”, “dormido”, “ganado”, “dumb”, “cobayas”, “subnormales”, “sheep”, “zombies”, “criminals” y muchos otros. En este contexto frecuentemente se apela a la necesidad de que “los borregos” o “los dormidos” se despierten (#wakeup) y se den cuenta de la conspiración global que supone la “plandemia”.

## 7. DISCUSIÓN Y CONCLUSIONES

Quizás uno de los aspectos más destacados de las narrativas de la “plandemia” tiene que ver con la similitud de los argumentos que a través de elementos como tokens, hashtags o emojis hemos encontrado en los datos en español e inglés, con la salvedad de diferentes acentos en cuanto a la preponderancia de unos discursos frente a otros, o cuestiones nacionales y locales. Hemos encontrado más sesgadas las narrativas en español asociadas a #plandemia hacia el movimiento antivacunas. Pero es evidente tanto la internacionalización que ha alcanzado la idea de “plandemia”, que se mantiene en las conversaciones que hemos recogido entre 2020-2023 en las dos lenguas consideradas. Se aprecia la internacionalización principalmente en los mensajes que se escriben usando ambas lenguas, internacionalizando las narrativas y logrando un mayor alcance.

Cuesta trabajo resumir cuáles son las principales narrativas negacionistas sobre la “plandemia” de COVID-19 pues en los mensajes de Twitter, junto a la idea negacionista de que la pandemia responde a un plan urdido con el fin de controlar a la población, o que el COVID-19 no existe y la pandemia es falsa, se van mezclando corrientemente otros mensajes que dan la mano, a veces de forma explícita, otras de forma borrosa, a otras corrientes conexas

de pensamiento, principalmente alineadas con los movimientos antivacunas, antimascarillas y anticuarentena, y que, en gran medida, rechazan a diferentes gobiernos o políticas y medidas sanitarias derivadas de los mismos. Las redes de co-hashtags, como estrategia que permite visibilizar algunas de las micronarrativas preponderantes, han posibilitado observar esta multiplicidad de ideas que se comparten en el contexto discursivo de la llamada “plandemia”, que conectan los discursos de varios movimientos sociales que niegan la pandemia o se resisten a las medidas establecidas para controlarla.

Como limitación de este trabajo, subrayamos su enfoque analítico basado de forma genérica en dos lenguas, sin haber llevado a cabo un análisis desagregado por países que por la escasez de información geolocalizada por Twitter hubiera sido difícil, aunque trabajos posteriores podrán profundizar en esta línea, a través de los mensajes, para conocer más a fondo, cuando es posible, algunos detalles de las narrativas que derivan de cada contexto nacional. A pesar de esta objeción, se han encontrado patrones comunicativos similares, y grandes paralelismos en las narrativas en las dos lenguas, con palabras clave, hashtags o emojis equivalentes al traducirlas o incluso idénticas algunas veces (por ejemplo, con el uso común internacional de algunos emojis clave en el discurso).

Pensando en el futuro, no podemos dejar de hacer una mención igualmente, aunque no ha sido el objeto de este capítulo, a que las narrativas que hemos encontrado se encuentran frecuentemente cargadas de odio y de mensajes que desacreditan a las autoridades sanitarias, deslegitimándolas. La construcción de democracias deliberativas en sociedades con gran influencia de internet y de las plataformas de redes sociales pasa necesariamente por reducir el tono emocional y la falta de civismo que frecuentemente se aprecia en los mensajes que hemos recogido sobre la plandemia de COVID-19. De esta forma, ante el uso de insultos o el lenguaje ofensivo de muchos tweets que caracterizan la forma de expresión del movimiento negacionista, antivacuna, anticuarentena o antimascarillas, se convierte en un reto el desarrollo de un camino que debe potenciar el respeto mutuo. Igualmente, según se sugiere en Ahmed (2023), otras líneas son viables, como la de intentar deslegitimar las fuentes de información falsa para reducir su impacto o, por ejemplo, aprovechar la posibilidad de denunciar contenidos inapropiados en diferentes plataformas.

Por último, en la medida en que las creencias conspirativas a las que hemos aludido en el capítulo no están respaldadas por la evidencia científica y que estas teorías pueden tener un impacto negativo en aspectos tan relevantes como la salud pública o el crédito a las autoridades sanitarias (Nefes et al., 2023), las relaciones intergrupales (Douglas, 2021) y la pervivencia de la democracia (Muirhead y Rosenblum, 2022), es muy relevante ser muy críticos con la información que se encuentra en línea. La rápida y extensa propagación internacional de teorías de la conspiración sobre la pandemia, común en diferentes lenguas

y que difunden desinformación, mentiras y odio, ha presentado y presenta un importante riesgo para las sociedades que es preciso neutralizar. Este reto marca algunos de los pasos a dar para desarrollar una agenda de investigación e intervención en el horizonte más inmediato y a corto plazo.

## REFERENCIAS

- Ahmed, W.; Önköl, D.; Das, R.; Krishnan, S.; Olan, F.; Hardey, M. y Fenton, A. (2023). Developing Techniques to Support Technological Solutions to Disinformation by Analyzing Four Conspiracy Networks During COVID-19. *IEEE Transactions on Engineering Management*. <https://ieeexplore.ieee.org/abstract/document/10136885>
- Baines, A.; Ittefaq, M. y Abwao, M. (2021). #Scamdemic, #Plandemic, or #Scaredemic: What Parler Social Media Platform Tells Us about COVID-19 Vaccine. *Vaccines*, 9(5):421. <https://doi.org/10.3390/vaccines9050421>
- Benoit, K.; Watanabe, K.; Wang, H.; Nulty, P.; Obeng, A.; Müller, S. y Matsuo, A. (2018). *quanteda*: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30), 774. <https://doi.org/10.21105/joss.00774>, <https://quanteda.io>
- Çalışkan, C. (2022). Who is behind the Plandemic? Analysing the different ideological conspiracy theory endorsements on Turkish Twitter. *Southeast European and Black Sea Studies*, 1-20. <https://doi.org/10.1080/14683857.2022.2137900>
- Christensen, M. y Au, A. (2023). The Great Reset and the Cultural Boundaries of Conspiracy Theory. *International Journal Of Communication*, 17, 19. <https://ijoc.org/index.php/ijoc/article/view/20128>
- Comisión Europea (2022). Detección de las teorías conspiratorias. [https://commission.europa.eu/strategy-and-policy/coronavirus-response/fighting-disinformation/identifying-conspiracy-theories\\_es](https://commission.europa.eu/strategy-and-policy/coronavirus-response/fighting-disinformation/identifying-conspiracy-theories_es)
- Douglas, K. (2021). Are conspiracy theories harmless? *Spanish Journal of Psychology*, 24. <https://doi.org/10.1017/SJP.2021.10>
- Gualda, E., Krouwel, A., Palacios-Gálvez, M., Morales-Marente, E., Rodríguez-Pascual, I. y García-Navarro, E.B. (2021). Social Distancing and COVID-19: Factors Associated With Compliance With Social Distancing Norms in Spain. *Frontiers of Psychology*, 14;12:727225. <https://doi.org/10.3389/fpsyg.2021.727225>
- Hemsley, J., Erickson, I., Jarrahi, M.H. y Karami, A. (2020). Digital Nomads, Coworking, and Other Expressions of Mobile Work on Twitter. *First Monday*, 25. <https://doi.org/10.5210/fm.v25i3.10246>
- Jaworsky, B.N. (2023). Everything's going according to Plan(demic): a cultural sociological approach to conspiracy theorizing. *American Journal of Cultural Sociology*, 11, 26-49. <https://doi.org/10.1057/s41290-021-00144-z>

- Jikeli, G. y Soemer, K. (2023). The value of manual annotation in assessing trends of hate speech on social media: was antisemitism on the rise during the tumultuous weeks of Elon Musk’s Twitter takeover? *Journal of Computational Social Science*. <https://doi.org/10.1007/s42001-023-00219-6>
- Jones-Jang, S.M., y Chung, M. (2022). Can we blame social media for polarization? Counter-evidence against filter bubble claims during the COVID-19 pandemic. *New Media & Society*, 0(0). <https://doi.org/10.1177/14614448221099591>
- Memetic.Media (2020). Tres mentiras del documental «Plandemic». *Elsurti.com*. <https://elsurti.com/coronavirus/chequeo/2020/05/12/plandemic-falso-coronavirus-paraguay/>
- Muirhead, R. y Rosenblum, N.L. (2022). Conspiracy in America. Recognizing and Confronting Assaults on Our Democracy. *AFT Health Care*. [https://www.aft.org/hc/fall2022/muirhead\\_rosenblum](https://www.aft.org/hc/fall2022/muirhead_rosenblum)
- Nassauer, A. y Legewie, N.M. (2020). Methodologische Entwicklungen in der Gewaltforschung. *Österreichische Zeitschrift für Soziologie*, 45 (Suppl 1), 135-156. <https://doi.org/10.1007/s11614-020-00412-1>
- Nazar, S. y Pieters T. (2021). Plandemic Revisited: A Product of Planned Disinformation Amplifying the COVID-19 “infodemic”. *Frontiers in Public Health*, vol. 9. <https://www.frontiersin.org/articles/10.3389/fpubh.2021.649930>
- Nefes, T.S., Präg, P., Romero-Reche, A. y Pereira-Puga, M. (2023). Believing in conspiracy theories in Spain during the COVID-19 pandemic: Drivers and public health implications. *Social Science & Medicine*, 336:116263. <https://doi.org/10.1016/j.socscimed.2023.116263>
- Neil, S.J.D. y Campbell, E.M. (2020). Fake Science: XMRV, COVID-19, and the Toxic Legacy of Dr. Judy Mikovits. *AIDS Research and Human Retroviruses*, 36(7), 545-549. <http://doi.org/10.1089/aid.2020.0095>
- Pan American Health Organization (2020). Understanding the infodemic and misinformation in the fight against COVID-19. Knowledge Tools. Digital Transformation Toolkit. [https://iris.paho.org/bitstream/handle/10665.2/52052/Factsheet-infodemic\\_eng.pdf](https://iris.paho.org/bitstream/handle/10665.2/52052/Factsheet-infodemic_eng.pdf)
- Pavela Banai, I., Banai, B. y Mikloušić, I. (2022). Beliefs in COVID-19 conspiracy theories, compliance with the preventive measures, and trust in government medical officials. *Current Psychology*, 41(10) 7448-7458. <https://doi.org/10.1007/s12144-021-01898-y>
- Sievert, C. (2020). *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC. <https://plotly-r.com>.
- Silge, J. y Robinson, D. (2022). *Text Mining with R. A tidy approach*. O’Reilly. <https://www.tidytextmining.com/>
- Sumner, S., Ferguson, B., Bason, B., Dink, J., Yard, E., Hertz, M., Hilker, B., Holland, K., Mercado-Cresco, M., Tang, S. y Jones, C. (2021). Association of online risk factors with subsequent youth suicide-related behaviours in the US. *JAMA Network Open*, 4(9). <https://doi.org/10.1001/jamanetworkopen.2021.25860>

- Wei, H., Hai, C., Shan, D., Lyu, B. y Wang, X. (2023). Text recognition and analysis of network public opinion focus events of a major epidemic: a case study of “COVID-19” in Sina Microblogs. *Multimedia Tools and Applications*, 82, 25811-25827. <https://doi.org/10.1007/s11042-023-14916-x>
- Welbers, K., Van Atteveldt, W. y Benoit, K. (2017). Text Analysis in R. *Communication Methods and Measures*, 11 (4), pp.245-265. <https://doi.org/10.1080/19312458.2017.1387238>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Golemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H, (2019). “Welcome to the tidyverse.” *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Xu, W. y Sasahara, K. (2022). A network-based approach to QAnon user dynamics and topic diversity during the COVID-19 infodemic. *APSIPA Transactions on Signal and Information Processing*, 11(2), Article e17. <https://doi.org/10.1561/116.00000055>
- Zhang, Y., Yue, Z., Yang, X., Chen, F. y Kwak, N. (2022). How a peripheral ideology becomes mainstream: Strategic performance, audience reaction, and news media amplification in the case of QAnon Twitter accounts. *New Media & Society*, 0(0). <https://doi.org/10.1177/14614448221137324>

# COVID-19 NARRATIVE THROUGH SOCIAL MEDIA: THE TELEGRAM CASE<sup>1</sup>

ALFONSO AMENDOLA

*Associate Professor of Sociology of Cultural Processes  
University of Salerno, Italy*

MARTINA MASULLO

*PhD Student in Politics and Communication (POLICOM)  
University of Salerno, Italy*

**DOI:** 10.14679/2756

## **Summary:**

1. The COVID-19 narrative through social: a general theoretical framework.
  2. Infodemic, communicative crisis and spontaneous self-management of communication.
  3. Telegram case: objectives, methodology and results.
  4. Conclusions.
- References.

## 1. THE COVID-19 NARRATIVE THROUGH SOCIAL: A GENERAL THEORETICAL FRAMEWORK

Between 2020 and 2021, through social networks, has developed a narrative of everyday life, that appears completely polarized towards a single macro-theme: the COVID-19 pandemic. At the same time, several micro-themes have developed – over the course of months and thus after becoming more familiar with the pandemic condition – which can also be defined as ‘dimensions’, as each of them has, in turn, given rise to a distinct imagery, with its own rules of belonging and trends. Therefore, the COVID-19 narrative through social networks has branched out into three main dimensions: the health dimension, the social dimension and the media dimension. The first one (health dimension) represents the original frame of the COVID-19 pandemic and includes the numbers and data of the contagions, the state of the national health care system and, subsequently, the updates regarding treatments and vaccines against the virus. The second one (social dimension) deals with the factors that have determined the new sociality generated after the covid: the rethinking of oneself within society by modifying most of the activities that previously presupposed a relationship with the other

---

<sup>1</sup> The essay is the result of the joint work of the authors: in particular, Alfonso Amendola dealt with § 1, 2; Martina Masullo § 3, 4.

and that in the post-covid become mostly personal activities. This includes all information related to the Prime Ministerial Decrees, official government communications with the consequent division of Italy into zones and social distancing rules. The third dimension (media dimension), finally, concerns the role the media played during the lockdown and the way they were rethought as well as their use adapted to the new individual needs. This dimension includes all those phenomena that have developed through the media and that, collaterally, were accelerated by the covid but whose development process was already underway: the explosion of streaming platforms, the new way of participating in online events, the collectivization of sport, the increasing use of video games and also the process of adapting study, training and work to the logic of smart working. In particular, the intention here is to focus on and analyze the way in which all this information – from the more technical and primary to the more creative and collateral – was communicated through different social platforms during the three months of the lockdown and those immediately following. These contents have given rise to news corpora grouped by specific topics that, replicated over time, have contributed to reinforcing certain ‘patterns of interpretation’ through which users have translated their ideas about a given fact or topic (Polesana and Risi, 2023). According to an American study by Romer and Jamieson (2020), conspiracy theories and fake news have had a strong impact not only on people’s individual behavior, but also on their collective behavior. In particular, the study shows that people who believed that the spread of the COVID-19 pandemic was due to the complicity of pharmaceutical companies or the Chinese government – two of the most widely pursued conspiracy theories – were more likely to have transgressed government regulations to contain the contagion (social distancing, use of safety devices and adherence to the vaccination plan). The preliminary idea of this contribution is, therefore, to identify a series of keywords and common traits that have characterized the narrative of the coronavirus through social networks in order to trace in them the origin of the main conspiracy theories and some cases of hate speech related to COVID-19 that have subsequently developed online, especially through the Telegram platform that – together with Facebook and the Whatsapp app – has contributed most to the spread of fake news on the covid issue. The use of social media became massive during the lockdown. This allowed conspiracy theories and fake news to spread more easily as they had a much wider audience (Marcelli and Valerii, 2021).

Closely related to the spread of conspiracy theories and fake news is the problem of hate speech, which is based on a process of dehumanization or distortion, the act of treating or perceiving a person or group of people as different and alien to oneself (Kelman and Hamilton, 1989). Dehumanization is a growing social problem since the new millennium in which “a migration crisis, political upheaval due to populism, misinformation and pandemics have increased feelings of insecurity and made the future unpredictable” (Bayer and Bård, 2020).

## 2. INFODEMIC, COMMUNICATIVE CRISIS AND SPONTANEOUS SELF-MANAGEMENT OF COMMUNICATION

Social media are today – together with television – the main platforms through which users inform themselves about the main contemporary topics. An analysis by Comscore, covering the year 2021, showed that 65% of respondents in the 18 to 24 age group (Generation Z) use digital media to a greater extent to inform themselves. And the trend is also growing for previous generations. However, information has never been so fragmented. During the pandemic and post-pandemic period, a huge amount of information about the virus and all related issues was produced through social media, which has contributed, as we have already pointed out, to generating a precise COVID-19-related narrative. A part of this information was verified and shared online by journalists and insiders, but another part of it, instead, turned out to be totally disintermediated (Pierangelo and Gorgoni, 2017) and allowed the condition of ‘infodemic’ – a neologism already in use since 2019 that originates from the union of the two words ‘information’ and ‘epidemic’ – to explode during the pandemic and to determine the health status of Italian information. If initially the large amount of information disseminated online had created a space for discussion on the possibility of preserving information pluralism that would allow people to construct their own opinion on the basis of different elements and voices, today it is making online information increasingly rarefied and compromised. The term infodemic (Manfredi, 2019; Ferrazzoli and Maga, 2021) indicates an overabundance of information that makes it difficult to distinguish reliable news from fake news. It is clear that during a crisis situation such as the management of epidemics, cushioning the risks related to misinformation can be more complex than usual since a number of elements must be taken into account that alter the normal conditions of perception and modify the processes of information acquisition. The World Health Organization (WHO) has declared the COVID-19 epidemic a serious ‘global risk’, recalling that – as predicted in the World Economic Forum’s 2017 Global Risk Report – these risks are interconnected. In particular, COVID-19 has resulted in a considerable increase in conspiracy theories, some of which underlie sometimes violent protests and demonstrations (Fabbri et al., 2022), and has brought to light the critical issues surrounding the dissemination of online information in a disintermediated news cycle. Such a phenomenon could accelerate or modify the epidemic process by ‘influencing and fragmenting the social response’ (Cinelli et al., 2020). In the case of the pandemic, in Italy but not only, the information leaked from non-authoritative sources and circulated mainly through social media provoked a social response that, in some cases, jeopardized the effectiveness of the measures put in place by the government to counter the contagion. The enormous difficulties in managing communication in an emergency provoked

a communication crisis that challenged the already unstable foundations on which online information rested. There are five fundamental elements that have characterized this crisis: the absence of coordination between official government sources and the press, the lack of coordination between official sources and the scientific community, the exasperation of tones on social networks such as Facebook, Instagram, and Twitter (which then caused the displacement of users onto Telegram channels and thematic chats), the construction and dissemination of fake news related to the covid theme, and, finally, a trivialization and underestimation of the effectiveness and need for a crisis management plan for communication (Borgonovo, 2020). On the various social networking platforms, digital micro-communities were formed that began to produce specific content on COVID-19 in which information was polarized towards a single or certain aspects of the issue: those bodies of news discussed at the beginning took shape and, at the same time, began to guide the interpretation of online information on the coronavirus. The various communities showed, from the outset, behavior that can be traced back to a form of 'spontaneous self-management of communication' (Borgonovo, 2020).

A distinctive feature of digital micro-communities is their capacity for self-management in communication. Since these groups are generally made up of individuals who share common interests, the members themselves organize and manage communication spontaneously. Self-management of communication in digital micro-communities is possible thanks to instant messaging platforms such as Telegram, WhatsApp or Discord, which offer a wide range of tools to facilitate interaction between members. In these micro-communities, communication is based on principles of equality, reciprocity and cooperation. Individuals are free to express their opinions, share relevant content and propose new ideas without hierarchical restrictions. It is the members themselves who decide which topics to discuss, how to organize events and what kind of content to share. Furthermore, self-management of communication encourages strong participation and greater individual responsibility. Each member can actively contribute to the conversation, offer support to other members and propose solutions to challenges the community may face. Digital micro-communities are also an inclusive environment where individuals can find a sense of belonging and connection with people who share the same interests. Through self-managed communication, these communities enable the emergence of online social support networks that can be particularly important in contexts where face-to-face interaction is limited or challenging. However, it is important to emphasize that self-management of communication may entail challenges. Without formal guidance, disputes or conflicts between members may arise. It is necessary to have clear guidelines and respect the principle of mutual respect to maintain a positive and constructive environment within the community. The spontaneous self-management of communication within digital micro-

communities allows for active participation of members, mutual support and the emergence of meaningful social networks. These groups demonstrate how technology can facilitate the involvement and connection of people, enabling them to form communities based on shared interests and managed autonomously.

Specifically, we can identify 8 points of development and digital communities:

1. **Autonomous community:** Spontaneous self-management is based on the idea of digital communities that are self-organizing and self-regulating without formal leadership. People freely gather around common interests and make decisions together.
2. **Voluntary participation:** Self-management is based on the voluntary participation of the people concerned. Each community member has the freedom to participate or withdraw according to his or her own needs and interests.
3. **Autonomous decision-making:** In self-managed digital micro-communities, decisions are made collaboratively and horizontally. There is no central authority figure, but rather a collective decision-making process involving all members.
4. **Shared knowledge:** Spontaneous self-management is based on the idea that each member of the community has something to contribute. Everyone's knowledge, skills and experience are shared and used for the common good.
5. **Transparency and openness:** In self-managed digital micro-communities, transparency and openness are important. Information is accessible to all members, and decision-making processes and community activities are open to the participation of all who wish to contribute.
6. **Shared responsibility:** Each member of the community is responsible both for himself/herself and for the well-being and functioning of the entire community. This means that members support each other and are committed to working together for the collective good.
7. **Trust and reciprocity:** Spontaneous self-management is based on mutual trust between community members. Individuals trust that others will act responsibly and contribute to the common good.
8. **Agility and adaptability:** Self-managed digital micro-communities are often characterized by the ability to adapt quickly to changes and make decisions in an agile manner. This flexibility allows the community to evolve and respond to emerging needs.

Michel Maffesoli (1988) theorizes a conception of the tribe that can easily be reinterpreted in the light of contemporary communities that take up its legacy and digitize its boundaries.

If the concept of proximity, thanks to the network, has overcome most of its physical criticalities, at the same time social networks turn out to be precisely circumscribed (digital) environments where the sharing of certain values is favored by that condition of homophily (Boccia Artieri, 2012) that makes it complicated, indeed highly unlikely, that in a flow of highly polarized information an idea or opinion contrasting with it can be inserted and challenged.

The digital micro-communities that have arisen during the pandemic follow exactly the same logic of cohesion: its members meet and group together like the tribes described by Maffesoli and give rise to closed, sealed dimensions in which visions and experiences stagnate following the laws of homophily as Boccia Artieri theorized it. It is also for these reasons that conspiracy theories and fake news constructed during the pandemic have spread so quickly and so widely: the mechanisms that regulate the algorithms of social platforms tend to propose content that is always similar to that in which users have shown the most interest, thus increasing the formation of echo chambers that on social media lead to a state of ideological isolation of the user. This gives rise to confirmatory biases that increase the power of fake news and its impact on online opinion formation (Quattrociocchi and Vicini, 2017). The further risk is the formation of true echo platforms (Quattrociocchi et al., 2022) that further foster the fake news phenomenon in every field of action.

Among the various platforms, Telegram is perhaps the social network that, both in terms of structure and policy, is better suited than the others to this condition and where confrontation with perspectives other than one's own is unlikely.

### 3. TELEGRAM CASE: OBJECTIVES, METHODOLOGY AND RESULTS

Telegram is a platform founded in 2013 by Russian entrepreneur Pavel Durov and his brother Nikolai, halfway between an application designed for instant messaging and a social network. It is characterized by its group chats that can host up to 200,000 members and its channels, which can be followed by an unlimited number of users, where there is a continuous flow of images, videos and content of all kinds: the app is currently used by around 550 million people worldwide. We decided to focus on Telegram because among the social networking platforms, it is the one that guarantees the greatest freedom in the dissemination of content to its users, since the company provides for minimal moderation of discussions and has refused, from the outset, to collaborate with governments that demand news about its members. This guarantees enormous freedom in the dissemination of all kinds of news and information and makes it the platform of choice for activists, journalists, organizations and

social and protest movements. During the pandemic, Telegram provided a fast and effective means of communication for people all over the world. Many Telegram groups and channels were created to share information and news about the pandemic situation in real time. This allowed people to stay up-to-date on safety protocols, local restrictions and the latest medical findings related to the virus. In addition, Telegram provided an important platform for virtual socializing during social distancing. Users were able to connect with friends and family through Telegram's group chats and video calls, giving them a sense of closeness and emotional support despite the physical distance. Telegram also facilitated the sharing of personal experiences, opinions and advice on the pandemic. Many people used the platform to tell their stories, exchange information about symptoms and treatment, and offer mutual support during this difficult time. In essence, Telegram played a significant role in influencing the pandemic imagination, providing a platform for communication, sharing and connection during complex times.

In recent years, the advent of new communication technologies has revolutionized the way people interact. The global pandemic has further accelerated this process, as individuals and communities constantly seek new and innovative ways to stay connected. In this context, Telegram, one of the world's most popular instant messaging platforms, has played a significant role in shaping pandemic and post-pandemic imagery.

Specifically, we can identify four nodal points:

1. The power of virtual connection

During the pandemic era, Telegram provided people with a crucial tool to maintain social connection despite the physical distancing required by security measures. Telegram's group chats, information channels and video calls allowed people to share information, emotions and support each other during times of crisis. This virtual connection has played a key role in maintaining people's emotional balance, and has become an essential alternative to face-to-face interactions.

2. Combating infodemics and disseminating accurate information

During the pandemic, Telegram also faced the challenge of infodemic: the dissemination of false or unverified information. The platform took measures to combat misinformation, such as introducing a self-destructing message sharing function and tagging suspicious or false content. These efforts were important in providing a safer and more reliable environment for users seeking accurate information on the epidemic, prevention and health guidelines.

3. Preparedness and adaptation in the post-pandemic environment

As the world approached a post-pandemic phase, Telegram proved to be a valuable ally in the preparation and adaptation phase. Health organizations, governments and communities use Telegram to share information on available vaccinations, epidemiological data, travel policies and evolving guidelines. This instant messaging platform provided a rapid and direct exchange of information that helped ensure effective post-pandemic management, easing the transition to a new ordinary life.

#### 4. Confidence and resilience in the digital age

The pandemic and post-pandemic imagery has also been influenced by the trust and resilience that users around the world have shown in using Telegram as a communication tool. The large-scale adoption of this platform was driven by the need to cope with isolation, to experiment with new ways of working and studying remotely, but also to create a feeling of unity and social cohesion in times of uncertainty. The resilience demonstrated by Telegram users has undoubtedly influenced the collective imagination, paving the way for a digitally connected and resilient future.

However, at the same time, a content exchange as dense as the one that took place on Telegram during the pandemic years also showed the risks of the daily use of this platform: the spread of fake news and ambiguous content in a digital environment that, while shedding light on the inherent potential of social networks, also highlighted as many critical issues to be taken into account for future analysis.

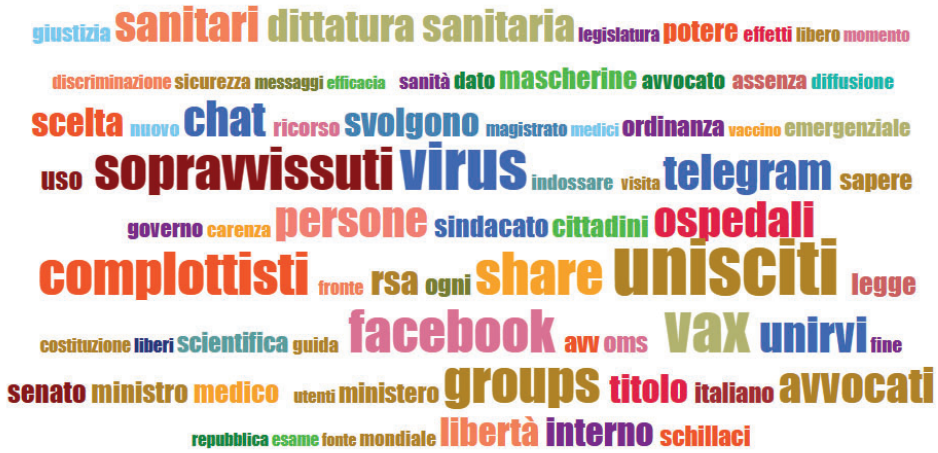
Therefore, the aim of our paper is to trace and analyze some of the main fake news, conspiracy theories and hate speech that have been generated on Telegram in particular and that have contributed to consolidate the disinformation imaginary that has developed between 2020 and 2021 around the topic of covid and to completely transform the society in which we live (Affuso et alii., 2020). Especially on the assumption that, according to preliminary data, Telegram is among the platforms most responsible for the rise of fake news and misinformation on covid. The observation of channel conversations and the analysis of the content and interactions that took place within Telegram channels were crucial not only for a semantic study of the keywords that guided discourses between users, but above all for understanding the context in which these took place.

During the first part of the pandemic, in the first months of the lockdown, in 2020, the Italian groups and channels that sprung up on Telegram were useful for gathering data on the situation Italy was experiencing by analyzing users' impressions and behavior. It was mostly a matter of keeping track of the number of contagions and the incessant changes in terms of the rules to be followed to avoid contagion. On 22 February 2020, the channel 'Tutto il

Corona virus minuto per minuto’ was created, which within a few weeks had 102,416 members and made use of the poll function to collect the opinions of its members. On 24 February 2020, ‘Italia 24 Flash’ was launched, which quickly proved to be a solid community and extended beyond Telegram. The ‘Merate/CoronaVirus’ channel, created on 4 March 2020, is a remarkable example of local institutional communication, while ‘Sapienza Coronavirus News’ is one of the most interesting cases of university emergency management. Finally, on 31 January 2020, the “CORONAVIRUS NEWS // ITA” channel was opened, which, unlike the channels born just before or immediately after the lockdown decree, seems to have a more conscious and long-term objective: to brand information on COVID-19 in Italy and to build up the community’s loyalty in order to make it stable beyond the emergency.

It is possible to identify the beginning of the second part of the pandemic with the moment when the vaccination campaign started between the end of 2020 and the beginning of 2021. Not only from a health point of view, but also from a communication point of view, this event represented a kind of point break in the representation of the covid through the media. The narrative frame shifted from the virus-theme to the vaccine-theme. This was the time when the no-vaccine movements were born, most of which first formed online and then crossed digital boundaries, producing important effects in everyday life as well: from protests in the streets to health workers opposing the administration of vaccines, to forging green passes to gain access to social activities. The Telegram channels and chats examined were created in 2021 and are still active today, although the frequency of messages has thinned out over time. After taking part in the channels, we started a non-participatory observation of the chats and were able to identify the most frequently used keywords in a time span of about two weeks, the last in June 2023, for the Telegram channel “*Noi i no vax, i complottisti, i sopravvissuti*” with 11,261 members and about two months, from June to July 2023, for the channel “*Vax: le cavie siamo noi?*” with 9755 members.

Image 1. Word cloud from Telegram channel 1



Source: Telegram. Observation period: June 2023.

As can be seen from the graph, some of the most frequently used words within the channel are ‘virus’, ‘health dictatorship’, ‘vax’ and ‘survivors’, which have become the key words characterizing the COVID-19 narrative via social media during 2021.

Image 2. Word cloud from Telegram channel 2



Source: Telegram. Observation period: June - July 2023.

Also in the second graph, it is possible to find some of the keywords already noted above such as ‘vax’. Interestingly, the other words we find in both graphs – and thus in both observed groups – are ‘freedom’ or ‘free’ and ‘lawyers’ or

‘doctors’. It seems clear that the intent of the users interacting on these channels is to prove with ambiguously scientifically defined testimonies and documents that covid vaccines are a way to control the population and deprive them of their freedom of choice. And this is one of the main conspiracy theories that have emerged during the pandemic.

It is clear that within this cloud it is possible to trace words that, although not contextualized, can be semantically linked to those ‘alternative aleturgies’ Fausto Colombo refers to in his ‘Verità e democrazia. Sulle orme di Michel Foucault’ (2022), describing them as fundamental to the construction of truths, since without a negative force there could be no positive one. Colombo does not refer to the terms ‘positive’ and ‘negative’ in an assertive and biased manner. However, it would be impossible here not to detect in the construction of the narrative we are trying to trace a negative or, at the very least, impactful purpose for society. The imagery that Telegram has helped to generate around the COVID-19 macro-theme has had profound consequences in the real lives of people who online have spread or come across the various fake news and conspiracy theories that have been written about.

Therefore, an important research challenge is to try to identify the mechanisms by which users seek out or, conversely, avoid information online that then influences their decisions offline (such as whether to vaccinate or not). The case of the COVID-19 pandemic has, first and foremost, shown the critical impact that these new information environments have had on society.

At the same time, the general observation of some Telegram channels appeared between 2020 and 2021 has brought to light a further figure related to online hate speech: it is possible to divide by stages the wave of hate speech that triggered during the pandemic. If, in the first part of 2020, China and the Chinese were the main targets of the haters for having created «a biological weapon», already from the end of 2020 until 2021, the attention has almost totally shifted towards journalists, doctors and institutions guilty of dramatizing the situation and sowing terror and panic among the population. During the second year of the pandemic the perception that the covid, in reality, was not as serious as it had been the initial prospect has grown considerably, going to undermine even more the image of health workers who have transformed, in the collective imagination, from heroes to villains.

A poll dating back to October 2020 showed that only 6.5% of Italians believed that the pandemic was built to justify some political and economic choices (Bucchi and Saracino, 2020). But the situation has changed a lot already in 2020. From March to November 2020, the percentage of Italians who carried on that idea that only the elderly and fragile people died of COVID-19 increased from 15% to 39% (Stefanoni, 2020). Feelings of hatred towards health workers intensified through social media already in the second half of 2020, to

explode at the beginning of 2021 (Viafora, 2020) even offline. The symbol of overcoming the digital boundaries of this current of hatred was certainly the famous graffiti that, in Milan, celebrated doctors and health workers as heroes that on November 2 was vandalized (Ranieri, 2020).

#### 4. CONCLUSIONS

According to the data, it is safe to say that Telegram, together with Facebook and WhatsApp, was one of the platforms that, during the pandemic, strongly influenced – actively – that flow of disinformation linked to fake news, conspiracy theories and hate speech that developed online and had strong offline repercussions. As we saw earlier, being organized in channels and thematic chats and providing a moderation of the content almost nothing, on Telegram the alternative narratives of a given event – as, in this case, the pandemic from COVID-19, tend to strengthen itself by using, among others, especially to that concept of homophilia that comes offline but is perfectly suited to online communities and social environments so closed and stagnant. Most of the actions created online have also materialized outside the digital borders, generating significant spillovers and also influencing the effectiveness of campaigns to contain contagion put in place by governments and supported by the scientific community. The narrative of the pandemic, especially through social platforms such as Telegram, which has been the subject of our analysis, have helped to consolidate that concept of information and disintermediation and make it a problem, not only for insiders, but for all those who want to approach the analysis or interpretation of the dynamics related to the communication of emergency COVID-19. Therefore, the study of the new platforms in relation to the consumption of hot topics, that is those contents defined «hot» as they are those more discussed online and therefore more proposed by the algorithm, is essential to develop more effective communication strategies to stem disinformation during emergencies and more efficient crisis management models for the near future.

#### REFERENCES

- Affuso, O., Parini, E. G., & Santambrogio, A. (2020). *Gli italiani in quarantena. Quaderni da un «carcere» collettivo*. Morlacci Editore.
- Bayer, J. & Bárd, P. (2020). *Hate speech and hate crime in the EU and the evaluation of online content regulation approaches*. European Parliament. [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL\\_STU\(2020\)655135\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU(2020)655135_EN.pdf)
- Boccia Artieri, G. (2012). *Stati di connessione. Pubblici, cittadini e consumatori nella (Social) Network Society*. Franco Angeli.

- Borgonovo, F. (2020, March 28). *COVID-19: Telegram, autogestione spontanea della comunicazione*. itstime. <https://www.itstime.it/w/covid-19-telegram-autogestione-spontanea-della-comunicazione-by-federico-borgonovo/>
- Bucchi, M. & Saracino, B. (2020, November 3). Vaccino Covid, un italiano su cinque non lo vuole. I dati di Observa Science. *Corriere della Sera*. [https://corriereinnovazione.corriere.it/cards/vaccino-covid-terzo-italiani-non-vuole-dati-sondaggio-osservatorio-observa-science/percezione-pericolo-fonti-informative\\_principale.shtml](https://corriereinnovazione.corriere.it/cards/vaccino-covid-terzo-italiani-non-vuole-dati-sondaggio-osservatorio-observa-science/percezione-pericolo-fonti-informative_principale.shtml)
- Cinelli, M., Quattrociochi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., Zola, P., Zollo, F., & Scala, A. (2020). The COVID-19 social media infodemic. *Scientific Reports*, *10*, 16598. <https://doi.org/10.1038/s41598-020-73510-5>
- Cinelli, M., Etta, G., Avalle, M., Quattrociochi, A., Di Marco, N., Valensise, C., Galeazzi, A., & Quattrociochi, W. (2022). Conspiracy theories and social media platforms. *Current Opinion in Psychology*, *101407*. <https://doi.org/10.1016/j.copsyc.2022.101407>
- Collard, M. (2023). Hate in the time of the Covid-19 pandemic: dehumanisation as a side effect; re-humanisation as a remedy. *Crime, Law and Social Change*. <https://doi.org/10.1007/s10611-022-10073-8>
- Colombo, F. (2022). *Verità e democrazia. Sulle orme di Michel Foucault*. Mimesis.
- Fabbri, L., Melacarne, C., Romano, A. & Rullo, M. (2022). *Teorie del complotto, radicalizzazione e ricerca del significato*. Franco Angeli.
- Ferrazzoli, M. & Maga, G. (2021). *Pandemia e infodemia. Come il virus viaggia con l'informazione*, Zannichelli.
- Kelman, H. C., & Hamilton, V. L. (1989). *Crimes of obedience: Toward a social psychology of authority and responsibilities*. Yale University Press.
- Maffesoli, M. (1988). *Il tempo delle tribù: il declino dell'individualismo nelle società di massa*. Armando Editore.
- Manfredi, G. (2019). *Infodemia. I meccanismi complessi della comunicazione nelle emergenze*. Guaraldi Editore.
- Marcelli, S. & Valerii, M. (2021). *Diciassettesimo rapporto sulla comunicazione. I media dopo la pandemia*. Franco Angeli.
- Pierangelo, F. & Gorgoni, S. (2017). *Disintermediazione e nuovi media. Come cambia la comunicazione*. Armando Editore.
- Polesana, M., A. & Risi, E. (2023). *(S)comunicazione e pandemia. Ricategorizzazioni e contrapposizioni di un'emergenza infinita*. Mimesis.
- Quattrociochi, W. & Vicini, A. (2017). *Misinformation. Guida alla società dell'informazione e della credulità*. Franco Angeli.
- Ranieri, G. (2020, November 3). Milano, lo "sfregio" agli eroi del coronavirus: vandalizzato il murales per medici e infermieri. (s.d.). *MilanoToday*. <https://www.milanotoday.it/attualita/coronavirus/murales-medici-infermieri-vandalisti.html>

- Romer, D., & Jamieson, K. H. (2020). Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. *Social Science & Medicine*, 263, 113356. <https://doi.org/10.1016/j.socscimed.2020.113356>
- Stefanoni, F. (2020, November 12). Sondaggio Covid, cambia la paura degli italiani e il governo Conte perde fiducia. *Corriere della Sera*. [https://www.corriere.it/politica/20\\_novembre\\_12/covid-cambia-paura-italiani-governo-conte-perde-fiducia-e030d07e-24e3-11eb-9615-de24e09c8a4a.shtml](https://www.corriere.it/politica/20_novembre_12/covid-cambia-paura-italiani-governo-conte-perde-fiducia-e030d07e-24e3-11eb-9615-de24e09c8a4a.shtml)
- Vergani, M., Martinez Arranz, A., Scrivens, R., & Orellana, L. (2022). Hate Speech in a Telegram Conspiracy Channel During the First Year of the COVID-19 Pandemic. *Social Media + Society*, 8(4), 205630512211387. <https://doi.org/10.1177/20563051221138758>
- Viafora, G. (2020, December 1). Post negazionisti e insulti ai medici: Facebook non li cancella «Rispettano gli standard» Il test. *Corriere della Sera*. [https://www.corriere.it/cronache/20\\_dicembre\\_01/covid-negazionismo-insulti-medici-facebook-non-cancella-post-rispettano-nostri-standard-3b7d3c72-33be-11eb-be82-c9839d3e98fa.shtml](https://www.corriere.it/cronache/20_dicembre_01/covid-negazionismo-insulti-medici-facebook-non-cancella-post-rispettano-nostri-standard-3b7d3c72-33be-11eb-be82-c9839d3e98fa.shtml)

# COVID-19 Y CAMPAÑAS POLÍTICAS EN TWITTER: UN ANÁLISIS DE LOS CANDIDATOS EN LAS ELECCIONES LEGISLATIVAS 2021 EN ARGENTINA

EMMANUEL BORTHIRY

*Becario Doctoral. INHUS - CONICET - CEHis  
Universidad Nacional de Mar del Plata, Argentina*

**DOI:** 10.14679/2757

## **Summary:**

1. Introducción.
  2. Antecedentes de investigación.
  3. Aspectos metodológicos.
  4. La evolución del tópico COVID-19.
  5. Candidatos COVID-19.
  6. Tópicos y enfoques partidarios.
  7. Conclusiones.
- Referencias.

## 1. INTRODUCCIÓN

La emergencia de la pandemia del COVID-19 ha significado una serie de transformaciones en las sociedades contemporáneas. Desde su propagación global a partir del año 2020, el virus ha desencadenado una serie de efectos sociales, políticos y económicos que trascendieron fronteras y colocaron en jaque a los estados nacionales. El aumento exponencial de casos supuso presiones inéditas sobre los sistemas de salud de todos los países, por lo cual los gobiernos debieron adoptar medidas de confinamiento y distanciamiento social, que implicaron a su vez tensiones y daños en las economías. En este contexto, las redes sociales han desempeñado un rol destacado en lo que refiere a la circulación de información, la formación de opiniones y la producción de discursos. Además, estos espacios digitales se han convertido en foros de discusión, donde se entrelazan realidades y percepciones, y donde los discursos sobre la pandemia se amalgaman con reflexiones políticas. El flujo constante de información ha generado un escenario propicio para la proliferación de discursos polarizados en torno a la pandemia. Plataformas como Twitter han sido empleadas en todo el mundo tanto para la difusión de información veraz y constructiva, como para la propagación de desinformación, noticias falsas y teorías conspirativas (Gualda, 2022; Romer y Jamieson, 2020).

Este panorama cobra especial relevancia en el marco de países en donde se llevaron a cabo elecciones en este contexto. En Argentina, los debates y controversias en torno a la pandemia jugaron un rol decisivo en las elecciones legislativas del año 2021. Los candidatos políticos utilizaron las redes sociales como sitios para expresar sus posiciones y confrontar ideas vinculadas a la situación epidemiológica, las medidas de cuidado, la economía y la campaña de vacunación. El fenómeno del distanciamiento social y las restricciones de circulación han confinado gran parte de la vida cotidiana al mundo digital, posicionando a las redes sociales como un espacio clave para la vinculación entre políticos y ciudadanos.

En este marco, la propuesta de investigación consiste en analizar los *timelines* de Twitter de los candidatos a diputados y senadores en las elecciones legislativas del año 2021, considerando un período amplio desde el comienzo de la pandemia. El análisis de los *timelines* de Twitter de los candidatos puede ofrecer un panorama de cómo se abordaron los desafíos derivados de la pandemia y cómo se entrelazaron con estrategias políticas y electorales, en un escenario signado por incertidumbres y urgencias. Guían esta investigación preguntas como ¿Cuáles fueron los discursos desplegados por los candidatos sobre la pandemia? ¿Qué estrategias comunicacionales emplearon en Twitter? ¿Cuáles fueron los momentos de mayor centralidad del tópico COVID-19? ¿Qué diferencias existieron entre partidos políticos y entre las distintas provincias en Argentina? ¿Qué rol jugó como tema en las elecciones legislativas? ¿Cuáles fueron los candidatos más activos y con mayor nivel de interacciones? ¿Existieron cuestionamientos a las medidas de cuidado, adhesiones a teorías de la conspiración o difusión de información falsa?.

## 2. ANTECEDENTES DE INVESTIGACIÓN

Desde mediados de la década de 1990, ha surgido una creciente literatura académica que examina cómo Internet se ha convertido en una herramienta fundamental para la comunicación política y electoral. Las campañas de Howard Dean en 2004 y las dos campañas de Barack Obama en 2008 y 2012 marcaron hitos significativos en este debate (Hendricks y Denton, 2010). Inicialmente, este enfoque se centraba en la reflexión sobre el uso de la web para la acción política, pero posteriormente se expandió para incluir el análisis de las redes sociales (Kreiss, 2012). Como resultado, ha surgido un campo multidisciplinario de estudios que abarca disciplinas como las ciencias políticas, la comunicación, el marketing, la sociología, la historia y la lingüística, entre otras. Estos estudios han abordado una amplia gama de enfoques y metodologías, desde el análisis del discurso y las estrategias de aproximación de los candidatos, la identificación de noticias falsas y el uso de bots y trolls, así como también los

efectos de la polarización política (Aruguete, 2019; Calvo y Aruguete, 2020; Juarez y Ferreyra, 2021).

También se han llevado a cabo investigaciones específicamente sobre los timelines de políticos en Twitter. Un proyecto destacado es el de Alshaabi et al. (2021), quienes crearon Storywrangler, una herramienta para recopilar y analizar millones de tweets desde 2008 hasta 2021, que permite explorar aspectos lingüísticos, culturales, socioeconómicos y políticos en los “timelines” de Twitter. Boireau (2014) examinó la actividad de políticos belgas, utilizando técnicas de minería y análisis de texto para evaluar sus posturas políticas. En Italia, Bracciale y Martella (2017) se enfocaron en los estilos de comunicación política al analizar los timelines de líderes políticos italianos para identificar tipos de liderazgo. En España, se llevaron a cabo investigaciones como la de López-Rabadán, López-Meri y Doménech-Fabregat (2016), en donde analizaron la gestión de Twitter por parte de políticos en el período preelectoral de 2015, centrandó su atención en el uso de la fotografía. Slimovich (2016) investigó en Argentina la digitalización del discurso político en redes sociales, enfocándose en las cuentas de Twitter de los candidatos presidenciales durante las elecciones de 2011. Recalde, Baquerizo y Zunino (2019) estudiaron el uso de Twitter por parte de las principales líderes políticas femeninas durante la campaña de las elecciones legislativas en Argentina 2017, analizando los temas que promovieron, sus interacciones con los medios tradicionales y el tipo de diálogo que establecieron entre ellas, teniendo en cuenta variables de género y ubicación.

Tal como señala la bibliografía, la presencia de políticos argentinos en la plataforma Twitter comenzó a tomar relevancia a partir de las elecciones de 2011 (Castro Rojas, Pérez Rivera y Amatta, 2016; Slimovich, 2016). El uso de Twitter no solo se ha convertido en una forma de mostrar su actividad, sino que también ha dejado un rastro valioso de datos para la investigación académica. Los tweets publicados generan lo que se denomina un “timeline” o línea temporal, que es una secuencia cronológica de publicaciones que proporciona una visión ordenada de los mensajes compartidos. A través de su análisis, se puede acceder a información sobre las estrategias de comunicación de los políticos, la agenda que promueven, su estilo retórico y las interacciones con otros usuarios y actores políticos.

En el contexto de la pandemia de COVID-19, se han llevado a cabo diversas investigaciones que exploran tanto el papel de Twitter como medio de comunicación y plataforma de difusión de información relacionada con el virus, como las creencias y actitudes de la población ante la pandemia y las teorías de conspiración. Por caso, Shurafa et. al. (2020) analizaron cómo Twitter se utilizó para enmarcar la pandemia desde una perspectiva política en la campaña presidencial de 2020 en los Estados Unidos. Romer y Jamieson (2020) investigaron cómo las creencias en teorías de la conspiración afectaron la disposición

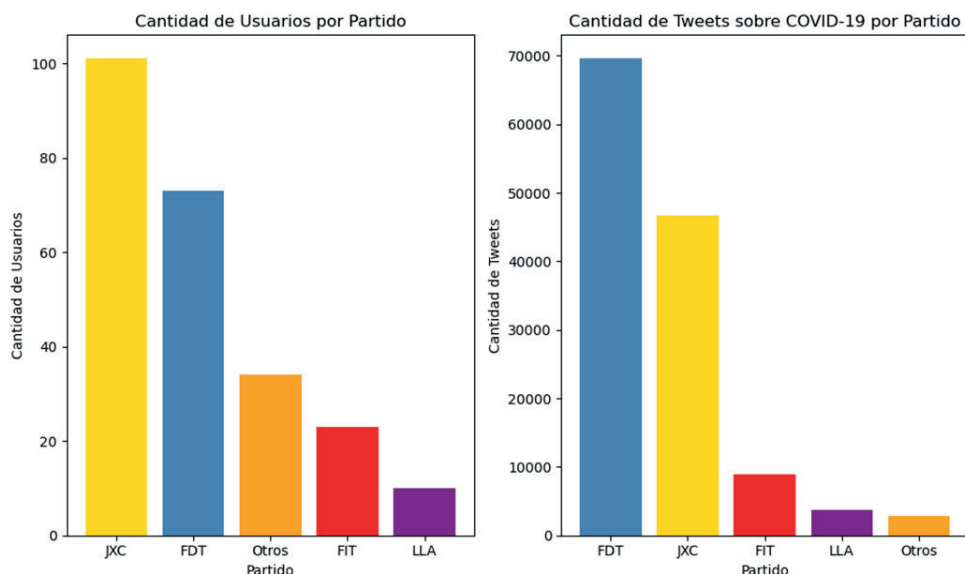
del público para tomar medidas preventivas y vacunarse contra el COVID-19. Estrella Gualda (2022) examinó el negacionismo y las teorías de conspiración en Twitter relacionadas con el virus, utilizando la manifestación de Madrid en agosto de 2020 como caso de estudio. En Argentina, Calvo y Aruguete (2020) investigaron cómo la polarización política y los encuadres mediáticos influyeron en la percepción del riesgo y las respuestas políticas durante la pandemia.

### 3. ASPECTOS METODOLÓGICOS

Para alcanzar los objetivos, esta investigación se vale de una metodología basada en el procesamiento y análisis de datos provenientes de 261 timelines de Twitter de candidatos en las elecciones legislativas argentinas de 2021. Los datos fueron obtenidos del repositorio del proyecto "Políticos en Twitter" (Políticos en Twitter, 2023), el cual proporciona acceso a un conjunto de datos de candidatos, políticos y otros actores relevantes del escenario político argentino. Para el procesamiento y la generación de gráficos se utilizó el lenguaje de programación Python. El corpus de datos comprende un total de 131.837 tweets y retweets relacionados con la pandemia de COVID-19. Para confeccionar este corpus se elaboró un diccionario de palabras claves relacionadas con la pandemia (aislamiento, barbijo/s, cuarentena, covid, infectadura, olivos, pandemia, plandemia, vacuna/s, vacunatorio), a partir del cual filtramos las publicaciones de los candidatos que empleaban esas palabras claves.

Este conjunto de datos abarca un período que se extiende desde principios del 2020 –año en que comienza la pandemia– hasta finales del 2021 –cuando se celebran las elecciones legislativas en Argentina. La selección de un período amplio busca, por un lado, poder observar las variaciones temporales de la temática COVID-19 (en términos de frecuencia de publicaciones) y, por otro, considerar que la campaña política real no se circunscribe únicamente a los tiempos formales y reglamentarios, sino que se produce de manera extendida o permanente (Ornstein y Mann, 2000). La profesionalización de la comunicación política, influenciada por técnicas de marketing, ha llevado la dinámica comunicativa de las campañas electorales a periodos no electorales, abarcando tanto a quienes ejercen el poder como a quienes forman parte de la oposición (López-Rabadán, López-Meri, Doménech, 2016).

Figura 1. Composición de la base por partidos y cantidad de tweets y retweets



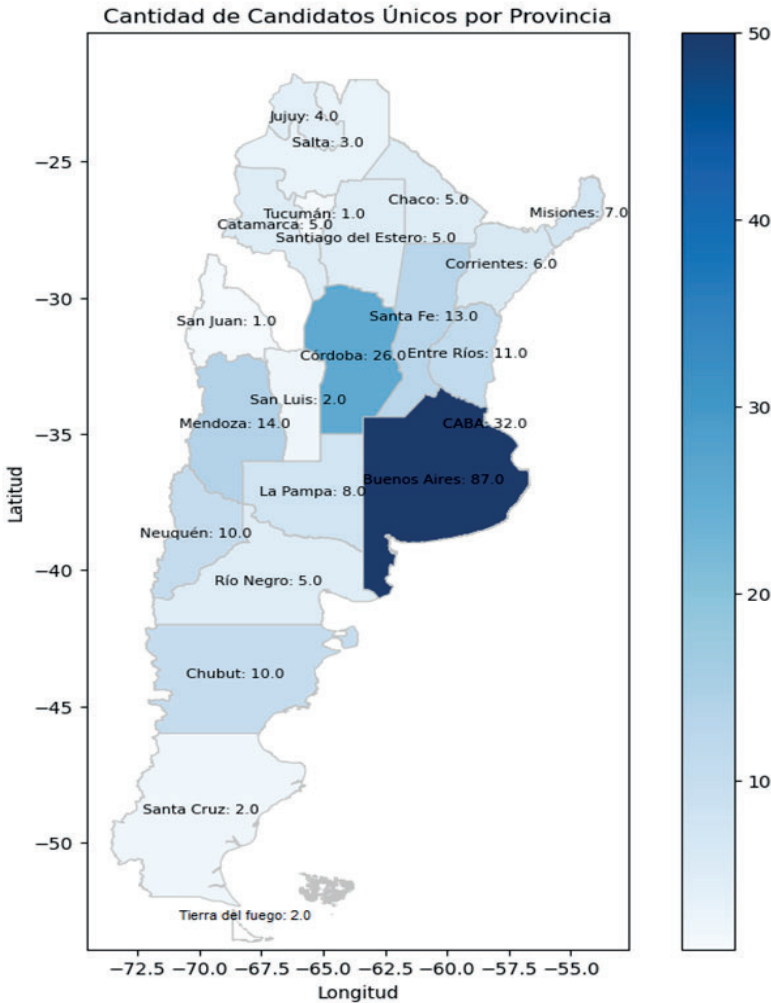
Fuente: Elaboración propia en base a políticos en Twitter.

En la figura 1 se pueden observar dos elementos centrales acerca de la composición de la base: la cantidad de usuarios y la de publicaciones por partido. Al tratarse de una base descargada desde un repositorio externo, se vuelve necesario examinar su composición para establecer métodos adecuados para la construcción de datos significativos. La base cuenta con una mayoría de candidatos de Juntos por el Cambio (JXC, 99), la principal coalición opositora. Por su parte, el gobierno del Frente de Todos (FDT) cuenta con 73 candidatos<sup>1</sup>. Dentro de “Otros” se agruparon candidatos de partidos provinciales sin relación directa con las coaliciones nacionales, incluyendo un total de 36 usuarios. En cuanto al Frente de Izquierda y de Trabajadores-Unidad (FIT), cuenta con 23 candidatos. Por último, La Libertad Avanza (LLA, 10) agrupa partidos que, bajo denominaciones heterogéneas, se presentaron a las elecciones como una alternativa de derecha liberal-libertaria, siguiendo la figura central de Javier Milei, candidato a diputado nacional por la Ciudad Autónoma de Buenos Aires (CABA).

<sup>1</sup> Para contextualizar, luego de 12 años de gobiernos peronistas/kirchneristas (2003-2015), en 2015 la coalición política Cambiemos (ahora Juntos por el Cambio) ganó las elecciones que consagraron a Mauricio Macri como presidente de la nación. 4 años más tarde, en unas nuevas elecciones, el peronismo (Frente de Todos) vence en las urnas a Juntos por el Cambio y Alberto Fernández es electo como el nuevo presidente que, a los pocos meses de asumir, debe administrar la pandemia en Argentina.

Si consideramos la cantidad de tweets publicados por partido, son los candidatos del oficialismo quienes más publicaciones realizaron, a pesar de tener una menor representación numérica en la base, con 69.553 tweets contra 46.236 de JXC. Por otro lado, si bien los candidatos de partidos provinciales se ubicaban en tercera posición en cuanto a número de usuarios, realizaron menos publicaciones (3.409) que el FIT (8.927) y que LLA (3.712).

Figura 2. Composición de la base, candidatos por provincia



Fuente: Elaboración propia en base a políticos en Twitter.

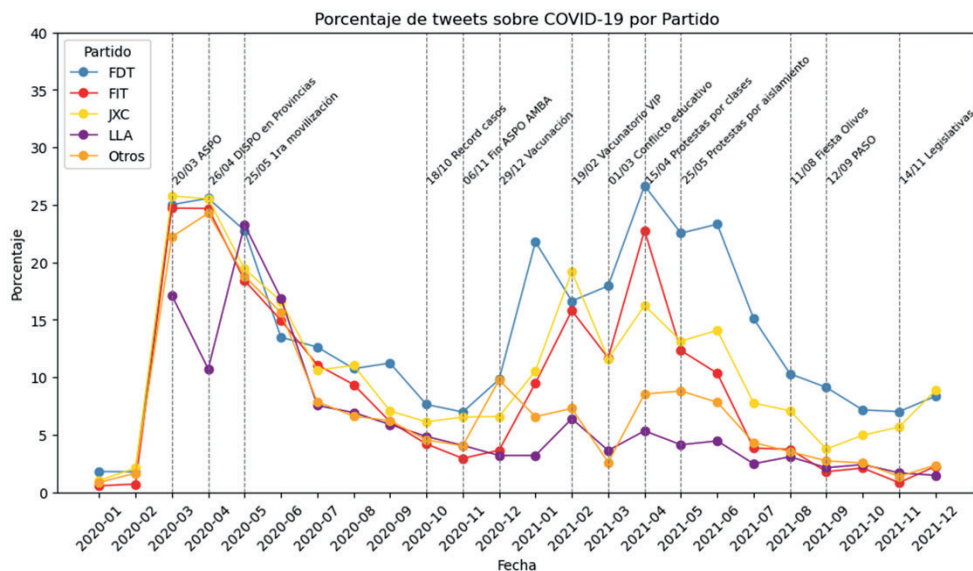
En lo que respecta a la distribución geográfica, la base está compuesta mayoritariamente por candidatos de la Provincia de Buenos Aires (PBA), lo cual se

condice con ser la provincia con mayor representación en la cámara de diputados. En segundo lugar se ubica la Ciudad Autónoma de Buenos Aires, seguido por Córdoba, Mendoza y Santa Fé. La base se compone por más candidatos/usuarios en las provincias más populosas y centrales del país. Por otro lado, las provincias situadas en el norte y sur del país tienen una menor representación en esta base, que puede explicarse tanto por una menor cantidad de candidatos a diputados, como por una menor centralidad mediática y uso de redes sociales.

#### 4. LA EVOLUCIÓN DEL TÓPICO COVID-19

A los fines de detectar momentos de mayor o menor centralidad de la pandemia como tópico en el debate público, se analizará el porcentaje de publicaciones de tweets acerca de COVID-19 sobre el total de publicaciones de los partidos<sup>2</sup>. Este enfoque de análisis ofrece la posibilidad de ver la evolución del tópico en relación a la curva sanitaria y los acontecimientos relevantes que transcurrieron en el período considerado. Los datos muestran que fue oscilando a lo largo del tiempo, sin embargo, hubo una presencia constante del tema.

Figura 3. Frecuencia de publicaciones porcentuales



Fuente: Elaboración propia en base a políticos en Twitter.

<sup>2</sup> Debido a las diferencias en valores absolutos de la base, y a los fines de establecer comparaciones entre candidatos y partidos, se utilizaron datos con valores porcentuales sobre el total de publicaciones, como puede observarse en el gráfico de la figura 3.

Es en el mes de marzo, con los primeros casos de COVID-19 y la oficialización del Aislamiento Social Preventivo y Obligatorio (ASPO), que el tema se instala fuertemente, ocupando cerca del 25% de las publicaciones de los partidos, a excepción de La Libertad Avanza. Se mantuvieron valores similares durante abril, en donde se destaca un pico máximo de los partidos provinciales en sintonía con la flexibilización de la cuarentena en aglomerados urbanos con menos de 500.000 habitantes, dando paso a la etapa de Distanciamiento Social, Preventivo y Obligatorio (DISPO). Si bien en mayo los valores continúan elevados, comienza una tendencia descendente que no se revertirá hasta finales del año 2020. Se destaca igualmente el pico máximo de La Libertad Avanza, que si bien coincide con el inicio de las movilizaciones anti-cuarentena, las cuales comenzaron a producirse en fechas patrias como el 25 de Mayo, no continuó en los meses siguientes.

La tendencia descendente continuó, situándose por debajo del 15% a partir de julio, y con valores cercanos al 5% para algunos partidos en meses como octubre y noviembre, en donde se producen respectivamente el pico de la primera ola y la flexibilización de las medidas de cuidado en el Área Metropolitana de Buenos Aires (AMBA). Sin embargo, en el mes de diciembre la tendencia se revierte, principalmente a partir de la iniciativa del gobierno por el arribo de las primeras vacunas. El 29 de diciembre comenzó formalmente la campaña de vacunación, lo que reactivó fuertemente el tema. JXC como principal oposición nacional se topó con una disyuntiva que enfrentó actores internos, entre quienes cuestionaron la campaña de vacunación debido a aspectos como por ejemplo el origen de las primeras vacunas que arribaron al país (provenientes de Rusia y China) y, por otro lado, quienes fomentaban la vacunación en las provincias o distritos en donde gobernaban (en particular CABA). En este sentido, la vacunación pasó a ser el eje más relevante.

La iniciativa del gobierno durante diciembre y enero se vio mermada durante febrero, cuando el día 18 se origina un escándalo político denominado como el “Vacunatorio VIP”, en donde se devela que una serie de personas (políticos, periodistas, entre otros) accedieron de manera privilegiada a vacunas. El escándalo culminó con la salida del Ministro de Salud Ginés González García y su reemplazo por la viceministra, Carla Vizzotti. Esta situación fue aprovechada en términos de estrategia comunicacional principalmente por Juntos por el Cambio, que se refleja en el gráfico con un pico de actividad.

A partir de entonces se produce una reactivación del tópico ligado a la vacunación que se superpone con el retorno a clases presenciales en las escuelas. Se inicia desde finales de febrero y principios de marzo un conflicto educativo en el cual la oposición exigía la vuelta a la presencialidad bajo la consigna “abran las escuelas”, mientras que el gobierno mostraba preocupación por el inicio de una nueva ola de contagios y el insuficiente avance en la campaña de vacunación.

En junio de 2021, pese a transitar el segundo pico de fallecimientos, se produce un descenso en las publicaciones de los candidatos, situándose de todos modos el oficialismo como principal promotor del tema debido a la continuidad de la campaña de vacunación y recurriendo al avance progresivo de la misma como uno de los ejes de campaña. No obstante, al momento de aproximarse a las elecciones Primarias, Abiertas, Simultáneas y Obligatorias (PASO), el tópico se encuentra en niveles por debajo del inicio de la pandemia o de comienzos del 2021. Incluso se produce otro escándalo político en el mes de agosto, conocido como “La fiesta de Olivos”, en donde se develó que durante la noche del 14 de julio de 2020, en la fase más severa de la cuarentena, se celebró el cumpleaños de la Primera Dama Fabiola Yáñez en la Quinta Presidencial de Olivos y de la cual formó parte el Presidente Alberto Fernández. A pesar de ello, los datos de la figura 3 reflejan que este escándalo tuvo menor nivel de repercusión que el Vacunatorio VIP. Este dato, sumado a la baja centralidad del tópico durante las elecciones, denotan un agotamiento y/o un cambio en las estrategias de comunicación de los partidos. Finalmente, en las elecciones legislativas 2021 la oposición de Juntos por el Cambio venció al Frente de Todos a nivel nacional con un 42.75% de los votos frente a un 34.56%.

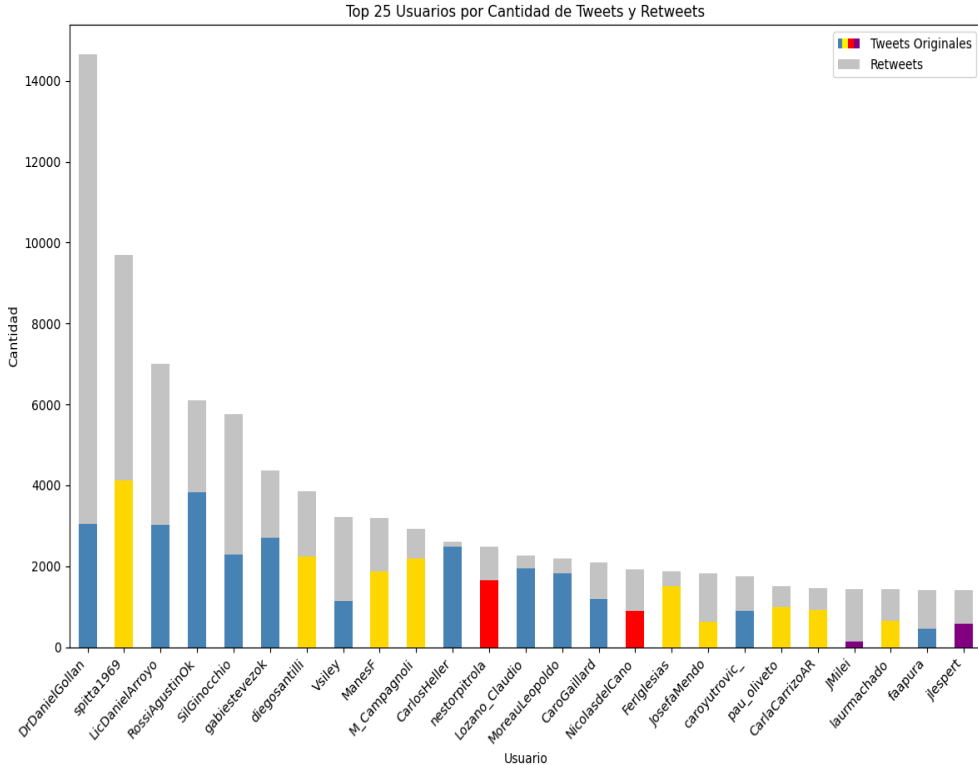
## 5. CANDIDATOS COVID-19

Hubo candidatos que mostraron más interés y tuvieron una mayor actividad en los tópicos relacionados al COVID-19. Estos candidatos se destacaron no solo por su presencia constante en la red social, sino también por su capacidad para participar activamente en el diálogo político a través de la producción y difusión de discursos. Resulta de interés relevar, por un lado, quienes fueron los usuarios más activos por partido y, por otro, dar cuenta de los usos que realizaron de la plataforma. Además, a partir de estos resultados, es posible observar luego con mayor detalle los discursos desplegados por los candidatos y medir su capacidad de intervención en el debate público.

La figura 4 revela cuáles fueron los 25 candidatos que más tweets y retweets publicaron con respecto a la pandemia. Se destaca que la mayor cantidad de usuarios corresponden tanto al FDT (12) como a JXC (9). Por su parte, tanto el FIT como LLA cuentan con 2 candidatos en el ranking. Resulta distintivo como ningún candidato del conglomerado de partidos provinciales forma parte, lo que da cuenta de que, al menos en términos de cantidad de publicaciones, la conversación estuvo acaparada por las dos principales coaliciones. En términos de representación geográfica, sólo 7 de los 25 se candidatos presentaron por provincias que no fueran Buenos Aires o CABA (Agustín Rossi por Santa Fe, Carolina Gaillard por Entre Ríos, Carolina Yutrovic por Tierra del Fuego, An-

tártida e Islas del Atlántico Sur, Laura Machado y Gabriela Estevez por Córdoba, Fabián Puratich por Chubut y Silvina Ginocchio por Catamarca).

Figura 4. Publicaciones originales y retweets por candidato y partido



Fuente: Elaboración propia en base a políticos en Twitter.

Tal como se observaba en la figura 1, los candidatos del oficialismo fueron los más activos en términos de publicaciones totales. Las estrategias empleadas fueron divergentes, habiendo casos con un amplio porcentaje de retweets (como Daniel Gollán o Carlos Arroyo) y otros que produjeron la mayor parte de sus publicaciones (Carlos Heller, Claudio Lozano o Leopoldo Moreau). Con notoria diferencia, el candidato con más publicaciones fue Daniel Gollán, quien fuera el Ministro de Salud de la Provincia de Buenos Aires durante la pandemia<sup>3</sup>. Tanto por su participación activa en redes durante la pandemia, como por su narrativa como candidato a diputado vinculada a la misma, es uno de los casos más destacado de “candidato COVID-19”, ya que basó su campaña

<sup>3</sup> Fabián Puratich se presentó con un perfil parecido al de Gollán (@faapura), luego de ser el Ministro de Salud de Chubut durante la pandemia, aunque su participación en redes tuvo un volumen menor.

en el accionar como ministro, haciendo énfasis en la campaña de vacunación en la provincia más populosa del país. Otro candidato con un perfil similar fue Daniel Arroyo, quien se desempeñaba como Ministro de Desarrollo Social de la Nación y se presentó como candidato a diputado en la PBA con un discurso ligado a las acciones de gobierno llevadas a cabo para paliar los efectos económicos adversos. Como dato a remarcar, se destaca en el ranking oficialista la ausencia de Victoria Tolosa Paz, quien fuera cabeza de lista por el Frente de Todos en la provincia.

En cuanto a JXC, de los 9 que forman parte del gráfico, 6 de ellos realizaron más publicaciones originales que retweets. La candidata más activa fue Sandra Pitta, quien hizo un fuerte uso del retweet, pero a su vez fue también quien más publicaciones originales produjo. Desde el año 2019 tuvo una amplia repercusión mediática al presentarse como una investigadora del CONICET, aunque crítica del organismo, y durante el transcurso de la pandemia tuvo diferencias con la gestión sanitaria<sup>4</sup>. Su discurso estuvo centrado, por un lado, en el cuestionamiento a las vacunas Sputnik-V de Rusia y Sinopharm de China aplicadas en el país y la exigencia de importar vacunas de laboratorios estadounidenses, y por otro, en la difusión del escándalo del vacunatorio VIP. Su perfil de investigadora experta en biotecnología la posicionó como una voz autorizada para cuestionar el accionar del gobierno. Por otra parte, se destacaron también candidatos como Diego Santilli y Facundo Manes, quienes disputaron una interna en las elecciones primarias en la PBA. Otros candidatos que tuvieron un amplio volumen de publicaciones fueron Marcela Campagnoli, Fernando Iglesias, Josefina Mendoza, Paula Oliveto, Carla Carrizo y Laura Machado.

Por parte del FIT, los candidatos más activos fueron Nestor Pitrola y Nicolás del Caño. Las publicaciones de Pitrola abarcaron distintos tópicos, como la situación laboral afectada por la pandemia, denuncias en torno a manifestaciones laborales, críticas hacia decisiones gubernamentales como el cierre del Parlamento y un cuestionamiento a la vacunación de funcionarios antes que del propio personal de salud. De modo semejante, Del Caño se pronunció por cuestiones como la falta de vacunas para docentes, la situación general de los trabajadores durante la pandemia, el vacunatorio VIP y la necesidad de liberar las patentes de las vacunas.

Por último, los candidatos de LLA Javier Milei y José Luis Espert<sup>5</sup> efectuaron un volumen bajo de publicaciones. Milei realizó pocas publicaciones originales para referir al COVID-19, la mayoría corresponden a retweets a otros usuarios. De todos modos, en ellos expresó una serie de críticas hacia la gestión gubernamental.

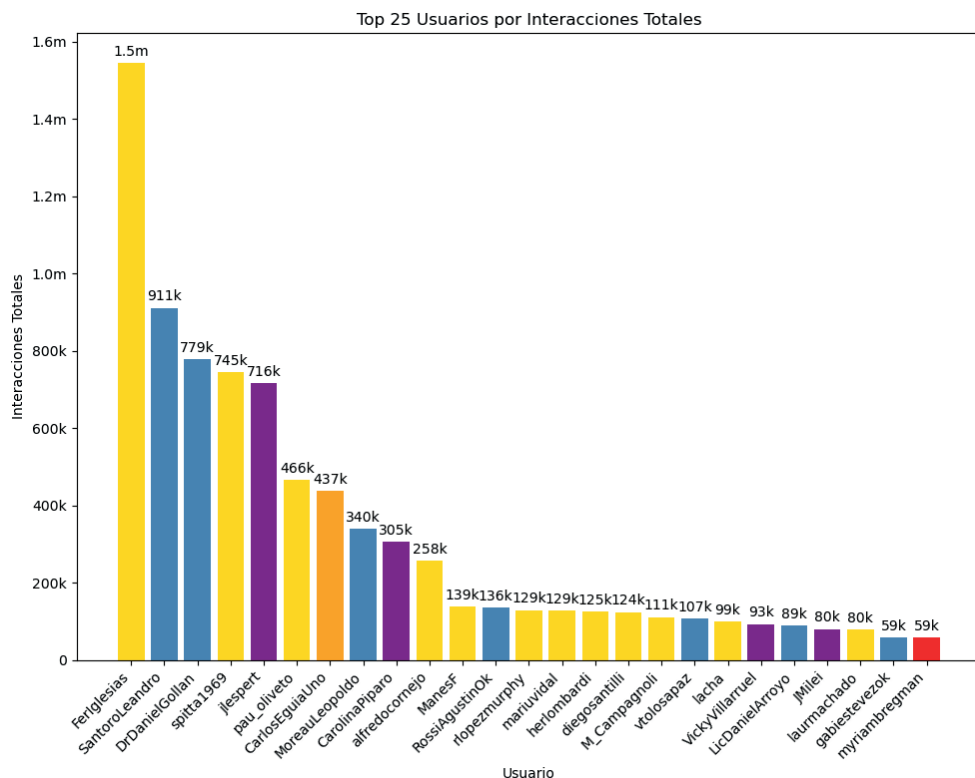
---

<sup>4</sup> En 2019 Infobae daba a conocer detalles del perfil de Sandra Pitta en una nota titulada "Quién es Sandra Pitta, la científica que se cruzó con Alberto Fernández".

<sup>5</sup> Vale aclarar que, si bien hicieron campaña juntos, José Luis Espert se presentó a las elecciones como candidato a diputado por la Provincia de Buenos Aires por Avanza Libertad.

mental de la pandemia, centradas en la tensión economía/salud, al argumentar que las medidas adoptadas eran perjudiciales para la economía y tampoco eran efectivas contra la pandemia.

Figura 5. Candidatos con mayor nivel de interacciones recibidas (likes + retweets)



Fuente: Elaboración propia en base a políticos en Twitter.

Ahora bien, aunque Gollan y Pitta ocupan un lugar relevante en ambos listados, un mayor número de publicaciones no implica necesariamente que esos tweets tengan impacto o difusión. Como puede observarse en la figura 5, candidatos como Fernando Iglesias (JXC), Leandro Santoro (FDT) o José Luis Espert (LLA), que no se ubicaban en las primeras posiciones del ranking de cantidad de publicaciones, fueron algunos de los que más *engagement*<sup>6</sup> generaron. El gráfico proporciona una visión de las interacciones, es decir, de la suma de “likes” y “retweets” recibidos. Analizarlas puede arrojar luz sobre la popularidad,

<sup>6</sup> En comunicación política en redes sociales el engagement remite al nivel de interacción de ciudadanos en relación a publicaciones de actores políticos influyentes en el debate público (Vaccari, Chadwick y O’Loughlin, 2015).

el impacto y el alcance de los diferentes candidatos en la plataforma y ayudar en la detección de tweets que tuvieron amplia circulación.

El candidato Fernando Iglesias (JXC) fue quien más interacciones totales recibió, con un total 1.5 millones. A través de su cuenta esbozó críticas al gobierno en distintos ejes. En primer lugar, cuestionó las medidas adoptadas por el gobierno. En una de sus publicaciones afirmaba: «Nos encerraron y fundieron a medio país por razones políticas, y no sanitarias ni epidemiológicas. Ahora decretan el fin de la cuarentena eterna por las mismas razones. #ElPeorGobierno-DeLaHistoria» (Fernando A. Iglesias [@FerIglesias], 2021). Por otra parte, con respecto a la tensión economía/salud, sostuvo que el gobierno no se encargó de ninguna de las dos cosas, en junio de 2021 afirmaba que en Argentina había «88.247 muertos. Solo 7% de la población vacunada. Más del doble de caída del PBI que Brasil y 50% más que Chile y Uruguay. Récord de cuarentena y escuelas cerradas. Robaron vacunas» (Fernando A. Iglesias [@FerIglesias], 2021). Con respecto a la vacunación enfatizó en el robo de vacunas, la procedencia de las mismas y la innecesariedad de implementar un pase sanitario, decía «Voy a inocularme cualquiera de las vacunas que aprueben los Estados Unidos o la Unión Europea. No me pongo como ejemplo ni soy antivacuna, todo lo contrario: creo en la ciencia y en la transparencia de los asuntos públicos. Fin del comunicado. Que cada uno haga lo que quiera» (Fernando A. Iglesias [@FerIglesias], 2020).

Por su parte, Leandro Santoro (FDT), principal candidato oficialista en CABA, utilizó su cuenta para efectuar críticas al jefe de gobierno de la ciudad, Horacio Rodríguez Larreta (JXC), las vacunas, la desinformación y la tensión economía/salud. Con respecto a la vacunación en la ciudad afirmaba: «Larreta arma un vacunatorio en el Luna Park –tan mal organizado– que se corre el riesgo real de que salgan todos infectados!» (Leandro Santoro [@SantoroLeandro], 2021). También deslizó acusaciones de corrupción al gobierno de la ciudad por la compra de barbijos a sobreprecio «Hola @horaciorlarreta! ¿Viste q compraron barbijos a \$3000 c/u? Bueno, es un escándalo!» (Leandro Santoro [@SantoroLeandro], 2020). Por otro lado, cuestionó la desinformación propiciada por la oposición con respecto a las vacunas: «Que la sputnik es veneno. Que las vacunas chinas q adquirimos son malas. Que para comprar Pfizer pidieron coimas. Que rechazamos vacunas de Covax por capricho. Todos los días arman una fake para meterle odio a la gente pero les dura un ratito» (Leandro Santoro [@SantoroLeandro], 2021). También vinculó las marchas anti-cuarentena y las quemadas de barbijo con la dirigente opositora Patricia Bullrich afirmando que «En el peor momento de la pandemia “los médicos x la verdad” convocan marchas anticuarentena –sin usar barbijos– levantando la consigna de Patricia Bullrich» (Leandro Santoro [@SantoroLeandro], 2021). En relación a la situación económica, destacaba que «Quieren levantar la cuarentena para “Salvar la Economía” pero la economía no está en crisis por la cuarentena, está en crisis

por el impacto social q causa la epidemia en el mundo» (Leandro Santoro [@SantoroLeandro], 2020).

En cuanto a LLA, quien más interacciones recibió fue José Luis Espert, candidato por PBA. Tanto Espert como su compañera de fórmula en Avanza Libertad, Carolina Piparo, tuvieron mayores niveles de interacción que el líder liberal-libertario Javier Milei. Desde comienzos de la pandemia en 2020 fue crítico de la cuarentena, ya en junio sostenía lo siguiente «Rumbo a los 100 días de cuarentena, la gente fundiéndose por no trabajar y sin embargo no paran de acelerarse los contagios de coronavirus. El fracaso del #YoMeQuedoEnCasa es total» (José Luis Espert [@jlespert], 2020). La denominación más empleada para referir a la extensión y el nivel de restricciones de las medidas fue “cuarentena cavernícola”, como contrapunto a la “cuarentena inteligente” que promovía el gobierno. En el marco de las movilizaciones anti-cuarentena del 2020, el gobierno había calificado a las mismas como una acción en contra de la salud de los argentinos, a lo cual Espert responde en Twitter «difícil que sea peor que los casi 5 meses que hace que estamos encerrados en una cuarentena cavernícola» (José Luis Espert [@jlespert], 2020). Durante el arribo de las primeras vacunas aseguraba que «ANMAT<sup>7</sup> todavía no autorizó la vacuna rusa que fuimos a buscar en vuelo épico con la quebrada Aerolíneas Argentinas y al mismo tiempo ANMAT ya aprobó la vacuna Pfizer que todavía no tenemos» (José Luis Espert [@jlespert], 2020). En reiteradas publicaciones solicitó un juicio político al presidente Alberto Fernández por los hechos del vacunatorio VIP, la relación con el gobierno ruso para la adquisición de vacunas y la gestión de la cuarentena. Señalaba que «Entre el vacunatorio VIP y la carta al gobierno ruso de la Nicolini, el presidente @alferdez está para el juicio político. Gobierno de irresponsables, casi...criminales» (José Luis Espert [@jlespert], 2021). Además, sostuvo que el gobierno adquirió las vacunas de los laboratorios estadounidenses Pfizer y Moderna y liberó las restricciones sanitarias luego de las elecciones primarias, afirmando que «Al final la política sanitaria del Gobierno queda en evidencia que fue un relato. Miles de PYMES fundidas por la cuarentena cavernícola, la mitad de la población pobre, y después de una mala elección abren todo? Brutos y perversos. Nos toman de boludos». (José Luis Espert [@jlespert], 2021).

## 6. TÓPICOS Y ENFOQUES PARTIDARIOS

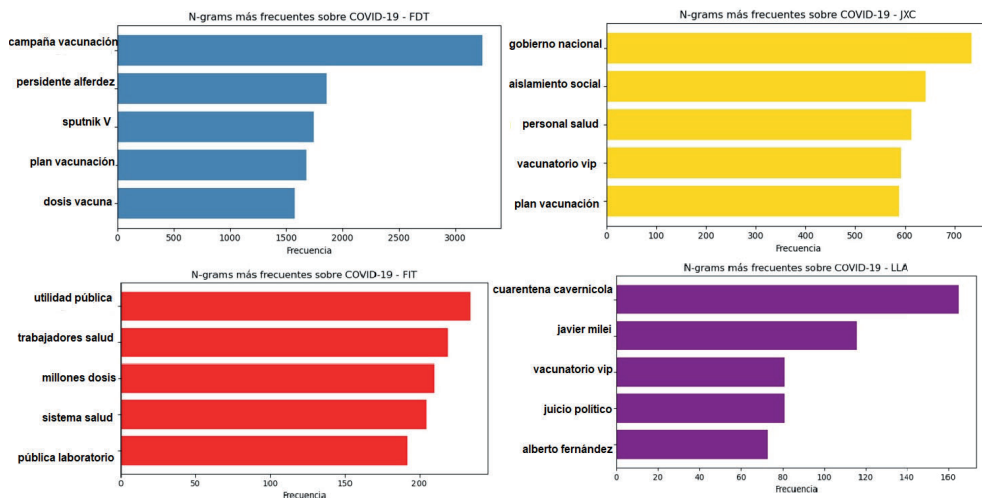
Mediante la visualización de los datos agregados por partido, expresados en *n-grams*<sup>8</sup> de 2 palabras (figura 6), es posible identificar las combinaciones

<sup>7</sup> Administración Nacional de Medicamentos, Alimentos y Tecnología Médica. Organismo descentralizado dependiente del Ministerio de Salud.

<sup>8</sup> Los *n-grams* son una técnica de procesamiento de lenguaje natural (NLP) utilizada para analizar la secuencia de palabras en un texto. Los *n-grams* son grupos consecutivos de palabras o

de palabras más frecuentes utilizadas para referir a la pandemia. Este tipo de análisis puede ayudar en la identificación de tópicos y en la comprensión de las estrategias adoptadas por los candidatos y los partidos.

Figura 6. N-grams más frecuentes por partido



Fuente: Elaboración propia en base a políticos en Twitter.

Los candidatos del FDT hicieron énfasis principalmente en la campaña de vacunación. 4 de los 5 n-grams más utilizados se corresponden con ese tópico. El bigrama más empleado –y con amplia diferencia con respecto al segundo– fue “campana vacunacion”, con un total de 3.423 menciones. Luego se encuentra “presidente alferdez” (1.857), dato que destaca la centralidad de la figura de Alberto Fernández como autoridad a cargo de la gestión de la pandemia por parte del gobierno. “Sputnik V” se menciona en 1.741 ocasiones, en referencia a una de las primeras vacunas en arribar al país y una de las más aplicadas. En cuarto lugar está el bigrama “plan vacunacion” (1.677), que podría interpretarse como un sinónimo de “campana vacunacion”, mientras que por último se encuentra “dosis vacuna” (1.572), utilizado para referir tanto al arribo de dosis como al seguimiento de su aplicación.

Por parte de JXC, el bigrama más destacado fue “gobierno nacional” (734), destinado a desplegar tanto críticas como reclamos vinculados a las medidas adoptadas por el gobierno. En segundo lugar, “aislamiento social” (612) fue empleado de manera ambivalente dependiendo de las etapas de la pandemia

---

caracteres que se extraen de un texto, donde “n” representa el número de elementos en el grupo. Se utilizaron *stopwords* en español de la librería nltk y se omitieron los términos “coronavirus”, “covid” y “pandemia” por su redundancia.

y las localidades en donde Juntos por el Cambio gobernaba, teniendo un significado positivo mientras las medidas de cuidado eran compartidas con el gobierno nacional y uno negativo luego de la prolongación de la cuarentena y del conflicto entre Horacio Rodríguez Larreta y Alberto Fernández por la quita de fondos de coparticipación. Por su parte, “personal salud” (600) tuvo la finalidad de, por un lado, reconocer las labores del personal (es destacable el uso de la palabra “personal” en lugar de “trabajadores”) y, por otro, de cuestionar las condiciones laborales y de los centros de atención de la salud durante la pandemia, responsabilizando al gobierno nacional. En cuarto término, “vacunatorio VIP” fue mencionado 592 veces por los candidatos de JXC, en alusión al escándalo mediático, enfatizando en la ligazón entre el peronismo y las acusaciones de corrupción.

En cuanto al FIT, los candidatos hicieron énfasis tanto en la labor y la situación de los trabajadores de la salud como en la liberación de patentes de las vacunas. El bigrama más utilizado fue “utilidad pública” (234) para exigir que el laboratorio del empresario Hugo Sigman del laboratorio mAbxience, encargado de la fabricación del principio activo de la vacuna AstraZeneca para América Latina, fuera declarado de utilidad pública para acelerar la fabricación. “Trabajadores salud” fue empleado 219 veces para expresar críticas relacionadas a las condiciones laborales, la protección y el apoyo del gobierno a los trabajadores durante la pandemia. En cuanto a “millones dosis” (210), estuvo ligado a la gestión, producción y distribución de vacunas, argumentando que Argentina estaba exportando a otros países cuando aún la población local no estaba vacunada, además de señalar que se debía priorizar la salud pública por sobre los intereses de los laboratorios privados. El cuarto bigrama fue “sistema salud” (205), en referencia a la necesidad de realizar transformaciones en el sistema, principalmente exigieron la “centralización del sistema de salud”, sobre la base de que la pandemia requería un enfoque unificado y coordinado en lugar de la dependencia de múltiples coberturas de salud. Por último, “pública laboratorio” (192) tuvo un sentido similar a “utilidad pública”, en tanto que señalaban la necesidad de la intervención del estado en la producción de vacunas.

En relación a LLA, los bigramas más frecuentes estuvieron vinculados con el cuestionamiento a la duración de la cuarentena, el señalamiento a Alberto Fernández como principal autoridad responsable, el escándalo del vacunatorio VIP y la centralidad de la figura de Javier Milei como verdadera oposición frente al gobierno nacional (FDT) y la jefatura de gobierno de la Ciudad Autónoma de Buenos Aires (JXC). Tal como se manifestaba en los tweets de Espert, el bigrama “cuarentena cavernícola” fue con 165 menciones el más empleado para criticar las medidas de cuidado de la salud adoptadas por las autoridades. La denominación “cavernícola” fue utilizada en las publicaciones para referir a que las medidas adoptadas restringieron las libertades individuales (tanto en lo

estricta que fue la cuarentena como en su prolongación), dañaron gravemente la economía, eran propias de países cerrados y respondían a los criterios de la Organización Mundial de la Salud (OMS), la cual fue tildada de “comunista”. En segundo lugar, el bigrama “Javier Milei” (116) da cuenta de la figura de autoridad del candidato como un opositor acérrimo a las dos coaliciones mayoritarias. Entre sus menciones se destacan, además de publicaciones sobre la situación económica, tweets de cuestionamientos a la aplicación de la vacuna China Sinopharm, debido a que no contaban con pruebas de fase 3, especialmente en niños. Si bien Milei se manifestó como “pro-vacunas”, sostuvo que las vacunas aún no estaban del todo probadas y que se vacunó solamente para poder realizar viajes al exterior. Por otra parte, “vacunatorio VIP” fue hallado en 81 tweets, cuestionando la vacunación de privilegio e intentando capitalizar políticamente las críticas. Por último, “juicio político” (81) y “Alberto Fernández” (73) son bigramas relacionados, en tanto que amalgaman tweets en donde se solicita efectuar un juicio político al presidente por la gestión de la pandemia, la cuarentena, los daños económicos y las relaciones internacionales establecidas para la adquisición de vacunas.

## 7. CONCLUSIONES

El análisis realizado permite arrojar luz sobre cómo los candidatos y los partidos abordaron la pandemia de COVID-19 en la plataforma Twitter en el marco de las elecciones legislativas. La pandemia desempeñó un papel crucial, ya que marcó un contexto en el cual la gestión de la crisis sanitaria y las estrategias para contenerla se convirtieron en temas centrales para los votantes. Existieron diferencias significativas entre los partidos y las provincias en relación con la pandemia. Los candidatos del FDT y JXC fueron los más activos en términos de publicaciones relacionadas con la pandemia. En cuanto a la representación geográfica, la mayoría de los candidatos eran de PBA y CABA, lo que sugiere que la conversación estuvo acaparada por las dos principales coaliciones.

Se observó que el tema pandemia mantuvo su relevancia en el discurso político argentino a lo largo del tiempo, aunque su importancia varió en diferentes momentos. Desde su inicio en marzo de 2020 hasta las elecciones en 2021, la pandemia fue un tema recurrente, con picos en el inicio y luego a partir del arribo de las vacunas, el escándalo del vacunatorio VIP y la vuelta a clases presenciales.

Los candidatos políticos adoptaron enfoques diversos en relación con la pandemia. Entre los candidatos del oficialismo, figuras como Gollán y Arroyo basaron sus campañas en su experiencia como funcionarios y enfatizaron en la campaña de vacunación. Los candidatos de JXC, como Pitta, Santilli y Manes, adoptaron discursos críticos hacia la gestión gubernamental de la pandemia.

El análisis de las interacciones en Twitter revela que algunos candidatos lograron un alto nivel de engagement en sus publicaciones. Iglesias (JXC) fue el candidato que más interacciones recibió. Santoro (FDT) también generó un alto nivel de interacciones al centrarse en críticas al gobierno de CABA y la desinformación sobre las vacunas. Espert fue quien más interacciones recibió por parte de LLA, con su enfoque en la duración de la cuarentena y la economía, así como su crítica al gobierno en relación con la gestión de la pandemia y la vacunación.

En cuanto a los n-grams, los candidatos del FDT se enfocaron en términos relacionados con la campaña de vacunación. En contraste, los de JXC emplearon términos que expresaban críticas al gobierno nacional y las medidas de aislamiento. Los candidatos del FIT se centraron en temas como la utilidad pública y los trabajadores de la salud. Por último, los candidatos de LLA utilizaron n-grams relacionados con la duración de la cuarentena, la figura de Javier Milei y críticas al gobierno nacional.

Por último, en términos de cuestionamientos a las medidas de cuidado, adhesiones a teorías de la conspiración o difusión de información falsa, ciertos candidatos de JXC y todos los de LLA cuestionaron la cuarentena, la efectividad de las medidas sanitarias, la procedencia de las vacunas y su seguridad. Si bien no hubo discursos anti-vacunas explícitos, los cuestionamientos realizados colocaron un manto de duda con respecto a la campaña de vacunación. En definitiva, este tipo de análisis pone de manifiesto que la pandemia de COVID-19 tuvo un impacto significativo en la política argentina, la crisis sanitaria no solo acaparó la agenda política y mediática, sino que actuó como un catalizador para dirimir diferencias políticas.

## REFERENCIAS

- Aruguete, N. (2019). ¿Twitter acrecienta la polarización política? *BPOLITICS*, N° 02, 5.
- Aruguete, N., y Calvo, E. (2020). Coronavirus en Argentina: Polarización partidaria, encuadres mediáticos y temor al riesgo. *Revista SAAP*, 14(2), pp. 280-310.
- Alshaabi, T., Adams, J.L., Arnold, M.V., Minot, J.R., Dewhurst, D.R., Reagan, A.J., Danforth, C.M., y Dodds, P.S. (2021). Storywrangler: A massive exploratorium for sociolinguistic, cultural, socioeconomic, and political timelines using Twitter. *Science Advances*, 7(29).
- Boireau, M. (2014). Determining Political Stances from Twitter Timelines: The Belgian Parliament Case. *Proceedings of the 2014 Conference on Electronic Governance and Open Society: Challenges in Eurasia*, pp.145-151.

- Bracciale, R. y Martella, A. (2017). Define the populist political communication style: The case of Italian political leaders on Twitter. *Information, Communication & Society*, 20(9), pp.1310-1329.
- Castro Rojas, S., Pérez Rivera, L. y Amatta, J. M. (2016). El enjambre digital en la política argentina. Twitter en las campañas 2011-2013-2015. *Rizoma*, 4(1), 90.
- Fernando A. Iglesias [@FerIglesias]. (2020, diciembre 26). «Voy a inocularme cualquiera de las vacunas que aprueben los Estados Unidos o la Unión Europea. No me pongo como ejemplo ni soy antivacuna, todo lo contrario: Creo en la ciencia y en la transparencia de los asuntos públicos». Fin del comunicado. Que cada uno haga lo que quiera». [Tweet]. Twitter <https://twitter.com/FerIglesias/status/1342876514145226755>
- Fernando A. Iglesias [@FerIglesias]. (2021a, junio 19). «88.247 muertos. Solo 7% de la población vacunada. Más del doble de caída del PBI que Brasil y 50% más que Chile y Uruguay. Récord de cuarentena y escuelas cerradas. Robaron vacunas. Y se hacen los escandalizados por una declaración sobre la cual ya pidieron disculpas». #TUGO 48% [Tweet]. Twitter. <https://twitter.com/FerIglesias/status/1406287951764635654>
- Fernando A. Iglesias [@FerIglesias]. (2021b, septiembre 21). «Nos encerraron y fundieron a medio país por razones políticas, y no sanitarias ni epidemiológicas. Ahora decretan el fin de la cuarentena eterna por las mismas razones #ElPeorGobiernoDeLaHistoria». #ElPeorGobiernoDeLaHistoria [Tweet]. Twitter. <https://twitter.com/FerIglesias/status/1440368548904275972>
- Fernando A. Iglesias [@FerIglesias]. (2021c, diciembre 6). «No soy antivacunas pero no se ve ninguna razón para el pasaporte sanitario. Ni cifras de mortandad altas, ni emergencia hospitalaria, ni resistencia a vacunarse. Difundan los resultados positivos de la vacunación, pongan más vacunatorios y respeten las decisiones de la gente». [Tweet]. Twitter. <https://twitter.com/FerIglesias/status/1467982269289619457>
- Gualda, E. (2022). Antivacunas, antimascarillas, negacionismo y teorías de la conspiración en tiempos de COVID-19. Simposio: “Redes, medios y pandemia”. I Jornadas Transdisciplinarias UNI2 Experiencias de investigación y vinculación relacionadas con la pandemia y la post-pandemia, Universidad Nacional Mar del Plata.
- Hendricks, J. A. y Denton, R. (2010). Political campaigns and communicating with the electorate in the Twenty-First Century. *Communicator-in-chief: How Barack Obama used new media technology to win the White House*, pp.1-18.
- Infobae (2023, 10 de septiembre). Quién es Sandra Pitta, la científica que se cruzó con Alberto Fernández. <https://www.infobae.com/politica/2019/08/01/quien-es-sandra-pitta-la-cientifica-que-se-cruzo-con-alberto-fernandez/>
- José Luis Espert [@jlespert]. (2020b, agosto 17). «Difícil que sea peor que los casi 5 meses que hace que estamos encerrados en una cuarentena cavernícola». [Tweet]. Twitter. <https://twitter.com/jlespert/status/1295416526326112258>
- José Luis Espert [@jlespert]. (2020c, diciembre 23). «ANMAT todavía no autorizó la vacuna rusa que fuimos a buscar en vuelo épico con la quebrada

- Aerolíneas Argentinas y al mismo tiempo ANMAT ya aprobó la vacuna Pfizer que todavía no tenemos Al Borda*». [Tweet]. Twitter. <https://twitter.com/jlespert/status/1341710460689801216>
- José Luis Espert [@jlespert]. (2021, septiembre 30). «*Alberto Fernández merece juicio político. Murió gente de COVID por demorar este anuncio que ocurre solo por haber perdido una elección. Miserable es poco*». [Tweet]. Twitter. <https://twitter.com/jlespert/status/1443699249791766535>
- José Luis Espert [@jlespert]. (2020a, junio 11). «*Rumbo a los 100 días de cuarentena, la gente fundiéndose por no trabajar y sin embargo no paran de acelerarse los contagios de coronavirus. El fracaso del #YoMeQuedoEnCasa es total InútilesK*». [Tweet]. Twitter. <https://twitter.com/jlespert/status/1271216493964013568>
- Juares, W. y Ferreyra, S. (2021). Twitter, internautas vecinos y los intersticios de la polarización política. Un estudio centrado en la campaña electoral de 2019 en el Partido de General Pueyrredon, Argentina. *InMediaciones de la Comunicación*, 16(2), pp. 103-129.
- Kreiss, D. (2012). Taking our country back: The crafting of networked politics from Howard Dean to Barack Obama. Oxford University Press, USA.
- Leandro Santoro [@SantoroLeandro]. (2020a, abril 9). «*Quiéren levantar la cuarentena para «Salvar la Economía» pero la economía no está en crisis por la cuarentena, está en crisis por el impacto social q causa la epidemia en el mundo y por las políticas de ajuste que promovieron los economistas que piden levantar la cuarentena*». [Tweet]. Twitter. <https://twitter.com/SantoroLeandro/status/1248397931532537859>
- Leandro Santoro [@SantoroLeandro]. (2020b, abril 14). «*Hola @horaciorlarreta! ¿Viste q compraron barbijos a \$3000 c/u? Bueno, es un escándalo! Desde el @FrenteDeTodos hicimos un pedido de informes (104). Me parece q lo tenes q responder ya. Hace como @alferdez y actúa rápido! Súmale q el hermano de un funcionario tuyo está atrás!* 🤨». [Tweet]. Twitter. <https://twitter.com/SantoroLeandro/status/1250127565408423937>
- Leandro Santoro [@SantoroLeandro]. (2021a, mayo 25). «*En el peor momento de la pandemia «los médicos x la verdad» convocan marchas anticuarentena -sin usar barbijos- levantando la consigna de Patricia Bullrich, luego de q Pfizer la desmienta en un comunicado oficial y Carrió la corra x el lado de la racionalidad. Todo subrealista*». [Tweet]. Twitter. <https://twitter.com/SantoroLeandro/status/1397299890917715969>
- Leandro Santoro [@SantoroLeandro]. (2021b, junio 2). «*Que la spuntnik es veneno. Que las vacunas chinas q adquirimos son malas. Que para comprar Pfizer pidieron coimas. Que rechazamos vacunas de Covax por capricho. Todos los días arman una fake para meterle odio a la gente pero les dura un ratito. Son tan malos como insistidores!*». [Tweet]. Twitter. <https://twitter.com/SantoroLeandro/status/1400077347873865742>
- López-Rabadán, P., López-Meri, A. y Doménech-Fabregat, H. (2016). *La imagen política en Twitter. Usos y estrategias de los partidos políticos españoles*.

- Ornstein, N. J. (2000). *The permanent campaign and its future*. American Enterprise Institute.
- Políticos en Twitter. (2023). <https://politicosen twitter.com/>
- Romer, D. y Jamieson, K. H. (2020). Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. *Social Science & Medicine*, 263, 113356.
- Slimovich, A. S. (2016). Política y redes sociales en Argentina. El caso de los candidatos presidenciales de 2011 en Twitter. *Signo y Pensamiento*, XXXV(68), pp.86-100.
- Shurafa, C., Darwish, K. y Zaghouani, W. (2020). Political Framing: US COVID19 Blame Game. En S. Aref, K. Bontcheva, M. Braghieri, F. Dignum, F. Giannotti, F. Grisolia, & D. Pedreschi (Eds.), *Social Informatics* (pp. 333-351). Springer International Publishing.
- Zunino, E., Recalde Cerda, L., Baquerizo (2019). Líderes políticas argentinas en Twitter. Un estudio en el marco de las elecciones 2017. *Perspectivas de la comunicación*, 12(2), pp.221-243.



**III.**  
**TEORÍAS DE LA CONSPIRACIÓN Y**  
**DISCURSOS DE ODIO EN LÍNEA SOBRE**  
**INMIGRANTES Y/O REFUGIADOS**



# CONSPIRACY THEORIES ON THE ORIGIN OF THE CORONAVIRUS AND ITS IMPACT ON THE PERCEPTION OF CHINA AND CHINESE ON TWITTER<sup>1</sup>

ALBA TABOADA-VILLAMARÍN

*Investigadora Predoctoral en Formación (FPI)*

*Universidad Autónoma de Madrid, España*

*Universidad de Huelva, ESEIS, España*

CAROLINA REBOLLO-DÍAZ

*Investigadora Posdoctoral Margarita Salas*

*Universidad de Huelva, ESEIS/COIDESO, España*

**DOI:** 10.14679/2758

## **Summary:**

1. Introduction.

2. Methods.

2.2. Data analysis.

2.2.1. Text cleaning and tokenisation.

2.2.2. Topic classification and modeling.

2.2.3. Sentiment analysis.

3. Main issues surrounding the origin of the coronavirus.

3.1. China's role in the debate on the origin of the SARS-CoV-2 virus.

3.2. China's impact on the perception of the pandemic in the United States.

3.3 India-China Geopolitical Rivalry in the Context of the Health Crisis.

3.4. Evaluation of Chinese vaccination strategies and their global impact.

4. Main sentiments and conspiracy.

5. Conclusions.

References.

## 1. INTRODUCTION

The COVID-19 pandemic, which has affected millions of people worldwide, has created a perfect breeding ground for conspiracy theories. As the virus spreads, there are unfounded claims, rumours, and unsubstantiated accusations. As the initial epicentre of the epidemic, China became the main target of these

---

<sup>1</sup> This publication is part of the Project 'NON-CONSPIRA-HATE!' (PID2021-123983OB-I00), funded by MCIN/AEI/10.13039/501100011033, and by "ERDF A way of making Europe". It also is part of the 'CONCERN' Project (PID2020-115095RB-I00), funded by MCIN/AEI/10.13039/501100011033. It also benefited from the European Union, nextGeneration-UE and the Ministerio de Universidades, and the support of ESEIS (Social Studies and Social Intervention Centre) and COIDESO (Centro de Investigación en Pensamiento Contemporáneo e Innovación para el Desarrollo Social) at the Universidad de Huelva, Spain.

theories, accused of being responsible for the outbreak and orchestrating a global cover-up.

Conspiracy theories suggest that a group of individuals secretly congregates to carry out malicious actions for a particular purpose (Van Prooijen, 2019). Douglas, Sutton and Cichocka (2017) define them as explanations for important events involving secret plots by powerful and malevolent groups. Typically, these theories involve powerful entities, such as prominent figures in society, high-level government bodies (e.g., intelligence agencies), influential industrial sectors (e.g., oil or pharmaceutical companies), or marginalized groups that have been subject to stigmatization (e.g., Muslim or Jewish communities) (Van Prooijen and Van Vugt, 2018).

Research on the psychology of conspiracy theories suggests that epistemic, existential, and social motives drive beliefs in conspiracies (Douglas et al., 2019). Among these, it has been argued that conspiracy theories flourish in times of instability, lack of control, and vulnerability. These are periods when people are in a vulnerable position, characterised by mistrust and an inability to understand difficult situations in which they lack control (Douglas et al., 2019). Social crises, defined as rapid and disruptive changes in society that challenge established power structures, norms of behaviour, or even the existence of certain individuals or groups, have led to conspiracy theories. The adverse feelings people experience during crises, such as fear, uncertainty, and a sense of lack of control, drive the need to find meaning in a situation, increasing the likelihood that they perceive conspiracies in particular social contexts (Van Prooijen and Douglas, 2017). People are more likely to endorse conspiracy theories when faced with existential threats, anxiety, or uncertainty (Van Prooijen, 2019), which is often triggered by distressing social events, such as the COVID-19 pandemic.

Some conspiracy theories emerge from pre-existing intergroup tensions that may be linked to prejudices towards specific groups (Douglas, 2021; Douglas et al., 2019), as at the core of conspiracy beliefs lies an attribution of malign intentions, which serves both to explain and make sense of complex social events and to point to the supposed culprit of negative events. Thus, the identification of this antagonist can trigger intergroup conflicts (Van Prooijen, 2019) manifesting in the form of racism and political and social tensions, among others. Additionally, people often have a deep need to maintain a positive perception of the groups to which they belong, such as their nationalities, political affiliations, and religious communities. Belief in the idea that others are conspiring against their own group tends to arise most likely when the group is perceived as undervalued, disadvantaged, or threatened (Uscinski and Parent, 2014).

A classic example is the case of anti-semitic conspiracy theories. Jews have been the subject of numerous conspiracy theories throughout history, including the idea that they conspire to dominate the world and attempt to establish a secret world government (Swawi, 2012). Recently, they have also been accused of organising and engineering the coronavirus as a biological weapon (Gerstenfeld, 2020). Swawi (2012) found that belief in Jewish conspiracy theory was associated with anti-Israeli attitudes, modern racism directed at the Chinese, right-wing authoritarianism and social dominance orientation. Muslims have also been the subject of conspiracy theories such as Eurabia (Carr, 2006), a secret project between European politicians and the Arab world for the Islamisation of Europe, or the Great Replacement (Davey and Ebner, 2019), which claims that there is a conspiracy to replace the native populations of Europe and elsewhere with Muslim immigrants, supposedly leading to the destruction of Western culture and identity. In such situations, hatred and racism are intertwined with conspiracy theories. In worst-case scenarios, they can even motivate acts as serious as the Oslo massacre of 2011 (Fekete, 2012). These conspiracies are seen as false, misleading, or misinformation, and function as a way of constructing a collective identity by identifying as one group in opposition to another (Gagliardone et al., 2021).

Conspiracy theories on COVID-19 have varied. Some have pointed to alleged plots between the Chinese government and international bodies, such as the World Health Organisation (WHO) (Kalil et al., 2021), while others have implicated politicians, businessmen, and public figures in a global plot to establish a 'new world order' (Campana, 2021). All these theories share their ability to sow hatred and mistrust between people in both online and offline contexts. Social media, particularly Twitter, are platforms where cyber-racism and hostility towards Asians have spread. Several studies have analysed the presence of negative sentiments and anti-Chinese/Asian attitudes in coronavirus-related tweets (Dubey, 2020; Nguyen et al., 2020). Other studies that investigated the main themes on Twitter related to COVID-19 in the early-to the mid-2020s also identified racism as a recurring theme (Abd-Alrazaq et al., 2020; Chandrasekaran et al., 2020). Offline hate has manifested itself in acts of discrimination and physical assault against Asian Americans (Stop AAPI Hate Reporting Center, n.d.). Olesky et al. (2021) found that general conspiracy theories were associated with the acceptance of anti-democratic policies towards people of other ethnic groups and negative attitudes towards people from countries associated with the spread of COVID-19. These included not only China, but also Italy, which was one of the first major outbreaks outside China (Sorokowski et al., 2020). Exposure to conspiracy theories can lead to discriminatory actions that can spill over to other groups not involved in the conspiracy, suggesting a generalisation of negative attitudes (Jolley, Meleady and Douglas, 2020).

This chapter deals with a paradigmatic case study in the context of racism experienced during the COVID-19 pandemic, focusing specifically on the English-speaking population and their perception of the Chinese population in the virtual environment. The period of analysis focused on a critical moment: the publication of the World Health Organization's (WHO) preliminary report on the origin of the SARS-CoV-2 virus at the end of March 2021. The purpose of this study is twofold. It seeks to reconstruct the virtual portrait that has emerged of the Chinese population throughout the COVID-19 pandemic, specifically regarding the controversy surrounding the origin of the virus. On the other hand, it seeks to understand the relationship between conspiracy theories and hate narratives in an anti-Chinese context during a health crisis.

This study makes use of a combined methodology that includes natural language processing and unsupervised machine learning techniques through topic modeling and sentiment analysis, along with qualitative analysis to delve deeper into the themes obtained through computational methods. A sample of 75 thousand tweets related to the COVID and China themes from the end of March to the end of April 2021 will be examined to shed light on the complex relationship between conspiracy theories and expressions of dislike directed towards the Chinese community during the pandemic. This study seeks to provide a deeper understanding of how online narratives and attitudes can fuel and reflect racism and xenophobia during times of crisis.

## 2. METHODS

### 2.1. Case study selection and data collection

On 30 March, 2021, the WHO released a preliminary report on «the zoonotic origin of the virus and the route of introduction into the human population, including the possible role of intermediate hosts, including through initiatives such as scientific and collaborative field missions» (WHO, 2021). Tentatively, this report intended to settle the controversy that underpinned the international health crisis over the possible origins of the SARS-CoV-2 virus and the main reason for the creation and spread of multiple conspiracy theories.

In another recent study by authors Rebollo-Díaz and Taboada-Villamarín (in press), in which three million tweets were collected from 2020 to 2022 containing the words “COVID” together with “China”, “Chinese” or “Wuhan”, an increase in the publication of tweets containing racist attitudes and rejection of the Chinese population was detected after the dissemination of this report. Therefore, in this paper, we want to explore this case study starting from the findings of the aforementioned publication and taking the tweets published from 30 March 2021 to 30 April 2021 from the total set of three million tweets, which were downloaded using the free Twitter API v2 for researchers with the

help of the Python programming language and the “tweepy” library. The sample for this study comprised 75,328 tweets. The criteria for downloading tweets are summarised in Table 1.

Table 1. Data collection features

<b>Access Api</b>	Api v2 Twitter
<b>Search syntax</b>	“COVID” AND (“china” or “chinese” or “wuhan”)
<b>Language of tweets</b>	English
<b>Time window</b>	30/03/2021 - 30/04/2021
<b>Total number of tweets in the sample</b>	75.328
<b>Retweets</b>	No

Source: Authors.

## 2.2. Data analysis

The sample was studied using a mixed model to carry out content analysis of the narratives expressed in the tweets under study. To this end, a first quantitative computational approach was integrated, applying natural language processing techniques and unsupervised machine learning through topic modeling and sentiment analysis. These analyses were performed in the programming languages Python 3.11.5 and R-4.3.1, respectively. Second, an in-depth analysis of a part of the sample was conducted following a qualitative approach to exemplify and empirically argue the results obtained through the computational models.

### 2.2.1. Text cleaning and tokenisation

Data preprocessing was performed using text cleaning and tokenisation. To enable the application of unsupervised learning techniques to data stored in text form, it is necessary to normalise and purify the text corpus, eliminating empty words such as conjunctions and prepositions, as well as links and emoticons that did not form part of the analysis. Text cleaning was performed by applying the NLTK Python library (Bird et al., 2009) in two phases. After the preliminary results, the keywords used in the download syntax “China”, “Chinese”, “COVID” and “Wuhan” were removed, as they were reiterated in the results of the models and did not provide relevant information in line with the objectives of this study. After debugging, the text was fragmented into smaller units following the tokenisation process (Saleem et al., 2021). This process was performed using the Python Gensim library considering two lengths: unigrams (n=1) and bigrams (n=2).

### 2.2.2. *Topic classification and modeling*

Unsupervised learning approaches were used to identify the main topics of discussion present in the sample, as well as the quantitative distribution in the analysed tweets and the degree of homogeneity or heterogeneity of these topics. A first classification exercise required the application of the K-means statistical model through the “metrics” module of the Sklearn library. Its implementation requires the prior specification of the number of clusters into which the tweets will be grouped. The result of the classification was evaluated by creating groupings composed of two to fifteen different clusters. According to the quantitative and theoretical evaluation, the total number of four clusters turned out to be suitable for this research. The application of the K-means model made it possible to classify each of the tweets into similarity clusters, as well as to visualise their distribution in a scatter plot (see Figure 1).

In the context of topic modeling, the algorithm known as LDA (Latent Dirichlet Allocation) was implemented through the “LDA” module provided by the “pyspark.ml.clustering” library. The LDA algorithm is widely used in text mining to identify patterns in word usage, which in turn enables the labelling and categorisation of text sets according to the topics addressed. The choice to carry out clustering before the implementation of topic modeling was based on empirical evidence of improved results, as documented in previous research (Gualda, Taboada-Villamarín and Rebollo-Díaz, 2023). Finally, the deduction of the titles assigned to each topic is based on the keywords that the model suggests as the most relevant (see Table 2). After the analysis of these keywords, the evaluation was carried out by two experts, who examined a random sample of the tweets included in each grouping to carry out a detailed analysis of them.

### 2.2.3. *Sentiment analysis*

To identify and understand in greater depth the predominant emotions in the tweets submitted for analysis, the tool provided by the “syuzhet” library was used in the context of the R programming language (Isasi, 2020). This multilingual package allows sentiment analysis to be applied with the NRC Emotion Lexicon from the same package (Mohammad & Turney, 2010). It represents an essential resource in the field of emotional analysis of texts, allowing a detailed assessment of up to eight different emotions. The emotions it can detect include fear, anger, joy, disgust, sadness, anticipation, trust and surprise. These emotions provide a nuanced emotional context that sheds light on the underlying reaction and tone of the analysed tweets. The “syuzhet” library works by assigning weighted values to the different emotions present in the text. In our case, this was achieved by processing and analysing the cleaned text corpus.

### 3. MAIN ISSUES SURROUNDING THE ORIGIN OF THE CORONAVIRUS

The sample analysed has provided insight into the significance of the publication and dissemination of the Official Global Report on the origins of the SARS-CoV-2 virus produced by the World Health Organisation (WHO, 2021) in the virtual public discussion environment, shedding light on the representation that civil society in English-speaking countries has outlined about the Chinese population and the health crisis. During April 2021, the classification and Topic Modeling allowed for the identification of four central axes of public discussion concerning China and COVID-19, through which various points of view were expressed. Table 2 shows the classification of the main topics resulting from the application of the LDA (Latent Dirichlet Allocation) model and the respective names given by the researchers to each cluster to identify the content of the tweets grouped in each category: 1) China's role in the debate on the origin of the SARS-CoV-2 virus, 2) China's impact on the perception of the pandemic in the United States, 3) India-China geopolitical rivalry in the context of the health crisis, and 4) Evaluation of Chinese vaccination strategies and their global impact.

Table 2. Topic Modeling result by cluster

Cluster number	Cluster name	Tweets	%	Keywords
1	China's role in the debate on the origin of the SARS-CoV-2 virus	24.528	32,5	"lab", "report_origins", "lak", "investigation", "animal"
2	China's impact on perceptions of the pandemic in the United States	20.766	27,6	"virus", "people", "trump", "us", "government"
3	India-China Geopolitical Rivalry in the Context of the Health Crisis	19.796	26,3	"coronavirus", "india", "world", "uk", "control"
4	Evaluation of Chinese vaccination strategies and their global impact	10.238	13,6	"vaccines", "data", "cases", "health", "sinopharm"

Source: Authors.

#### 3.1. China's role in the debate on the origin of the SARS-CoV-2 virus

The predominant theme in the analysis corresponds to the debate on the origin of the SARS-CoV-2 virus, comprising 32.5% of the tweets examined. The narratives investigated exhibit a critical attitude towards the publication of the report on the origin of the SARS-CoV-2 virus, expressing significant distrust towards the World Health Organisation and the government of the People's Republic of China. Questions are raised about possible prior alliances motivated

by conflicts of interest, both economic and political, and alternative theories are proposed concerning the “creation” or spread of COVID-19. The content analysis shows that the mention of ‘China’ on numerous occasions carries a direct reference to the government of the People’s Republic of China and thus to the Chinese Communist Party (CCP). The remarks include a pejorative discourse on the policies implemented in the Chinese population to address the health crisis, delineating a government characterised by manipulation, censorship of its researchers and withholding of information. This government is portrayed as dehumanised, with masked objectives, such as consolidating its position as a leading economic and technological power through the virus, which also exerts significant influence and dominating power over other governments and institutions with global reach.

### **3.2. China’s impact on the perception of the pandemic in the United States**

Second, the tweets grouped under the cluster “China’s impact on the perception of the pandemic in the United States” represent 27.6% of the total sample set. In this cluster, China’s responsibility for the adverse effects of the pandemic experienced by the US population is examined. The content analysis points to a marked disagreement in terms of how different political representatives have handled policies related to the health crisis in different US states, and a discordance between the political parties that have been in power from the beginning of the pandemic until the time of analysis. In particular, reference is made to the alleged connection between the US Democratic Party and Chinese Communist government elites, based on relations between the US president’s former medical adviser, Anthony Fauci, and Chinese leader Xi Jinping. Mentions of a conspiracy between the Chinese party and the Democratic party to unseat President Trump in the last election are frequent, even going so far as to hint at an alleged Communist Party-financed “theft” of the election.

### **3.3. India-China Geopolitical Rivalry in the Context of the Health Crisis**

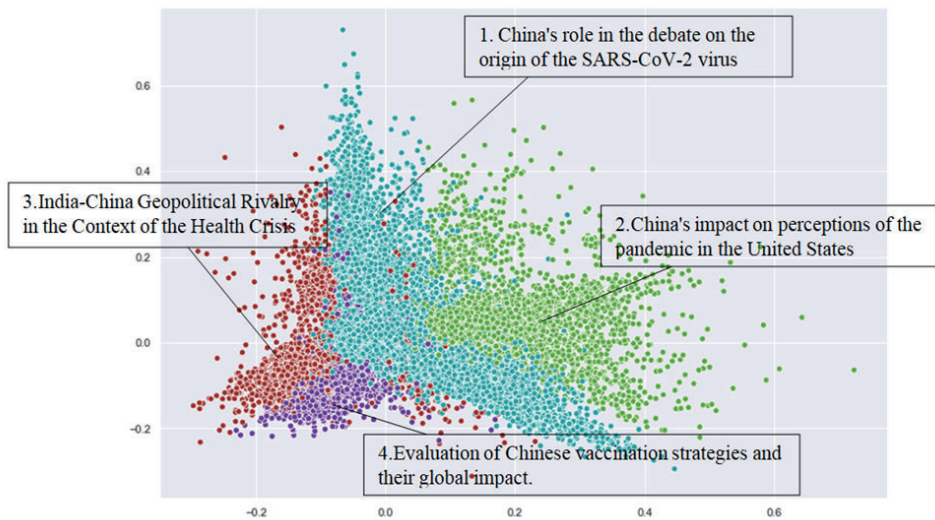
In third place, with 26.3% of the tweets in the set analysed, a latent geopolitical conflict between India and China stands out as another of the central topics of discussion. The predominant theme in the messages is related to a historical conflict in terms of territory and political autonomy between the two countries. The main perspective of these tweets is that of the Indian population and is directly linked to the idea of China’s creation of a biological weapon to exert control over a part of Indian territory. In addition, China’s interest in achieving economic domination of India through economic support and medical supplies to combat COVID-19 is mentioned. Allusions are also

made to internal conflicts on Indian territory related to these relations. The People's Republic of China is depicted as a political rival that needs to be defended against because of its quest for silent colonisation.

### 3.4. Evaluation of Chinese vaccination strategies and their global impact

Finally, the fourth cluster comprises 13.6% of the total number of Twitter posts under analysis. This theme is characterised by divergent opinions expressed in the tweets regarding the vaccination strategies adopted by the Government of the People's Republic of China and their impact on Western countries. Content analysis reveals concerns about the reliability of Chinese vaccines, raising questions about why other governments are willing to use them. The tweets include information about possible side effects and portray the Chinese government as a suspicious actor, referring to the theory that it may have introduced the COVID-19 virus and then presented a premature solution. Recurrent in these tweets are links and news stories mentioning faulty vaccines, their possible redundancy and the possibility of a hidden agenda that the Chinese government does not disclose.

Figure 1. K-means Scatter Graph



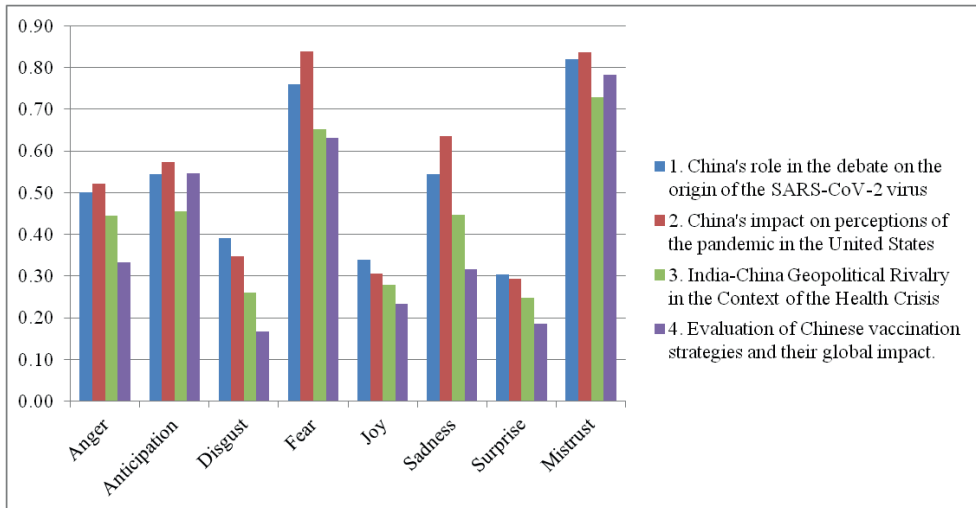
Source: Authors.

It is important to highlight that each identified theme shares transversal axes or perspectives. The scatter graph (Figure 1) provides a visualisation of the distance between the themes, represented in different colours, indicating



or uncertainty around China’s actions on these issues. On the other hand, in the theme “India-China Geopolitical Rivalry in the Context of the Health Crisis,” negative emotions such as fear (0.65) and distrust (0.73) are relevant, reflecting a perception of conflict and mistrust in India-China relations in the context of the health crisis. There is also a moderate presence of emotions such as anger and sadness. In the theme “Evaluation of Chinese vaccination strategies and their global impact,” the emotion of mistrust stands out (0.78), suggesting a high degree of scepticism about Chinese vaccination strategies and their influence globally. Anticipation is also relevant in this context, which could indicate an expectation regarding China’s future actions in the field of vaccination.

Figure 3. Sentiment analysis by cluster



Source: Authors.

Additionally, a close examination of the tweets in the sample revealed both direct and indirect mentions of conspiracy theories. Two conspiracy theories have emerged recurrently in all thematic areas and have been widely disseminated at the national level during the period of the health crisis. These two main theories are intrinsically related to the origin of the virus and are intertwined with the particular nuances of China observed in each cluster. In Table 3, we add representative examples of the tweets analyzed in the sample that contain the highest score in emotions such as mistrust and fear.

Table 3. Examples of tweets in the sample

<b>COVID-19 as biological weapon</b>	<i>@indiatvnews The second wave of Covid 19 is a biological warfare created by China to destabilise India and China. make India lag behind with the support of India's Anti Modi Group. Even the idea was to blame the central government. It was started on Decem</i>
	<i>@POTUS Now China starts biochemical world war 3ed, get ready to send medical teams to all countries. Covid is not a pandemic, it's a biological weapon and China made more and harder weapons.</i>
	<i>Breaking news: China caused the covid. It knowingly spread it to the world. Yet it faced no consequences. We all seem to have forgotten. Related: Politicians from 100 countries exploited the tragedy to seize power on a grand scale.</i>
	<i>Fauci funded Wuhan research that created Covid-19 Republican Representative Scott Perry of Pennsylvania said it is deeply troubling that the federal infectious disease research organisation headed by Dr Anthony Fauci has bypassed federal oversight of a grant...</i>
	<i>The WHO report on Covid's origins is best understood as a cover-up strongly influenced by the Chinese Communist Party and Westerners with conflicts of interest."</i>
	<i>Fauci funded the Wuhan lab and authorised the procurement of genetically engineered functions at Wuhan to make bat viruses more dangerous. If SARS-CoV-2 came from the Wuhan lab, Fauci is the father. American biologist Bret Weinstein's 90 per cent chance that it leaked from the lab.</i>
<b>COVID-19 as a result of a leak from the Wuhan laboratory</b>	<i>What a coincidence: "Bill Gates linked to Chinese company 'extracting' DNA data from Americans through covid testing"</i>
	<i>There is no firm conclusion on the origins of Covid, World Health Organisation report. stop lying cv19 came from the wuhan lab, we all know that, in fact, the question is whether they did it on purpose.</i>
	<i>Wuhan lab aided Chinese military in secret project to find animal viruses #Newsmax via @Newsmax</i>
	<i>Why do the mainstream media quote Chinese Communist Party-funded consultants to discredit theories of the origin of the COVID laboratory?</i>
	<i>WHO report on Covid-19 written by the Chinese Gestapo government - WHO officials must be held accountable for the whitewashed rubbish they will publish tomorrow! The true origins of the virus are the Wuhan laboratory.</i>
	<i>26 out of 27 scientists who dismissed the Wuhan link, leading to the Wuhan theory being framed as a racist conspiracy", had links to the Wuhan lab, yet these reports had been classified as "conspiracy". Bullshit. Journalism is there to challenge official narratives.</i>

Source: Authors.

First, significant weight is given to the theory that the SARS-CoV-2 virus was created as a biological weapon devised by China. This theory hypothesises the artificiality of the virus and its intentional creation by the Chinese state, with the aim of using it as a biological weapon in other countries, with objectives

ranging from territorial issues to control, surveillance and even anti-democratic objectives, among others. Secondly, there are multiple mentions of the theory that the virus escaped unintentionally from a laboratory. In this theory, it is argued that the spread of COVID-19 among the civilian population was the result of an accidental leak at the Wuhan Institute of Virology, where it is assumed that the virus was already being handled and studied beforehand.

## 5. CONCLUSIONS

In this work, our aim was to analyze the online representation that has emerged in relation to China during the COVID-19 pandemic, with a particular focus on the controversy surrounding the origin of the virus and associated conspiracy theories. The analysis of the sample of tweets has provided insight into the main narratives and emotions present in the online public discourse.

The discussion on Twitter following the release of the WHO report has primarily centered around the origin of the virus, its impact on international relations, and Chinese vaccination strategies. In all these topics, there is a notable expression of high distrust towards the Chinese government (as well as towards the WHO), which is attributed with hidden agendas, malicious intentions, and secret conspiracies with other actors, leading to belief in various conspiracy theories. Particularly, the theory of intentional creation of the virus as a biological weapon and the theory of accidental leakage from a laboratory play significant roles in Twitter conversations and contribute to the distrust towards China. Furthermore, geopolitical tensions and rivalries between countries, such as China and India, have also been reflected in the discussions. An interesting aspect of these beliefs is reflected in the second cluster (Impact of China on the perception of the pandemic in the United States) about a supposed alliance between the Democratic Party and the Chinese government. These results confirm what has already been discussed by Uscinski and Parent (2014) and Smallpage, Enders, and Uscinski (2017) regarding how people on the losing side of political processes also seem more prone to believe and spread conspiracy theories about their political opponents.

The data has revealed the presence of negative emotions such as fear and distrust, suggesting a negative perception in discussions and opinions surrounding China's actions in various aspects of the health and geopolitical crisis. Lack of trust is primarily directed towards China and the WHO, while fear may also be related to the spread of the virus and geopolitical implications. For example, the conflict between India and China and the belief that the virus is a biological weapon to exert control over the country may be responses to feelings of threat, vulnerability, and lack of control, creating a fertile ground for conspiracy theories (Douglas et al., 2019). Additionally, anticipation also

plays a relevant role in the perception of China's future actions in these areas. Sentiment analysis can be a useful resource when approaching large unlabeled datasets where conspiracy theories may exist, as it can help detect them through emotions related to this type of discourse (feelings of fear, distrust, and anticipation).

Despite the increase in racism and Sinophobic attitudes during the COVID-19 pandemic, in our results, it is difficult to determine the extent to which a racist or hateful attitude is shown towards the Chinese civilian population as opposed to criticism directed at their government. In other words, distrust, aversion, and hostile attitudes are primarily directed towards the Chinese Communist Party. This may be due to the historical-political relationships that some Western countries like the United States have had with China due to geopolitical and economic competition, which can portray China as a threat. However, it is important to highlight the risk this poses when these sentiments permeate the perception of the Chinese population as a whole in a context of intergroup conflicts.

There are further opportunities to investigate the online representation of China during the COVID-19 pandemic in different contexts and platforms, which would provide a more comprehensive understanding of how public perceptions may vary worldwide. For example, since this study focused on English-language data with an emphasis on the American and Indian context, it would be valuable to conduct similar analyses in other languages and countries with different political and cultural relationships with China. This could shed light on variations in public perception and online narratives in different regions of the world. In addition to Twitter, exploring other social media platforms such as Telegram, Instagram, or Weibo in China could provide a more complete view of how the Chinese population is represented online and how conspiracy theories are disseminated in different digital environments.

## REFERENCES

- Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Z. (2020). Top Concerns of Tweeters During the COVID-19 Pandemic: Inveigillance Study. *J Med Internet Res*, 22(4), e19016. <https://doi.org/10.2196/19016>
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python* (1st ed). O'Reilly.
- Campana, M. N. (2021). Covid19 y el "nuevo orden mundial": el rol de las teorías conspirativas durante la pandemia. *Anuario del Centro de Investigaciones Jurídicas y Sociales*, XIX, 11-2021, 495-508. <http://hdl.handle.net/11336/148449>

- Carr, M. (2006). You are now entering Eurabia. *Race & Class*, 48(1), 1-22. <https://doi.org/10.1177/0306396806066636>
- Chandrasekaran, R., Mehta, V., Valkunde, T., & Moustakas, E. (2020). Topics, Trends, and Sentiments of Tweets About the COVID-19 Pandemic: Temporal Inveigilance Study. *J Med Internet Res*, 22(10), e22624. <https://doi.org/10.2196/22624>
- Davey, J., & Ebner, J. (2019). The Great Replacement': The violent consequences of mainstreamed extremism. *Institute for Strategic Dialogue*, 7, 1-36.
- Douglas, K. M. (2021). COVID-19 conspiracy theories. *Group Processes & Intergroup Relations*, 24(2), 270-275. <https://doi.org/10.1177/1368430220982068>
- Douglas, K. M., Sutton, R. M., & Cichocka, A. (2017). The psychology of conspiracy theories. *Current directions in psychological science*, 26(6), 538-542. <https://doi.org/10.1177/0963721417718261>
- Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding conspiracy theories. *Political psychology*, 40, 3-35. <https://doi.org/10.1111/pops.12568>
- Dubey, A. D. (2020). The Resurgence of Cyber Racism During the COVID-19 Pandemic and its Aftereffects: Analysis of Sentiments and Emotions in Tweets. *JMIR Public Health and Surveillance*, 6(4), e19833. <https://doi.org/10.2196/19833>
- Fekete, L. (2012). The Muslim conspiracy theory and the Oslo massacre. *Race & Class*, 53(3), 30-47. <https://doi.org/10.1177/0306396811425984>
- Gagliardone, I., Diepeveen, S., Findlay, K., Olaniran, S., Pohjonen, M., & Tallam, E. (2021). Demystifying the COVID-19 infodemic: Conspiracies, context, and the agency of users. *Social Media+ Society*, 7(3), 20563051211044233. <https://doi.org/10.1177/20563051211044233>
- Gerstenfeld, M. (2020). Anti-Jewish coronavirus conspiracy theories in historical context. *BESA Center Perspectives Paper*, 1(513), 1-4. <https://www.jstor.org/stable/resrep26356.12>
- Gualda, E., Taboada-Villamarín, A. & Rebollo-Díaz, C (2023). 'Big data' y ciencias sociales. Una mirada comparativa a las publicaciones de antropología, sociología y trabajo social. *Gazeta de Antropología* 39(1), artículo 09. <https://doi.org/10.30827/Digibug.79779>
- Jolley, D., Meleady, R., & Douglas, K. M. (2020). Exposure to intergroup conspiracy theories promotes prejudice which spreads across groups. *British Journal of Psychology*, 111(1), 17-35. <https://doi.org/10.1111/bjop.12385>
- Kalil, I., Silveira, S. C., Pinheiro, W., Kalil, Á., Pereira, J. V., Azarias, W., & Amparo, A. B. (2021). Politics of fear in Brazil: Far-right conspiracy theories on COVID-19. *Global Discourse*, 11(3), 409-425. <https://doi.org/10.1332/204378921X16193452552605>
- Levinsson, A., Miconi, D., Li, Z. Y., Frounfelker, R. L., & Rousseau, C. (2021). Associations between endorsement of conspiracy theories and sympathy

- for violent radicalization in young adults during the COVID-19 pandemic: moderation by psychological distress. *International Journal of Environmental Research and Public Health*, 18(15), 7846. <https://doi.org/10.3390/ijerph18157846>
- Mohammad, S., & Turney, P. (2010, June). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text* (pp. 26-34).
- Nguyen, T. T., Criss, S., Dwivedi, P., Huang, D., Keralis, J., Hsu, E., Phan, L., Nguyen, L. H., Yardi, I., Glymour, M. M., Allen, A. M., Chae, D. H., Gee, G. C., & Nguyen, Q. C. (2020). Exploring U.S. Shifts in Anti-Asian Sentiment with the Emergence of COVID-19. *International Journal of Environmental Research and Public Health*, 17(19), Article 19. <https://doi.org/10.3390/ijerph17197032>
- Oleksy, T., Wnuk, A., Maison, D., & Łyś, A. (2021). Content matters. Different predictors and social consequences of general and government-related conspiracy theories on COVID-19. *Personality and individual differences*, 168, 110289. <https://doi.org/10.1016/j.paid.2020.110289>
- Rebollo-Díaz, C. y Taboada-Villamarín, A. (forthcoming). Anti-Chinese attitudes and nationalism during the coronavirus pandemic. A longitudinal analysis of Twitter data. En S. Wang y T. Franca (Eds.), *Chinese Diasporas in the COVID-19 Pandemic*. Brill.
- Saleem, Z., Alhudhaif, A., Qureshi, K. N., & Jeon, G. (2021). Context-aware text classification system to improve the quality of text: A detailed investigation and techniques. *Concurrency and Computation: Practice and Experience*, 35(15), e6489. <https://doi.org/10.1002/cpe.6489>
- Smallpage, S. M., Enders, A. M., & Uscinski, J. E. (2017). The partisan contours of conspiracy theory beliefs. *Research & politics*, 4(4), 2053168017746554. <https://journals.sagepub.com/doi/full/10.1177/2053168017746554>
- Sorokowski, P., Groyecka, A., Kowal, M., Sorokowska, A., Białek, M., Lebuda, I., ... & Karwowski, M. (2020). Can information about pandemics increase negative attitudes toward foreign groups? A case of COVID-19 outbreak. *Sustainability*, 12(12), 4912. <https://doi.org/10.3390/su12124912>
- Stop AAPI Hate Reporting Center. (s. f.). *Website for the Asian American Pacific Islander Equity Alliance*. Recuperado 6 de marzo de 2023, de <https://aapiequityalliance.org/>
- Swami, V. (2012). Social psychological origins of conspiracy theories: The case of the Jewish conspiracy theory in Malaysia. *Frontiers in Psychology*, 3, 280. <https://doi.org/10.3389/fpsyg.2012.00280>
- Uscinski J. E. & Parent J. M. (2014). *American conspiracy theories*. New York, NY: Oxford University Press.
- Van Prooijen, J. W. (2019). An existential threat model of conspiracy theories. *European Psychologist*. <https://doi.org/10.1027/1016-9040/a000381>

- Van Prooijen, J. W., & Douglas, K. M. (2017). Conspiracy theories as part of history: The role of societal crisis situations. *Memory studies*, *10*(3), 323-333. <https://doi.org/10.1177/1750698017701615>
- Van Prooijen, J. W., & Van Vugt, M. (2018). Conspiracy theories: Evolved functions and psychological mechanisms. *Perspectives on psychological science*, *13*(6), 770-788. <https://doi.org/10.1177/1745691618774270>
- WHO (2021). WHO-convened Global Study of the Origins of SARS-CoV-2 <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/origins-of-the-virus>



# CHEQUEANDO DISCURSOS DE ODIOS EN LÍNEA SOBRE LA MIGRACIÓN VENEZOLANA EN ECUADOR<sup>1</sup>

ISIDRO MARÍN-GUTIÉRREZ

*Profesor Ayudante Doctor*

*Universidad de Málaga, España*

MÓNICA HINOJOSA-BECERRA

*Profesora Ayudante Doctora*

*Universidad de Málaga, España*

ANGEL LUIS TORRES-TOUKOUMIDIS

*Docente - Investigador*

*Universidad Politécnica Salesiana, Ecuador*

**DOI:** 10.14679/2759

## **Sumario:**

1. Introducción.
  2. Marco teórico.
  3. La crisis migratoria venezolana.
  4. Metodología.
  5. Resultados.
  6. Conclusiones.
- Referencias.

## 1. INTRODUCCIÓN

La investigación se enfoca en la relación entre la desinformación y la migración venezolana en Ecuador. Nuestro objetivo es indagar en las noticias sobre los venezolanos expuestas desde la agencia oficial de verificación ecuatoriana, *Ecuador Chequea*, para contrastar las tipologías de noticias falsas que se presentan junto con el discurso que se encuentra en las mismas. Se busca analizar cómo la desinformación ha afectado el comportamiento de la población ecuatoriana hacia los venezolanos y cómo las agencias de fact-checking pueden combatirla.

Los discursos de odio son expresiones que buscan denigrar, discriminar o incitar a la violencia contra un grupo de personas debido a su origen étnico,

---

<sup>1</sup> Este trabajo forma parte del Proyecto I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGBTI [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDER A way of making Europe”. También agradecemos el apoyo de los centros de investigación de la Universidad de Huelva, el grupo de investigación “Estudios Sociales E Intervención Social, ESEIS” y el centro “Pensamiento Contemporáneo e Innovación para el Desarrollo Social, COIDESO”.

nacionalidad, religión, género, orientación sexual u otra característica (Torres y Taricco, 2019). Estos discursos pueden manifestarse de diversas formas, desde comentarios ofensivos y chistes despectivos hasta amenazas o ataques violentos (Falxa, 2014).

En el contexto de la migración venezolana en Ecuador, los discursos de odio se han vuelto cada vez más comunes en línea, especialmente en las redes sociales (Altamirano y Torres-Toukourmidis, 2021; Párraga Pilay, 2021). Muchos ecuatorianos, a través de las redes sociales, han expresado su frustración y resentimiento hacia los migrantes venezolanos, acusándolos de quitarles empleos y recursos, ser delincuentes o propagar enfermedades (Sánchez Moscoso, Candela & Torres-Toukourmidis, 2022). Estos discursos no solo son falsos y estereotipados, sino que también alimentan la xenofobia y el racismo, lo que puede llevar a la discriminación y la violencia en la vida real (Ramírez Salado, 2022).

## 2. MARCO TEÓRICO

Una de las herramientas más importantes para combatir la desinformación han sido las agencias de fact-checking. La verificación de noticias se define como un filtro operativo de datos para evidenciar mentiras, errores, falta de rigor e inexactitudes de contenidos difundidos por los medios (Morejón-Llamas, 2023).

Una agencia de fact-checking es una organización o empresa que se dedica a verificar la veracidad de la información y los hechos presentados en los medios de comunicación, discursos políticos, publicaciones en línea y otros contextos. En resumen, el propósito de una agencia de fact-checking es examinar las afirmaciones y declaraciones para determinar su grado de precisión y proporcionar una evaluación imparcial basada en la evidencia disponible (Almeida Santos, Teresa Peixinho, Lopes & Araújo, 2023).

Estas agencias utilizan métodos de investigación y verificación rigurosos para examinar la información y verificar su exactitud. Esto puede incluir el análisis de fuentes confiables, la consulta de expertos, la revisión de datos y evidencia científica, la comparación de afirmaciones con registros históricos y la evaluación de la consistencia lógica de la información presentada (Pérez Curiel & Velasco Molpeceres, 2020).

Una vez que una agencia de fact-checking ha realizado su investigación, emite un informe o una calificación que clasifica la declaración verificada como verdadera, falsa, engañosa, exagerada o insuficiente en evidencia para tomar una determinación definitiva. Estos hallazgos se suelen publicar en sitios web o en medios de comunicación para que el público pueda acceder a ellos y tomar decisiones informadas basadas en hechos verificados (Suárez-Roca & Vélez-Bermello, 2022).

Las agencias de fact-checking desempeñan un papel importante en el periodismo de verificación y en la lucha contra la desinformación, ya que ayudan a promover la transparencia, la responsabilidad y la precisión en el discurso público. Sus esfuerzos contribuyen a mejorar el acceso a información confiable y a reducir la propagación de noticias falsas (Olmos Alcaraz, 2023).

La agencia de fact-checking que se enfoca en las noticias sobre los venezolanos en Ecuador es *Ecuador Chequea*, una agencia ecuatoriana avalada por Poynter, International Fact-Checking Network (Palacios & Cusot, 2019). *Ecuador Chequea* se dedica a verificar la veracidad de las afirmaciones hechas por políticos, figuras públicas y medios de comunicación en el contexto ecuatoriano proporcionando información precisa y confiable para que el público pueda tomar decisiones informadas y combatir la desinformación (Vélez Bermello, 2020).

*Ecuador Chequea* utiliza métodos de investigación y verificación rigurosos para examinar las declaraciones y afirmaciones. Verifican la información a través de la consulta de fuentes confiables, la revisión de datos y evidencia, y la realización de entrevistas con expertos en los temas pertinentes. Una vez completada la verificación, emiten informes detallados que explican su evaluación y califican la veracidad de las afirmaciones (Sánchez Moscoso, Candela & Torres-Toukoumidis, 2022).

Esta organización tiene un sitio web donde publican sus verificaciones (<https://ecuadorchequea.com/>), ofreciendo acceso gratuito a la información verificada y a sus conclusiones. A través de su trabajo, *Ecuador Chequea* contribuye a promover la transparencia y la rendición de cuentas en el discurso público en Ecuador.

### 3. LA CRISIS MIGRATORIA VENEZOLANA

La crisis migratoria venezolana se refiere al flujo masivo de ciudadanos venezolanos que han abandonado su país en los últimos años debido a la grave crisis política, económica y social que enfrenta Venezuela. Esta crisis ha tenido un impacto significativo en varios países de la región, incluido Ecuador (Burbano Alarcón, Zaldívar & Vera Puebla, 2019).

Miles de venezolanos han buscado refugio y oportunidades en Ecuador, convirtiéndolo en uno de los destinos principales para los migrantes venezolanos. La llegada masiva de venezolanos ha generado desafíos y tensiones en Ecuador, tanto a nivel social como económico (Heredia & Battistesa, 2018).

En el ámbito social, el flujo migratorio ha planteado retos en términos de integración y convivencia. La rápida llegada de un gran número de migrantes ha generado tensiones en algunas comunidades receptoras, especialmente en ciudades fronterizas y zonas urbanas con recursos limitados. Existen preocu-

paciones relacionadas con el acceso a servicios básicos como vivienda, salud, educación y empleo, así como tensiones culturales y sociales (Bahar, Dooley & Selee, 2020).

Desde el punto de vista económico, la migración venezolana ha tenido un impacto en el mercado laboral y en la economía en general. La oferta de mano de obra ha aumentado, lo que ha generado competencia en sectores como el comercio, la construcción y el trabajo informal. Al mismo tiempo, la llegada de emprendedores y profesionales venezolanos ha contribuido al dinamismo económico en algunos sectores (Ripoll & Navas-Alemán, 2018).

El gobierno ecuatoriano ha implementado diversas medidas para hacer frente a la crisis migratoria venezolana. Esto incluye la introducción de políticas de regularización migratoria, como el otorgamiento de visas especiales y un plan de regularización. Estas iniciativas tienen como objetivo facilitar la regularización de la situación migratoria de los venezolanos y promover su integración en la sociedad ecuatoriana (Ruiz Muriel & Álvarez Velasco, 2019). Se han establecido mecanismos de coordinación con organizaciones internacionales y la sociedad civil para brindar asistencia humanitaria, atención médica y apoyo a los migrantes venezolanos más vulnerables (Ramírez & Ospina, 2021).

La crisis migratoria venezolana sigue siendo un desafío en Ecuador y en la región en general. La situación requiere una respuesta integral que involucre tanto a los gobiernos como a la comunidad internacional para garantizar la protección de los derechos de los migrantes, abordar las necesidades humanitarias y promover la integración y el desarrollo sostenible en los países receptores (Ramírez, 2013).

#### 4. METODOLOGÍA

La metodología utilizada en la investigación se basa en la extracción de noticias de *Ecuador Chequea* de los últimos seis años, entre 2018 y 2023, en donde aparezcan las palabras claves “venezolano”, “venezolana” o “Venezuela” y relacionados con la inmigración a Ecuador. Se han obtenido información de *Ecuador Chequea* con respecto a la inmigración venezolana desde la propia página web de *Ecuador Chequea*, de su cuenta de Twitter (X) y de su cuenta de Facebook.

Posteriormente se han revisado mediante un análisis de contenido la autoría, variables cualitativas y variables discursivas. Se ha utilizado el análisis documental para recopilar dichos datos. Nuestra investigación se desarrolló tanto bajo un enfoque cuantitativo, ya que las noticias fueron analizadas con una hoja bajo un protocolo de observación. Las variables a analizar fueron las siguientes: Autoría (si eran mensajes de usuarios individuales y no de organizaciones con

nombre y apellidos; si eran mensajes de anónimos o si los mensajes analizados eran de prensa digital). Otra variable fue la temática de los mensajes, si contenían texto, imágenes o vídeos. También hemos querido analizar las noticias bajo una perspectiva cualitativa. En otra variable estaban las observaciones (tipos de discursos, si precedía algún tipo de antecedente, cuál era el contexto o el tipo de lenguaje utilizado).

Finalmente se analizaron tan solo 50 noticias que competían con nuestro estudio entre el 1 de enero de 2018 hasta el 1 de julio de 2023. Nuestras hipótesis son:

- a) La mayor parte de las noticias que se verifican en *Ecuador Chequea* es de medios de comunicación tradicional.
- b) La mayor parte de las noticias que aparecen sobre los inmigrantes venezolanos en Ecuador que aparecen en *Ecuador Chequea* son falsas.

## 5. RESULTADOS

De acuerdo con los datos obtenidos, existe una prevalencia de noticias falsas, ya que el 70% (35 noticias) de las noticias analizadas eran falsas. Las noticias verdaderas fueron el 30% del total (15 noticias). Existe una alta prevalencia de noticias falsas con respecto a los venezolanos en Ecuador. El hecho de que el 70% de las noticias analizadas sean falsas es una cifra significativamente alta pero se entiende en el contexto de un verificador de noticias.

Esto sugiere que existe desinformación o noticias falsas en el contexto ecuatoriano, lo que puede tener graves consecuencias para la sociedad, la toma de decisiones y la confianza en algunos medios de comunicación. Un ejemplo de noticia falsa puede ser que Ecuador para el 24 de mayo de 2020 había acogido a unos 500.000 migrantes venezolanos (<https://acortar.link/nDAyfk>).

El hecho de que solo el 30% de las noticias sean verdaderas indica que las noticias veraces son una minoría en este contexto. Un ejemplo de noticia cierta ha sido la regularización de los migrantes venezolanos (<https://acortar.link/HP6H0g> del 27 de agosto de 2022). Esto resalta la necesidad de un esfuerzo adicional para verificar la información y promover la divulgación de noticias basadas en hechos verificables. Estos resultados subrayan la importancia de contar con agencias de verificación de noticias como *Ecuador Chequea*. Su labor es fundamental para identificar y desacreditar noticias falsas, ayudando así a mantener la integridad de la información y a informar al público de manera precisa y responsable. Existen consecuencias para la sociedad. Las noticias falsas pueden tener un impacto perjudicial en la sociedad al difundir información errónea y engañosa. Pueden influir en la opinión pública, distorsionar la per-

cepción de los eventos y contribuir a la polarización. Es importante abordar este problema de manera efectiva (Vázquez Luna, 2020).

La mayoría de las noticias son de autoría desconocida (50%) (25 noticias), de autores en redes sociales (30%) (15 noticias) y de prensa digital (20%) (10 noticias). La mayor proporción de noticias son con autoría desconocida representando el 50% del total. Esto significa que un número significativo de noticias en este contexto no se puede atribuir a una fuente confiable o verificable. Las noticias de autoría desconocida a menudo plantean interrogantes sobre su credibilidad y autenticidad, ya que es difícil verificar la información sin conocer la fuente. El hecho de que el 30% de las noticias provengan de autores en redes sociales es un hallazgo relevante. Esto podría indicar que las redes sociales desempeñan un papel importante en la difusión de noticias falsas. Las redes sociales a menudo permiten la rápida propagación de información sin necesidad de una verificación adecuada, lo que aumenta la exposición a noticias falsas. Aunque la proporción de noticias de prensa digital (20%) es menor en comparación con las otras categorías, es importante destacar que las noticias de este tipo también pueden ser susceptibles a la desinformación. Esto subraya la importancia de la integridad y la ética periodística en la era digital. Como ejemplo tenemos la noticia del 5 de octubre de 2018 en el que se indica en el periódico digital *La República* el titular “YouTube: ecuatorianos amenazan con quemar vivos a venezolanos” (<https://acortar.link/iDYFvR>).

La prevalencia de noticias de autoría desconocida y de autores en redes sociales plantea riesgos significativos para la confianza del público en las fuentes de información. Esto podría llevar a la desconfianza generalizada en las noticias y tener un impacto negativo en la percepción de la veracidad de la información.

Uno de los principales discursos de odio que se pueden encontrar en línea sobre la migración venezolana en Ecuador es el de culpar a los venezolanos por el aumento de la delincuencia en el país. A menudo se les acusa de ser ladrones y criminales peligrosos, sin tener en cuenta que la mayoría de los venezolanos son personas honestas y trabajadoras que han venido a Ecuador en busca de una vida mejor. Otro ejemplo común de discurso de odio en línea es el de culpar a los venezolanos por la falta de trabajo y oportunidades en el país. Se les acusa de quitar empleos a los ecuatorianos y de ser una carga para la economía del país, sin tener en cuenta que muchos venezolanos están dispuestos a trabajar en cualquier cosa para poder sobrevivir.

## 6. CONCLUSIONES

Las conclusiones del estudio indican que, al valorar 50 noticias sobre venezolanos emitidas en el verificador *Ecuador Chequea*, el 70% (35 noticias) corresponde a noticias falsas y el 30% (15 noticias) a noticias verdaderas. Los

datos muestran una alta prevalencia de noticias falsas en el contexto analizado. Esto destaca la importancia de la verificación de noticias y la promoción de noticias basadas en hechos para abordar este problema y mitigar sus efectos negativos en la sociedad. Por lo tanto se corrobora la hipótesis inicial. No ocurre lo mismo con la segunda hipótesis en la que queda descartado que la mayoría de noticias que se chequean son de medios de comunicación contrastados, al contrario. Se chequean más noticias sin autoría verificable o de personas a través de redes sociales.

Estos resultados sugieren la necesidad de mejorar la educación mediática para que las personas puedan discernir entre noticias falsas y verdaderas. Promover la alfabetización mediática puede ayudar a reducir la difusión de noticias falsas y empoderar a la sociedad para tomar decisiones informadas (Mendoza Zambrano, Marín-Gutiérrez & Rivera-Rogel, 2018).

La mayoría de las noticias son de autoría desconocida (50%), de autores en redes sociales (40%) y de prensa digital (20%). Estos resultados resaltan la importancia de contar con agencias de verificación de noticias como *Ecuador Chequea* para evaluar y validar la información de fuentes desconocidas o cuestionables. Se debe de enfatizar la necesidad de educar al público sobre cómo discernir entre noticias confiables y no confiables, especialmente en el contexto de las redes sociales.

La educación es fundamental para combatir los discursos de odio en línea. Es importante que las personas comprendan la gravedad del problema y aprendan a reconocer los diferentes tipos de discursos de odio. La educación también puede ayudar a fomentar la empatía y la comprensión hacia las personas que son objeto de estos discursos.

Otra estrategia efectiva a desarrollar es la denuncia. Las personas deben sentirse seguras al denunciar los discursos de odio que encuentran en línea. Esto puede ayudar a crear conciencia sobre el problema y presionar a los responsables de las redes sociales para que tomen medidas para detenerlos. La denuncia puede llevar a la eliminación de los discursos de odio y a la sanción de los responsables.

Al fin y al cabo, la promoción de la diversidad y la inclusión puede ser una forma poderosa de combatir los discursos de odio. Al celebrar la diversidad y promover la inclusión, se puede crear un ambiente más positivo y acogedor para todas las personas, independientemente de su origen o nacionalidad. Esto puede ayudar a reducir la incidencia de los discursos de odio y a fomentar la tolerancia y el respeto mutuo.

Se recomienda que en futuras investigaciones, se extienda el muestreo hacia otros fact-checking de otras latitudes como Perú, Colombia, Estados Unidos o España, países que contiene una tasa relevante de migrantes venezolanos.

## REFERENCIAS

- Almeida Santos, C., Teresa Peixinho, A., Lopes, F. y Araújo, R. (2023). Fact-checks: liquidez de um género. Estudo de um caso português em contexto pandémico. *Estudios sobre el Mensaje Periodístico*, 29(2). <https://doi.org/10.5209/esmp.85827>
- Altamirano, G., y Torres-Toukourmidis, Á. (2021). Análisis del discurso xenófobo hacia la migración venezolana en los comentarios de las publicaciones de Facebook pertenecientes a los diarios locales. *GIGAPPEstudios Working Papers*, 8(190-212), 310-325.
- Bahar, D., Dooley, M., y Selee, A. (2020). Inmigrantes venezolanos, crimen y percepciones falsas: Un análisis de los datos en Colombia, Perú y Chile. Migration Policy Institute. <https://policycommons.net/artifacts/3793625/inmigrantes-venezolanos-crimen-y-percepciones-falsas/4599453/>
- Burbano Alarcón, H.M., Zaldívar, A. y Vera Puebla, M. F. (2019). La política pública migratoria ecuatoriana en el caso de la crisis migratoria venezolana. *Revista Facultad de Jurisprudencia*, (6), 121-137. <https://www.redalyc.org/journal/6002/600263450012/600263450012.pdf>
- Falxa, J. (2014). Redes sociales y discursos de odio, un enfoque europeo. En Lina Mariola Díaz Cortés; Fernando Pérez Álvarez (ed. lit.). *Moderno discurso penal y nuevas tecnologías: memorias [del] III Congreso Internacional de Jóvenes Investigadores en Ciencias Penales*, 17, 18 y 19 de junio de 2013, págs. 89-106.
- Heredia, J. Y. y Battistesa, D. (2018). Nueva realidad migratoria venezolana. *REIB: Revista Electrónica Iberoamericana*, 12(1), 15-46. <https://dialnet.unirioja.es/servlet/articulo?codigo=6833189>
- Mendoza Zambrano, D.; Marín-Gutiérrez, I. y Rivera-Rogel, D. (2018). TIC y alfabetización mediática digital en la nueva ecología de la comunicación en Ecuador. En *Nova Ecologia dos Meios e Tecnologia*, Ria Editorial, 85-96.
- Morejón-Llamas, N. (2023). Características y ejes discursivos de la desinformación y el proceso de fact-checking sobre las vacunas COVID-19 en Latinoamérica. *Revista Española de Comunicación en Salud*, 47-61. <https://doi.org/10.20318/recs.2023.7005>
- Palacios, I. y Cusot, G. (2019). Las fake news y las estrategias de verificación del discurso público: Caso Ecuador Chequea. #PerDebate, 3(1), 88-107. <https://doi.org/10.18272/pd.v3i1.1558>
- Párraga Pilay, A. P. (2021). El discurso de odio sexista a través de las redes sociales como reacción al himno “Un violador en tu camino”. *Razón y Palabra*, 25(111). <https://doi.org/10.26807/rp.v25i111.1807>
- Pérez Curiel, C. y Velasco Molpeceres, A. M. (2020). Impacto del discurso político en la difusión de bulos sobre COVID-19. Influencia de la desinformación en públicos y medios. *Revista latina de comunicación social*, (78), 65-97. <https://doi.org/10.4185/RLCS-2020-1469>

- Ramírez Salado, M. (2022). Análisis lingüístico del discurso de odio en redes sociales. *Visual Review. International Visual Culture Review/Revista Internacional de Cultura Visual*, 12(1), 1-11. <https://doi.org/10.37467/revvisual.v9.3720>
- Ramírez, J. (2013). *La Política migratoria en Ecuador: rupturas, continuidades y tensiones*. IAEN. <https://www.aacademica.org/jacques.ramirez/2>
- Ramírez, J. y Ospina, M.P. (2021). La política migratoria en Colombia y Ecuador: decisiones y respuestas a la migración venezolana. En Nejamkis, Lucila, Conti, Luisa y Aksakal, Mustafa, *(Re)pensando el vínculo entre migración y crisis. Perspectivas desde América Latina y Europa*. CALAS-CLACSO. <https://www.aacademica.org/jacques.ramirez/42>
- Ripoll, S. y Navas-Alemán, L. (2018) Xenofobia y Discriminación Hacia Refugiados y Migrantes Venezolanos en Ecuador y Lecciones Aprendidas para la Promoción de la Inclusión Social, IDS and UNICEF. <https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/14130>
- Ruiz Muriel, M. C., y Álvarez Velasco, S. (2019). Excluir para proteger: la “guerra” contra la trata y el tráfico de migrantes y las nuevas lógicas de control migratorio en Ecuador. *Estudios sociológicos*, 37(111), 689-725. <https://doi.org/10.24201/es.2019v37n111.1686>
- Sánchez Moscoso, A. A., Candela, S. y Torres Toukoumidis, Á. (2022). Desinformación y migración venezolana. El caso Ecuador. *Cuadernos del Centro de Estudios en Diseño y Comunicación. Ensayos*, (161), 107-123. <https://dialnet.unirioja.es/servlet/articulo?codigo=8692392>
- Suárez-Roca, J. E., y Vélez-Bermello, G. L. (2022). Verificación de los hechos: aplicación metodológica en el medio de comunicación El Bacán. *Revista Científica Arbitrada de Investigación en Comunicación, Marketing y Empresa REICOMUNICAR*, 5(9), 163-184. <https://doi.org/10.46296/rc.v5i9.0042>
- Torres, N., y Taricco, V. (2019). Los discursos de odio como amenaza a los derechos humanos. *Cuadernos CELE*, 18. <https://mail.intgovforum.org/multilingual/sites/default/files/webform/ddo.pdf>
- Vázquez Luna, J. L. (2020). Las noticias falsas (fake news), la desinformación y la infodemia durante la pandemia de la COVID-19 Fake news, disinformation and infodemy during the COVID-19 pandemic. *SintAxis, Revista Científica del Centro de Investigación para la Comunicación Aplicada*. 189-202. <https://doi.org/10.36105/stx.2020edespCOVID-19.09>
- Vélez Bermello, G. L. (2020). Inmediatez y fact-checking: análisis del portal Ecuador Chequea. *Revista ABRA*, 40(61), 52-76. <http://dx.doi.org/10.15359/abra.40-61.3>



**IV.**  
**ESTUDIOS DE GÉNERO, DISCURSOS DE**  
**ODIO EN LÍNEA Y CIBERVIOLENCIAS**



# UNA APROXIMACIÓN AL ORIGEN DE LA MISOGINIA Y EL DISCURSO DE ODIOS EN LÍNEA HACIA LAS MUJERES<sup>1</sup>

ELENA RUIZ-ÁNGEL

*Profesora Sustituta Interina*

*Universidad de Huelva, ESEIS/COIDESO, España*

PATRICIA RUIZ-ÁNGEL

*Profesora Sustituta Interina*

*Universidad de Huelva y Universidad Pablo de Olavide, Sevilla, España*

FRANCISCO JAVIER SANTOS FERNÁNDEZ

*Investigador Predoctoral en Formación (FPU20/02848)*

*Universidad de Huelva, ESEIS/COIDESO, España*

DOI: 10.14679/2760

## **Sumario:**

1. Introducción.
  2. Aproximación al origen del discurso misógino.
  3. Conceptualización de la misoginia.
  4. El fenómeno de la manofera y la proliferación del discurso de odio en línea hacia las mujeres.
- Referencias.

## 1. INTRODUCCIÓN

En los últimos años, las redes sociales y las nuevas tecnologías de la comunicación tienen cada vez más importancia a la hora de relacionarnos y comunicarnos en la cotidianidad de las prácticas sociales, especialmente, de adolescentes y jóvenes (Gómez-Miguel y Calderón-Gómez, 2022) hasta tal punto que se han constituido como parte del accionar social global, en el que participan muchos actores (Ugalde, 2019).

---

<sup>1</sup> Este trabajo forma parte del Proyecto I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGBTI [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDERA way of making Europe”. También agradecemos el apoyo de los centros de investigación de la Universidad de Huelva, el grupo de investigación “Estudios Sociales e Intervención Social, ESEIS” y el centro “Pensamiento Contemporáneo e Innovación para el Desarrollo Social, COIDESO”.

Agradecemos también el apoyo del Ministerio de Educación y Formación Profesional, con la ayuda predoctoral para la Formación de Profesorado Universitario (FPU20/02848) que beneficia a uno de los autores.

En la sociedad digital se reproducen todo tipo de conductas sociales que tienen lugar en el mundo real como es el comportamiento misógino, aquel cargado de sentimiento lleno de rechazo y desprecio de los hombres hacia las mujeres y, en general, hacia todo lo relacionado con lo femenino (Ferrer y Bosch, 2000; Varela, 2012). El comportamiento patriarcal es un fenómeno estructural que forma parte de la sociedad desde sus orígenes, y que, por tanto, no queda relegado a manifestaciones esporádicas (Lorente, 2020). La aparición de las redes sociales, plataformas digitales, webs, foros, etc., como espacios de interacción colectiva, hace que tenga continuidad este tipo de conducta patriarcal en el entorno digital. Dando lugar a una exposición casi cotidiana de discursos de odio en línea contra las mujeres, que producen escenarios virtuales donde la confrontación y la polarización se hacen cotidianas.

El presente capítulo trata sobre la aproximación del origen del discurso misógino desde una perspectiva histórica hasta el nacimiento de la misoginia en línea que tiene lugar en los entornos virtuales y su proliferación en Internet. El machismo responde a un fenómeno socioestructural y sistémico que ha permanecido de forma transversal a lo largo de la historia fruto de la cultura patriarcal que impera en las sociedades. La misoginia se ha adaptado al medio donde opera alimentándose del discurso de odio en línea hacia las mujeres. La misoginia en línea ha proliferado de forma exponencial con el nacimiento de internet y las redes sociales, como espacio de interacción personal, información y socialización entre la población actual (Choi y Noh, 2020; Festl, 2021).

Estos espacios de interacción han visto incrementados el número de usuarios y usuarias durante y después de la pandemia del COVID-19 (Caldevilla-Domínguez et al., 2021). Estudios realizados en 2023 señalaron un crecimiento de la influencia de las redes sociales en nuestras vidas donde más de 4 740 millones de personas de todo el mundo utilizan las redes sociales, siendo un 75% mayor de 13 años. Además, indicaron que en 2022 el número de usuarios y usuarias de las redes sociales aumentó un 4,2 % donde el uso de las redes sociales aumentó una media de 2 horas y media al día en todo el mundo (We Are Social, 2022; Hootsuite, 2023).

El espacio virtual donde tiene lugar este tipo de discursos de odio en línea contra las mujeres se le denomina “manosfera”, donde se utilizan los foros, webs, blogs y redes sociales, con el objetivo de marcar un itinerario misógino y sexista (Cannito y Mercuri, 2022).

La Organización de Naciones Unidas ha manifestado la preocupación de dicha realidad social en línea. Así, su Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y expresión, en su último informe del 7 de agosto de 2023 indica que

«a pesar de la creciente preocupación por las consecuencias desproporcionadas y perjudiciales de la desinformación en línea sobre cuestiones de género, no hay acuerdo sobre lo que constituye el fenómeno ni sobre la mejor manera de abordarlo. También falta claridad sobre cómo la desinformación de género fomenta la violencia de género en línea y se ve estimulada por otras desigualdades. La ausencia de una definición consensuada y las lagunas de conocimiento ponen de manifiesto la naturaleza compleja, controvertida e intrínsecamente política del problema» (Khan, 2023, p. 3).

Para tener conocimiento de la magnitud de la realidad social de la misoginia en línea es importante tener una perspectiva histórica del nacimiento del discurso misógino y machista en nuestra sociedad, y cómo se ha adaptado a los entornos virtuales donde tiene lugar el discurso de odio en línea hacia las mujeres, espacio que se ha convertido en un torrente circulatorio de arterias y venas por donde les llegan los nutrientes misóginos a los violentos y del que todavía se tiene poca información.

## 2. APROXIMACIÓN AL ORIGEN DEL DISCURSO MISÓGINO

La gran variedad de expresiones, conceptos y nociones que engloba el discurso misógino de una forma amplia, compacta y sólida, se ha hecho común en cualquier tipo de sociedad contemporánea, tanto en la vida cotidiana como en la vida virtual.

Pero, ¿dónde se origina el discurso misógino? Es importante conocer sus raíces para aproximarnos al conocimiento de su forma de operar, para poder analizar y diagnosticar su patrón o *modus operandi* en la actualidad y, de una forma más concreta, en las redes sociales.

En el romanticismo se desarrolla gran parte del discurso misógino, entendido éste como argumento a base de rasgos generales menospreciativos contra las mujeres, supuestamente compartidos por todas, y principalmente acuñado por la tradición filosófica. El romanticismo fue un movimiento cultural reaccionario que duró más de un siglo, desde finales del siglo XVIII hasta finales del siglo XIX, ejerciendo una gran influencia a lo largo del siglo XX. Época que comprende principalmente dos momentos históricos, el primero desde la Revolución Francesa a la Revolución del 48 y el segundo con un romanticismo decadentista que alcanza hasta la Primera Guerra Mundial y la aparición del Vanguardismo. El movimiento del romanticismo defendía las raíces ancestrales, la vuelta al pasado, los rasgos diferenciales, los nacionalismos, los elementos pasionales y pre-conscientes (Valcárcel, 1993).

La filósofa Valcárcel muestra la misoginia romántica como una secularización y un pensamiento reactivo a la vez. El discurso de la superioridad de los varones frente a las mujeres vinculado a través del discurso religioso fue funda-

mentado de nuevo por el pensamiento laico e incluso ateo, donde la misoginia romántica hace suyo el lema de Napoleón: “las mujeres no tienen categoría”. Las mujeres debían permanecer fuera de la jerarquía social y política contra la práctica común del Antiguo Régimen. Aunque la Ilustración trajo consigo literatura igualitaria sobre la base de la futura sociedad democrática, lo hizo, pero con la omisión de derechos para las mujeres, una desigualdad defendida por la sociedad y por el propio movimiento romántico como algo natural. Dando lugar al pensamiento democrático patriarcal a costa de la inferioridad de las mujeres (Valcárcel, 1993).

Se les negaba así a las mujeres ser ciudadanas de primera para convertirlas en ciudadanas de segunda categoría, expulsándolas de una ciudadanía de derechos y libertades reconocidos como sí disfrutaban los varones. Además, con un ideal femenino de mujer construido, y en singular, como si de una instalación de categoría del género se tratara: lo masculino y “lo otro”, relegado a la categoría femenina, relatada sobre esa desigualdad natural, te identificaras o no con él. Pero nos interesa saber qué ocurre para que ese discurso misógino haya sido aceptado tanto por conservadores como por los sectores más progresistas en aquella época. El quid de la cuestión lo aportan la filósofa Amelia Valcárcel y la socióloga Rosa Cobos donde las reflexiones de ambas autoras, basadas en sus investigaciones científicas, apuntan que el discurso misógino es un discurso democrático (Valcárcel, 1993) basado en la igualdad (masculina) (Cobo Bedía, 2012).

Los creadores de los principios del Estado moderno democrático, el inglés John Locke y el francés Jean-Jacques Rousseau, sostuvieron un potente principio de igualdad donde ambos se detuvieron en las mujeres. Rousseau defendió un nuevo contrato social basado en la autonomía/individualismo de las personas (en masculino) y la ausencia de dominio de unos sobre otros como elementos fundamentales del estado natural, que debían dirigir el Estado, y por consecuencia, la sociedad moderna. El espacio definido y relegado para la mujer era el espacio doméstico, como espacio natural para ellas. Así lo definió la Ilustración patriarcal cuyo máximo exponente fue dicho filósofo ginebrino. Para ambos filósofos las mujeres eran el “otro sexo”. No consideraban a las mujeres ni sujetos políticos ni sujetos de razón. Convirtiéndose así, Rousseau, como unos de los fundadores del patriarcado moderno (Cobo, 2012).

El autor define en su obra *Emilio O de La Educación* el papel diferenciado por sexo que debían ejercer en la sociedad moderna futura, en la que dicha desigualdad estaba basada en la razón ilustrada, tal y como referencia su obra:

«[...] la educación de las mujeres debiera ser siempre relativa a los hombres. Complacernos, sernos útiles, hacer que las amemos y estimemos, que nos eduquen cuando somos jóvenes y nos cuiden cuando seamos viejos,

nos aconsejen, nos consuelen, para que así nuestras vidas sean fáciles y agradables; estos son los deberes de las mujeres de todos los tiempos y para lo que debieran ser enseñadas durante la infancia. Tan pronto como dejemos de recurrir a tales principios, nos alejaremos ampliamente de las normas, y todos los preceptos que se les ofrezcan no contribuirán ni a su felicidad ni a la nuestra» (Rousseau, 1983).

La escritora inglesa Mary Wollstonecraft, llevada por el espíritu de la Ilustración y heredera en clave feminista de la obra de Rousseau *Emilio O de La Educación*, en su obra *Vindicación de los derechos de la mujer*, descubre toda la construcción creada en torno al otro sexo, tal y como expone la autora:

«[...] la mujer debe ser débil y pasiva, porque posee una fuerza corporal menor que el hombre; y de ahí deduce que ha sido creada para proporcionarle placer y someterse a él, y que su deber es ofrecer a su dueño como un ser grato, siendo éste el gran fin de su existencia<sup>2</sup> [...]» (Wollstonecraft, 1998 p. 109).

No es hasta 1949, cuando la filósofa existencialista de principios del siglo XX, Simone de Beauvoir, precursora del feminismo tal y como lo conocemos hoy en día, en su obra más universal *El Segundo Sexo*, se introduce en la problematización de lo que supone ser mujer; donde reflexiona y toma conciencia de las relaciones asimétricas entre hombres y mujeres a lo largo de la historia y en todas las esferas de la vida (Beauvoir, 1949). Hasta tal punto que el hombre representaba, lo positivo y lo neutro, definitivamente representaba el género humano. La mujer, por el contrario, representaba lo negativo:

«Se mire por donde se mire, en la filosofía clásica, en los textos de las religiones o en las teorías de la ciencia moderna, la mujer se representa en negativo; el carácter de la mujer sería así naturalmente defectuoso; la mujer es un hombre fallido; la mujer es un ser relativo; el cuerpo de la mujer aparece desvalido, etcétera» (Morant, 2017, p. 106).

Como dice Valcárcel, la misoginia filosófica no es inmotivada, por cuanto que destacados teóricos del pensamiento filosófico intentaron redefinir el rol establecido a las mujeres porque las formas de vida tradicional se estaba desdiciendo por los dos elementos fundamentales que debían regir la sociedad moderna y que la Ilustración había activado, donde todos los hombres son libres e iguales y que la sociedad es un contrato (Valcárcel, 1993), tal y como defendieron con gran influencia Locke y Rousseau.

Llegados a este punto, nos planteamos cómo en pleno siglo XXI, el discurso misógino es aún aceptado en sociedades desarrolladas y avanzadas en todo el mundo. Haciendo hincapié en el discurso de odio en línea hacia las muje-

---

<sup>2</sup> *Emilio*, de Rousseau (Libro V. Cap. LXXXVII)<sup>1</sup>.

res, que está calando en la ciudadanía más joven con el fin de aglutinar más seguidores (Baena, 2022). Una realidad lejos de ser cuestionada y rebatida por muchos usuarios y usuarias en redes sociales, siendo este perfil poblacional el que mayor tiempo usa las nuevas tecnologías, y en especial, las redes sociales, como canal de información y comunicación (We Are Social, 2022; Hootsuite, 2023). El informe de EU Kids Online survey destaca que “una media de un 80 % de niños/as y jóvenes entre 9 y 16 años usan diariamente el teléfono inteligente. El 77 % de los jóvenes de entre 15 y 16 años los usan de manera diaria. La actividad más realizada por dos tercios de la población objeto es ver vídeos por medio de plataformas dedicadas a compartir estas producciones” (Smahel *et al.*, 2020, p.6).

En la actualidad vemos cómo en la sociedad en general y, de una forma más viral e inmediata en las redes sociales (Megías *et al.*, 2020), ha tenido lugar un movimiento reaccionario a los avances en materia de igualdad y diversidad, llegando a cuestionar políticas en este sentido, consolidadas en muchos países, como es el ejemplo de España. (Lorente, 2023).

Este movimiento reaccionario contra los avances en derechos de las mujeres ha venido de la mano, principalmente, del surgimiento de los grupos políticos de extrema derecha. Precisamente, en uno de los momentos álgidos en la consolidación de estos derechos, también en derechos LGTBIQ+, coincidentemente con la aparición de estas posiciones ideológicas comienza, con más fuerza, un discurso misógino en la calle, pero de una forma más intensa en las redes sociales, denominándolo discurso misógino en línea o discurso de odio en línea hacia las mujeres. Convirtiéndose, además, en un torrente circulatorio de arterias y venas por donde les llegan los nutrientes misóginos a los violentos.

Han pasado casi 3 siglos desde el nacimiento del patriarcado, y con él, de la creación de las bases del discurso misógino. El conocimiento experimental que nos aporta la perspectiva histórica, desde una mirada longitudinal y transversal, nos hace reflexionar sobre cómo el patriarcado se rearma cuando el feminismo y los derechos humanos avanzan. Hasta llegar a refundirse el machismo, adaptándose a cada tiempo y medio, para seguir manteniendo el orden establecido de las sociedades androcéntricas (Lorente, 2023). Generándose nuevas formas de misoginia en línea dependiendo de las tecnologías, la cultura, las personas y comunidades que se aprovechan de ellas (Ging y Siapera, 2018).

### 3. CONCEPTUALIZACIÓN DE LA MISOGINIA

El término misoginia está formado por la raíz griega “miso”, que significa odiar, y “gyne” cuya traducción sería mujer y se refiere al odio, rechazo, aversión y desprecio de los hombres hacia las mujeres y, en general, hacia todo lo relacionado con lo femenino. Todas aquellas creencias o expresiones emocio-

nales, psicológicas e ideológicas de odio hacia las mujeres y a lo femenino. Ese odio (sentimiento) ha tenido frecuentemente una continuidad en opiniones o creencias negativas sobre las mujeres y lo femenino, y, en conductas negativas hacia ellas (Ferrer y Bosch, 2000; Varela, 2012).

El Diccionario de la Lengua Española, define misoginia como “aversión a las mujeres” (ASALE, 2023), el diccionario de Oxford la define como «sentimiento de odio o desagrado hacia las mujeres, o un sentimiento de que las mujeres no son tan buenas como los hombres» (Diccionario de inglés Oxford, s. f.) y según el diccionario de Cambridge, el término misoginia es definido como «la falta de confianza y rechazo para con las mujeres» (Diccionario Cambridge Inglés y Tesoro gratuitos, 2023). Por lo que se evidencia que una persona misógina es aquella que no quiere a las mujeres y que además manifiesta un sentimiento de rechazo y repugnancia hacia ellas.

Para la Real Academia de la Lengua Española, el misógino es alguien que odia a las mujeres, manifiesta aversión hacia ellas o rehúye de su trato. Al acceder en él y en otros diccionarios de ámbito internacional como Cambridge Dictionary u Oxford English Dictionary, observamos cómo una realidad social tan presente en la vida cotidiana y en los medios de comunicación en los últimos años, en cuanto a hechos y/o acontecimientos sociales se refiere, no se refleja aún en tan ilustres espacios académicos. Nos referimos a la misoginia en línea u online.

La actitud misógina es una realidad social que tiene lugar tanto en la vida real (fuera de línea) como en el espacio virtual (en línea), esta última desde el nacimiento de las TICs (tecnologías de la información y de la comunicación) y, más concretamente, desde el nacimiento de lo que conocemos como la manófera o también denominada como androsfera (FundéuRAE, 2019).

Desde la historia del pensamiento feminista, se ha intentado definir o explicar el término misoginia relacionándolo con el discurso del mito del amor romántico, dándose a conocer como misoginia romántica. En los años 50, 60 y 70 del siglo XX, pensadoras feministas como Simone de Beauvoir, Kate Millett y otras, pusieron de manifiesto como el estudio de las subjetividades estaba sesgado por el discurso sexista del sistema sexo-género: los varones eran considerados superiores a las mujeres en aquellas cualidades abstractas, intelectuales y culturales, mientras que las mujeres eran consideradas aptas para la reproducción (natural), el cuidado de otros, el afecto y la responsabilidad familiar (Errázuriz Vidal, 2009).

En el presente siglo XXI, teóricas o filósofas feministas, como Celia Amorós, Amelia Valcárcel, Alda Facio o Marcela Lagarde han profundizado e investigado en sus estudios sobre misoginia. Así Celia Amorós expone cómo

«la hipertrofia esencializadora del género femenino que caracteriza el discurso de la misoginia romántica lleva hasta el límite la anulación de toda diferencia individual entre las mujeres, la negación más radical en ellas del principio de individuación. Ciertamente, las mujeres han sido siempre las idénticas -a diferencia de los varones que han constituido siempre ellos «espacios de los iguales» en el sentido, no precisamente de ámbitos igualitarios, sino de pares en tanto que varones, en tanto que miembros del genérico que a título de tal tiene el poder y en el que, por lo tanto, se vuelve importante discernir quién es quién: en este sentido la individualidad como categoría política y ontológica pertenece a los ámbitos en que se juega el poder y para ubicarse en los cuales se precisa tener información acerca del modo en que éste se distribuye y circula. Por el contrario, en el espacio de las idénticas, como espacio simbólico que se adjudica al género femenino, al ser éste un ámbito donde no están en cuestión poderes de peso, se hace irrelevante la determinación de la individualidad: no hay razón suficiente para que la atención se tome la molestia de realzarla, sino que queda subsumida en la indiscernibilidad, bajo predicaciones que afectan indistintamente a todos los ejemplares pertenecientes al genérico. Pues bien: en los discursos de la misoginia romántica esta distinción, un tanto esquemática, entre el «espacio de los iguales» y el de «las idénticas» se va a volver más contundente, y ello en relación con las motivaciones políticas reactivas a que antes hemos hecho referencia» (Amorós, 1992 p. 132).

La misoginia ha sido estudiada a lo largo de la historia, y concretamente desde el ámbito de la sociología y la psicología, en búsqueda de una explicación científica de esta conducta social. Así, psicólogas como Ferrer y Bosch concluyeron que nos referimos a ella como una actitud conectada a lo que denominó (Glick y Fiske, 1996), sexismo tradicional u hostil y sexismo benévolo. Entendiendo por sexismo hostil (viejo sexismo) como prejuicio, estereotipo y conducta discriminatoria contra las mujeres. El sexismo benévolo hace referencia a un conjunto de actitudes sexistas (considerándolas estereotipadas y limitadas a ciertos roles) más sutil, bondadoso y afectuoso hacia las mujeres que la definición de sexismo hostil (Ferrer y Bosch, 2000).

Este fenómeno social no solo tiene lugar en la vida cotidiana, es decir, en la esfera offline, sino que es un problema estructural y cultural transferido de generación en generación que impregna la conducta de muchos usuarios en los diferentes entornos digitales (Martín y López, 2021). Dando lugar a lo que conocemos como misoginia en línea. Así, la redactora especial de Naciones Unidas señala que el uso de las nuevas tecnologías va más allá de esta discriminación sistémica que sufren las mujeres y las niñas, dando lugar a una sistematización de esta conducta en los entornos digitales (Human Rights Council, 2018).

#### 4. EL FENÓMENO DE LA MANOSFERA Y LA PROLIFERACIÓN DEL DISCURSO DE ODIO EN LÍNEA HACIA LAS MUJERES

Con el nacimiento de la *sociedad red*, término acuñado por el sociólogo Manuel Castells, tiene lugar «la nueva estructura social de la Era de la Información, basada en redes de producción, poder y experiencia» (Castells, 1998, p. 350). Si nos adentramos en el dicho término según Uriarte y Acevedo, vemos como este espacio hace referencia a un

«ámbito en constante transformación que incide a la vez que se retroalimenta de cambios en percepciones, actitudes y comportamientos cotidianos, convirtiendo a ciudadanos, organizaciones y mercados en propulsores individuales de cambios colectivos tan vertiginosos como imprevisibles» (tomado de Uriarte y Acevedo, 2018, p. 35).

La aparición de las redes sociales, plataformas digitales, webs, foros, etc., como espacios de interacción colectiva, hace que tenga continuidad este tipo de conducta patriarcal en el entorno digital. Dando lugar a una exposición diaria de discursos de odio en línea contra las mujeres, que producen escenarios virtuales donde la confrontación y la polarización se hacen cotidianas.

El discurso de odio en línea hacia las mujeres es entendido como expresión de la misoginia que prolifera en el mismo espacio público que toda democracia pretende gobernar (Porchia y Lévano, 2021), teniendo un grado de coincidencia con la violencia de género, reconocido así por la O.N.U, la Unión Europea y el Consejo de Europa (United Nations, 2018).

Las distintas formas de ejercer y manifestarse en línea la violencia de género hacen pensar que hay una relación con el discurso de odio, que podría ser previo a episodios más peligrosos fuera de línea. Las mujeres aparecen como principales objetivos del discurso de odio en línea, por lo que están más expuestas de sufrir violencia durante sus experiencias como usuarias y a sufrir comportamientos perversos asociados con la violencia de género en la categoría de delito o discurso de odio (Porchia y Lévano, 2021).

El propio Comité del Consejo de Europa ha definido el discurso de odio como

«toda forma de expresión que difunda, incite, promueva o justifique el odio racial, la xenofobia, el antisemitismo u otras formas de odio basado en la intolerancia, incluyendo: la intolerancia expresada por el nacionalismo agresivo y el etnocentrismo, la discriminación y la hostilidad contra las minorías, los inmigrantes y las personas de origen inmigrante» (Comité de Ministros a los Estados miembros, 1997, p. 1).

Alertando actualmente que este discurso de odio se está difundiendo de una forma más veloz por la fácil disponibilidad que tiene la ciudadanía de comunicarse electrónicamente (Megías *et al.*, 2020).

La vida de las personas se está caracterizando cada vez más por el uso de las redes sociales ocupando un lugar casi omnipresente en sus vidas, de donde podemos extraer una razón para explicar su rápido desarrollo y crecimiento. Las redes sociales ejercen principalmente dos cometidos importantes: comunicar e intercambiar información con otros usuarios y usuarias y crear comunidades (Blasco, 2021).

Internet no es una réplica exacta del espacio público fuera de línea (*offline*), así como las tecnologías de la información y la comunicación (TICs) facilitan o agregar formas de misoginia existentes fuera de línea, pero además generan otras nuevas dependiendo del desarrollo tecnológico, las políticas algorítmicas de las plataformas digitales, las culturas administrativas que producen las nuevas tecnologías y las personas y comunidades que se aprovechan de ellas (Ging y Siapera, 2018).

El análisis del discurso de odio merece especial atención por sus implicaciones sociales, culturales y educativas, su influencia en la construcción de un clima social y político, y también por el vínculo entre el aumento del odio en línea y la perpetración de crímenes de odio (Müller & Schwarz, 2021).

La Organización de las Naciones Unidas, en su *Estrategia y Plan de Acción de las Naciones Unidas para la Lucha contra el Discurso de Odio*, define discurso de odio como

«cualquier forma de comunicación de palabra, por escrito o a través del comportamiento, que sea un ataque o utilice lenguaje peyorativo o discriminatorio en relación con una persona o un grupo sobre la base de quiénes son o, en otras palabras, en razón de su religión, origen étnico, nacionalidad, raza, color, ascendencia, género u otro factor de identidad» (ONU, 2019, p. 3),

incluyendo, de acuerdo con la Comisión Europea contra el Racismo y la Intolerancia en su recomendación 15 de 8/12/2015, formas específicas de odio como el acoso, la humillación, el menosprecio, el descrédito, la difusión de estereotipos negativos, la estigmatización y la amenaza (La Comisión Europea contra el Racismo y la Intolerancia, 2016, p. 4).

No obstante, no existe como tal una definición del discurso de odio en el marco del derecho internacional, estando el concepto ampliamente debatido y cuestionado, sobre todo en relación con la libertad de opinión y expresión, la no discriminación y la igualdad (ONU, 2023). Respecto a las características principales de dicho discurso, cabe destacar que éste (1) puede materializarse

a través de diversas manifestaciones como las imágenes, los dibujos animados, las ilustraciones, los memes, los objetos, los gestos y los símbolos, pudiendo difundirse tanto en la esfera en línea como fuera de esta, (2) tiene un carácter discriminatorio (sesgado, fanático e intolerante) o peyorativo, estando basado en prejuicios y adoptando un tono despectivo o humillante, y (3) se centra en factores de identidad reales o percibidos de un individuo o grupo, incluyendo las creencias religiosas, la etnia, la raza, el color, la ascendencia, el género, el idioma, el status socioeconómico o la clase social, la discapacidad, el estado de salud y la orientación sexual, entre otras características (ONU, 2023).

Según el *Informe sobre la Evaluación de los Delitos de Odio en España* (2021), elaborado por el Ministerio del Interior del Gobierno de España, los delitos de odio más cometidos a través de internet y las redes sociales estuvieron motivados por razón de ideología, orientación sexual e identidad de género y racismo y xenofobia, siendo estos los que mayor incidencia presentan en 2021, donde la discriminación por motivos de sexo/género ocupan el cuarto lugar. Respecto al tipo de delito cometido, las amenazas, las injurias y la promoción y/o incitación pública al odio son computadas como las conductas delictivas más repetidas, siendo internet, la telefonía/comunicaciones y las redes sociales los medios y plataformas más empleados para la ejecución de dichos delitos (López et al., 2021, p.35).

La Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y expresión de la Organización de las Naciones Unidas, señala la importancia y relevancia de

«un fenómeno social cada vez más extendido en lo que se denomina la manofera: la misoginia online, es decir, aquella misoginia que tiene lugar en los entornos virtuales. Así, en su último informe sobre la justicia de género y libertad de expresión indicó que: “la violencia de género, el discurso del odio y la desinformación se están utilizando ampliamente tanto en línea como en la vida cotidiana para suprimir las opiniones de las mujeres» (Khan, 2021, p.26), y alertó al mismo tiempo que

«el sexismo y la misoginia son factores dominantes en la censura de género que se han agudizado con el surgimiento de fuerzas populistas, autoritarias y fundamentalistas en todo el mundo» (Khan, 2021, p.6).

En el último informe del 7 de agosto de 2023, la ONU indica cómo la tecnología digital tiene una doble vertiente, una positiva y otra negativa. Por un lado, simboliza la democratización de la expresión de opinión, participación e intercambio de información de forma ilimitada y, por otro lado, multiplican los riesgos y las amenazas en línea (Khan, 2023).

La manofera, llamada en castellano “androsfera” o “machosfera”, se trata de un anglicismo formado por ‘man’+ ‘sphere’ que se refiere a un “conjunto

abierto de blogs, foros, canales de YouTube, plataformas de juegos online, etcétera, en la que los hombres debaten en la Red sobre la masculinidad” (Franco & Bernárdez, 2023). Según Charo Lacalle, existe cuatro tipos de subculturas más representativas de manosfera: MGTOW (Men Going Their Own Way), MRA (Men Rights Activists), Incels (Involuntary Celibates) y PUA (Pick up Artists), de las cuales, las comunidades más antiguas (MRA y PUA) están decreciendo frente a un fuerte crecimiento de las más radicales como son MGTOW e Incels (Horta Ribeiro et al., 2021).

Una de las comunidades más populares en la subcultura virtual por su radicalidad son los *Incels (Involuntary Celibates)*. Orquestan el discurso de odio en línea sobre violencia sexual contra las mujeres. Otras de especial relevancia son las comunidades misóginas denominadas *Forocoches* y *Burbuja.info*.

El discurso de odio ha aumentado en las redes sociales y en la manosfera durante y después del confinamiento de la pandemia del COVID-19 (El-Awaisi et al. 2020; Nemes y Kiss, 2021) siendo los temas relacionados con el género, la raza (Wong, 2022) y la religión, principalmente, los que motivaron ese discurso de odio.

## REFERENCIAS

- Amorós C. (1992). Feminismo, Ilustración y misoginia romántica. En Birulés, F. y otros. *Filosofía y género. Identidades femeninas*. Pamiela.
- Baena, M. (2022, diciembre 1). La misoginia digital coge fuerza entre los hombres jóvenes. EFEMINISTA. <https://efeminista.com/misoginia-digital-hombres-jovenes-25n/>
- Beauvoir, S. de. (2005, e.o.1949). *El segundo sexo*. Cátedra.
- Blasco Fontecilla, H. (2021). El impacto de las redes sociales en las personas y en la sociedad: Redes sociales, redil social, ¿o telaraña? *Tarbiya: Revista de investigación e innovación educativa*, 49, 97-110.
- Caldevilla-Domínguez, D., Barrientos-Báez, A., y Padilla-Castillo, G. (2021). Twitter as a Tool for Citizen Education and Sustainable Cities after COVID-19. *Sustainability*, 13(6), Article 6. <https://doi.org/10.3390/su13063514>
- Cannito, M., y Mercuri, E. (2022). Fatherhood and gender relations in the manosphere: Exploring an Italian non-resident fathers’ online forum. *European Journal of Cultural Studies*, 25(4), 1010-1029. <https://doi.org/10.1177/13675494211036967>
- Chaves-Montero, A. (2018). La utilización de una metodología mixta en investigación social. Rompiendo Barreras en la Investigación, 164-184. <https://dialnet.unirioja.es/servlet/articulo?codigo=7643236>
- Choi, D.H., y Noh, G.-Y. (2020). The influence of social media use on attitude toward suicide through psychological well-being, social isolation, and social

- support. *Information, Communication & Society*, 23(10), 1427-1443. <https://doi.org/10.1080/1369118X.2019.1574860>
- Cobo, R. (2012). Las paradojas de la igualdad en Jean-Jacques Rousseau. *Avances del Cesor*, 9, 109-121. <https://wearesocial.com/es/blog/2022/01/digital-2022/>
- Comité de Ministros a los Estados miembros. (1997, octubre 30). Recommendation No. R (97) 20 of the Committee of Ministers to member states on “hate speech”-Freedom of Expression-Www.coe.int. Freedom of Expression. [https://www.coe.int/en/web/freedom-expression/committee-of-ministers-adopted-texts/-/asset\\_publisher/aDXmrol0vvsU/content/recommendation-no-r-97-20-of-the-committee-of-ministers-to-member-states-on-hate-speech](https://www.coe.int/en/web/freedom-expression/committee-of-ministers-adopted-texts/-/asset_publisher/aDXmrol0vvsU/content/recommendation-no-r-97-20-of-the-committee-of-ministers-to-member-states-on-hate-speech)
- Diccionario Cambridge Inglés y Tesoro gratuitos. (2023, 30 de agosto). <https://dictionary.cambridge.org/es/>
- Diccionario de inglés Oxford (2023, 2 de septiembre). <https://www.oed.com/?tl=true>
- El-Awaisi, A., O’Carroll, V., Koraysh, S., Koummich, S., & Huber, M. (2020). Perceptions of who is in the healthcare team? A content analysis of social media posts during COVID-19 pandemic. *Journal of Interprofessional Care*, 34(5), 622-632. <https://doi.org/10.1080/13561820.2020.1819779>
- Ministerio del Interior de España (2021). *Informe de Evolución de los Delitos de Odio en España*. [https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/publicaciones-periodicas/informe-sobre-la-evolucion-de-los-delitos-de-odio-en-Espana/Informe\\_evolucion\\_delitos\\_odio\\_Espana\\_2021\\_126200207.pdf](https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/publicaciones-periodicas/informe-sobre-la-evolucion-de-los-delitos-de-odio-en-Espana/Informe_evolucion_delitos_odio_Espana_2021_126200207.pdf)
- Errázuriz Vidal, P. (2009). La misoginia romántica europea en la construcción de la subjetividad de la mujer moderna. Tesis Doctoral Universidad de Valladolid. <https://dialnet.unirioja.es/servlet/tesis?codigo=157660>
- Ferrer, V. A., y Bosch, E. (2000). Violencia de género y misoginia: Reflexiones psicosociales sobre un posible factor explicativo. *Papeles del psicólogo*, 75, 2.
- FundéuRAE. (2019, 6 de agosto). manosfera. <https://www.fundeu.es/consulta/manosfera/>
- Festl, R. (2021). Social media literacy & adolescent social online behavior in Germany. *Journal of Children and Media*, 15(2), 249-271. <https://doi.org/10.1080/17482798.2020.1770110>
- Franco, Y. y Bernárdez, A. (2023). Misoginia online: La cultura de la manosfera en el contexto español. Tirant Humanidades. <https://www.tirant.net/Docs/RSC Tirant.pdf>
- Ging, D., y Siapera, E. (2018). Special issue on online misogyny. *Feminist Media Studies*, 18(4), 515-524. <https://doi.org/10.1080/14680777.2018.1447345>
- Glick, P., y Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70(3), 491-512. <https://doi.org/10.1037/0022-3514.70.3.491>

- Gómez-Miguel, A y Calderón-Gómez, D. (2022) El ocio digital de la población adolescente. Madrid: Centro Reina Sofía sobre Adolescencia y Juventud. Fundación FAD Juventud. <https://www.adolescenciayjuventud.org/download/9747/>
- Hootsuite (2023, septiembre 13). Estadísticas de uso de las redes sociales. Hootsuite. <https://blog.hootsuite.com/es/125-estadisticas-de-redes-sociales/>
- Horta Ribeiro, M., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., Long, S., Greenberg, S., y Zannettou, S. (2021). *The Evolution of the Manosphere Across the Web*.
- Human Rights Council (2018). Report of the Special Rapporteur on violence against women, its causes and consequences on online violence against women and girls from a human rights perspective (en línea). <https://www.ohchr.org/en/special-procedures/sr-violence-against-women>
- Khan, I. (2021). Promoción y protección del derecho a la libertad de opinión y de expresión (A/76/258; p. 30). Naciones Unidas. Asamblea General. <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/212/19/PDF/N2121219.pdf?OpenElement>
- Khan, I. (2023). Promoción y protección del derecho a la libertad de opinión y de expresión (A/78/288; p. 28). Naciones Unidas. Asamblea General. <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N23/233/68/pdf/N2323368.pdf?OpenElement>
- La Comisión Europea contra el Racismo y la Intolerancia (ECRI). (2016, marzo 21). ECRI General Policy Recommendation No.15 on combating hate speech [Spanish translation]. <https://hudoc.ecri.coe.int/eng#%7B%22sort%22:%5B%22cripublicationdate%20descending%22%5D,%22criidentifier%22:%5B%22REC-15-2016-015-ESP%22%5D%7D>
- López, J., Sánchez, F., Fernández, T., Mániz, C. J., Herrera, D., Martínez, F., Yamir, Rubio, M., Gil, V., Santiago, A. M., Gómez, M. Á., Gómez, J., Prieto, R., Méndez, G., Amado, M. P., & Matilla, A. (2021). Informe sobre la evolución de los delitos de odio en España. <https://www.interior.gob.es/opencms/ca/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones/publicaciones-descargables/publicaciones-periodicas-anuarios-y-revistas/informe-sobre-la-evolucion-de-los-delitos-de-odio-en-espana/>
- Lorente, M (2020). *Autopsia al machismo*. Comares.
- Lorente, M. (2023). *La refundación del machismo: Poscultura y guerra cultural*. Comares.
- Martín, M. J. R., y López, Á. G. (2021). La perspectiva tecnosocial feminista como antídoto para la misoginia online. *Revista Española de Sociología*, 30(3), Article 3. <https://doi.org/10.22325/fes/res.2021.64>
- Megías, I. (coord.); Amezaga, A.; García, M.C.; Kuric, S.; Morado, R. y Orgaz, C. (2020). *Romper cadenas de odio, tejer redes de apoyo: los y las jóvenes ante los discursos de odio en la red*. Madrid: Centro Reina Sofía sobre Adolescencia y Juventud, Fad. <https://zenodo.org/records/4288486>

- Ministerio del Interior de España (2021). *Comisión Europea Contra el Racismo y la Intolerancia. (2016). Recomendación General N° 15 relativa a la lucha contra el Discurso de Odio y Memorándum explicativo*. <https://rm.coe.int/ecri-general-policy-recommendation-n-15-on-combating-hate-speech-adopt/16808b7904a>.
- Misoginia. (2023, 31 de agosto.). Inmujeres. <https://campusgenero.inmujeres.gob.mx/glosario/terminos/misoginia>
- Morant, I. (2017). El segundo sexo, de Simone de Beauvoir, y el feminismo contemporáneo. *Saitabi: revista de la Facultat de Geografia i Història*, 67, 105-134.
- Nemes, L., & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication*, 5(1), 1-15. <https://doi.org/10.1080/24751839.2020.1790793>
- Observatorio Español del Racismo y la Xenofobia (2022). Boletín de monitorización del discurso de odio en redes sociales. [https://www.inclusion.gob.es/oberaxe/ficheros/ejes/discursoodio/Boletin\\_marzo-abril2022\\_monitorizacion.pdf](https://www.inclusion.gob.es/oberaxe/ficheros/ejes/discursoodio/Boletin_marzo-abril2022_monitorizacion.pdf)
- Otzen, T., y Manterola, C. (2017). Técnicas de Muestreo sobre una Población a Estudio. *International Journal of Morphology*, 35(1), 227-232. <https://doi.org/10.4067/S0717-95022017000100037>
- Porchia, M., y Lévano G., P. F. (2021). More than words: Aspects of online misogyny. Justice, trade, security, and individual freedoms in the digital society, 399-440. <https://dialnet.unirioja.es/servlet/articulo?codigo=8119782>
- Real Academia Española (RAE), R.-, & Asociación de Academias de la Lengua Española (ASALE). (2023). Misoginia. *Diccionario de la lengua española*. <https://dle.rae.es/misoginia>
- Rousseau, J.J. (1983). *Emilio O De la educación* ([4a ed.]). Bruguera.
- Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S., y Hasebrink, U. (2020). *EU Kids Online 2020: Survey results from 19 countries*. <https://doi.org/10.21953/lse.47fdeqj01ofo>
- Ugalde, C. (2019). Redes sociales y nuevas tecnologías. *Obra digital*, 17, Article 17. <https://doi.org/10.25029/od.2019.248.17>
- United Nations (2018). *38th session of the Human Rights Council: Reports (A/HRC/38/2; p. 176)*. <https://www.ohchr.org/en/hr-bodies/hrc/regular-sessions/session38/list-reports>
- Uriarte, L. M., y Acevedo Ruiz, M. (2018). Sociedad Red y transformación digital: Hacia una evolución de la consciencia de las organizaciones. *Economía industrial*, 407, 35-49. <https://dialnet.unirioja.es/servlet/articulo?codigo=6535706>
- Valcárcel, A. (1993). *Misoginia romántica: Hegel, Schopenhauer, Kierkegaard, Nietzsche*. En Alicia H. Puleo García (Ed.). *La filosofía contemporánea desde una perspectiva no androcéntrica* (pp. 13-32). Secretaría General de Educación y Formación Profesional. <https://dialnet.unirioja.es/servlet/articulo?codigo=2063499>

- Varela, N. (2012). La nueva misoginia. *Revista europea de derechos fundamentales*, 19, 25-48.
- Varela, N. (2020). *Feminismo 4.0: La cuarta ola* (1ª reimp.). Ediciones B.
- We Are Social (2022, enero 26). *Digital 2022. We Are Social Spain*.
- Wollstonecraft, M. (1998). *Vindicación de los derechos de la mujer. Mary Wollstonecraft. Debate*.
- Wong, D. (2022). The future is ours to build: Asian American abolitionist counterstories for Black liberation. *Politics, Groups, and Identities*, 10(3), 493-502. <https://doi.org/10.1080/21565503.2021.1982737>

# ONLINE HATE SPEECH AS A FORM OF CYBERVIOLENCE: COMMUNICATION PRACTICES, SOCIAL PERCEPTION AND VULNERABILITIES IN THE EXPERIENCE OF YOUNG PEOPLE<sup>1</sup>

FRANCESCA CUBEDDU

*Research fellow*

*University of Roma Tre, Italy*

MARTA SCOCCO

*Research fellow*

*University of Macerata, Italy*

DOI: 10.14679/2761

## **Summary:**

1. Introduction.
  2. Hate speech on the web and social media.
  3. The social impacts of hate speech.
  4. Method.
  5. Online hate speech and cyberviolence: knowledge of the phenomenon and risk perception among young people.
  6. Individual and social group vulnerabilities in online forms of violence.
  7. Conclusions.
- References.

## 1. INTRODUCTION

The dissemination of digital platforms and social media has made it possible to observe a change in the concept of hate speech (Whillock & Slayden, 1995; Waldron, 2012) through which a culture of hatred of different social groups has been constructed over the years. Digital platforms have replaced real environments by becoming real social spaces beyond the virtual (Castells, 2001).

Digital and social media can not only be a potential communication resource and open space for socialising (Boyd, 2014), but also a place of risk (Livingstone & Haddon, 2009; Smahel et al., 2020), where the vulnerabilities of certain social groups are more evident and targeted. A phenomenon that

---

<sup>1</sup> The authors collectively conceived and wrote this contribution. In particular, Francesca Cubeddu is responsible for paragraphs 1, 2 and 3; Marta Scocco is responsible for paragraphs 4, 5 and 6; the paragraph 7, conclusions, were written together.

goes by the name of cyberviolence, a multidimensional construct encompassing numerous phenomena with specific characteristics that includes cyberbullying, cyberstalking and hate speech dynamics (Backe et al., 2018; Peterson & Densley, 2017; Hellsten et al., 2021).

Starting from a theoretical analysis of hate speech, the aim is to analyse its relationship with cyberviolence and the impacts it has on the social system. In the first part of the contribution, the concept of hate speech and its social implications will be explained, while in the second part this issue will be analysed using data from a national survey focusing on young people and the diverse forms of digital violence. The research carried out between 2020 and 2021 aims to explore the cyberviolence phenomenon and communication practices in a sample of 2200 young Italians, born between 1995 and 2002.

The aim is to deepen our knowledge of online hate speech within the broader contemporary phenomenon of cyberviolence. Given that in young people's experience media platforms play a major social role as factors of social influence and also generators of vulnerability, it is important to build and promote digital citizenship education (Marinelli, 2021) in which, in addition to digital skills, there is also a social awareness of the mechanisms of online violence.

The proliferation of digital platforms and social media has facilitated the transformation of the concept of hate speech (Whillock & Slayden, 1995; Waldron, 2012). These platforms have, over time, engendered a culture of animosity towards various social groups. Digital platforms have partially supplanted physical environments, evolving into genuine social spaces extending beyond the virtual realm (Castells, 2001). Digital and social media serve as not only a potential means of communication and open spaces for social interaction (Boyd, 2014) but also as potential grounds for risk (Livingstone & Haddon, 2009; Smahel et al., 2020). In these spaces, the vulnerabilities of certain social groups become more conspicuous and are specifically targeted. This phenomenon is commonly referred to as "cyberviolence," a multifaceted construct encompassing various phenomena with distinct characteristics, including cyberbullying, cyberstalking, and hate speech dynamics (Backe et al., 2018; Peterson & Densley, 2017; Hellsten et al., 2021). Beginning with a theoretical examination of hate speech, our objective is to analyse its association with cyberviolence and its ramifications for the social system. The first segment of this contribution elucidates the concept of hate speech and its societal implications. Subsequently, the issue is examined in the second part, utilising data from a nationwide survey centered on young individuals and the diverse manifestations of digital violence. The research, conducted between 2020 and 2021, seeks to investigate the phenomenon of cyberviolence and communication practices within a sample of 2200 young Italians born between

1995 and 2002. Our goal is to deepen our comprehension of online hate speech within the broader contemporary context of cyberviolence. Given that media platforms play a significant social role in the lives of young people as influencers and generators of vulnerability, it is imperative to establish and promote digital citizenship education (Marinelli, 2021). This education should encompass not only digital skills but also foster social awareness of the mechanisms underlying online violence.

## 2. HATE SPEECH ON THE WEB AND SOCIAL MEDIA

Hate speech is of increasing relevance in the contemporary social context. This term is used to define expressions that denigrate, ridicule or dehumanise individuals or groups based on specific characteristics, such as religion, gender and sexual orientation, disability, nationality, and social inequalities.

The term hate speech coined in the 1920s (Mchangama, 2022) – a socio-historical period characterised by a robust theorisation of racial and political superiority for nationalist and racist movements – to protect minorities who suffered hateful repercussions due to social and religious inequalities and skin colour. This historical period is also known for spreading the first anti-racist theories from which the lemma (Allport, 1963) took root.

In the 1980s, Critical Race Theory (Symposium, 1994; Delgado & Stefancic, 2001; Zuberi, 2011), a movement of jurists that arose to combat various forms of hate speech, spread in America. The goal is freedom of expression and identity (Özbilgin et al., 2011). Hate speech is culturally an expression of social contempt for diversity and differences (Young, 1990; Carbado & Gulati, 2003). Hate Speech brings together several elements known from sociological studies, such as: discrimination and prejudice; violence; threats; denigration and spreading falsehoods (van Dijk, 1987; Alietti & Padovan, 2005; Bentivegna & Rega, 2020; Boccia Artieri, 2021).

With social media and digital platforms, a different diffusion of hate speech can be observed. Indeed, dynamics, such as cyberviolence, show how the different existing forms of hate speech are implemented (Backe et al., 2018; Peterson & Densley, 2017; Hellsten et al., 2021). Phenomena with specific characteristics rooted in dynamics of power, domination and social inequality, where power disparities exist between social groups (Waldron, 2012). When a dominant group perceives a threat to its status or privileges, it may use hate speech to belittle and marginalise groups perceived as others (Delgado & Stefancic, 2017).

Due to its digital nature, online hate speech is characterised by certain peculiarities. The vastness and anonymity of the web can amplify the scope and

intensity of hate speech, making victims more vulnerable and attackers often unpunished (Petrilli 2019; Ferrini & Paris, 2019).

Considering it as a subset of cyberviolence, one can understand how it represents a form of digital attack that, although based on words, can have tangible and harmful consequences (Barendt, 2019).

The diffusion of hate speech online can lead to some negative repercussions, including social exclusion, psychological harm, and in some cases, violent actions in the real world. Cyberviolence encompasses a wide range of online aggressive behaviour, hate speech being its manifestation.

In 2003, with the expansion of hate messages on the Internet, the Council of Europe signed the Additional Protocol to the Budapest Convention on Cybercrime. A document that obliges member states to adopt criminal sanctions if racist or xenophobic material is found on computer systems.

In the last few years, hate speech has become more concentrated online, particularly on social media such as Facebook, YouTube and Twitter.

Numerous interventions are planned to prevent and counteract online hate speech, but still with doubtful effectiveness (Santerini, 2022). As Maria Ressa said during her Nobel Peace Prize award speech: «Online violence does not stay online. Online violence leads to real world violence».

Hate speech as defined by the Council of Europe (2014) aims at spreading, fomenting, promoting or justifying racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, including intolerance expressed in the form of aggressive nationalism and ethnocentrism, discrimination and hostility towards minorities, migrants and persons of immigrant origin.

Hate speech is one of the forms of discrimination prohibited by the European Convention for the Protection of Human Rights and Fundamental Freedoms (CEDU, 1950), Article 14 – Prohibition of Discrimination – expresses how no form of discrimination should be present or highlighted. At the European level, the Code of conduct on countering illegal hate speech online, a code of conduct promoted in May 2016 by the European Commission and aimed at web service companies – such as Facebook and Twitter – to take reports of racist content and block them out within twenty-four hours. In 2011, Italy signed the Protocol to the Budapest Convention on Cybercrime, committing states to consider the following crimes: dissemination of racist and xenophobic material when it promotes or incites violence (Art. 3); racist and xenophobic threats and insults (Art. 4 and 5); denial, gross trivialisation, approval or justification of genocide or crimes against humanity (Art. 6). In February 2020, the Working Group on Online Hate was established with the Ministry of Justice and the Publishing Department of the Prime Minister's Office by the Minister for Technological Innovation and Digitalisation. The aim, through regulations

and laws, is to counter online hate. Digital agencies and platforms need to be able to maintain freedom of expression by censoring hate speech.

The policies of Facebook, Twitter, Google and Instagram themselves limit the dissemination and treatment of certain topics and speech, but these imposed limits can be circumvented in various ways. Although there is a focus on regulating hate speech and protecting freedom of expression, social awareness and education are essential to counteract the cultural roots that provoke hate speech (Parekh, 2012). Social media has made hate speech a social problem that needs to be regulated (Hietanen & Eddebo, 2022), in fact, the current proposal of the European Commission is to introduce hate speech as a new area of crime (European Commission, 2021).

Hate in hate speech is also often associated with some subjective experience that characterises the intention of the speech act, which is often emotionally framed. The United Nations Strategy and Plan of Action on Hate Speech (United Nations, 2019) highlights that hate is always spreading, undermining peace, democracy and the human condition itself. Brudholm and Johansen (2018) note that hate and hate speech are expressions that, through strong expressions, aim to fuel terror and extremism. However, they highlight that there is no precise expression of hate speech and how crucial it is to highlight its facets.

Despite attempts to culturally counteract the basis of hate speech, according to research by the European Commission, anti-Semitic messages, hostile messages against gender orientation and the elderly have increased in recent years (European Commission, 2021).

### 3. THE SOCIAL IMPACTS OF ONLINE HATE SPEECH

According to Boccia Artieri (2021), the violent attitude is determined by the network and the type of widespread but indiscriminate access that attacks the overall structures of knowledge and coexistence.

Boccia Artieri (2021) in his article on the Digital Agenda, expresses that hate speech can be defined as a further confirmation of the negative consequences of the Internet in the current level of development and in relation to changes in the social structure.

Hate speech can have both individual and collective consequences (Whillock & Slayden, 1995; Weber, 2009). Indeed, it causes psychological damage to those affected, who often experience stress, anxiety and depression due to hate speech and discrimination. Online forms of hatred can lead to isolation and social exclusion not only of individuals but of groups, the targets of discrimination and hatred, also causing a reinforcement of negative stereotypes that cause a strong

social intolerance, brought about by a social spread of a form of hate culture (MacKinnon, 1996; Brown, 2015) that is discriminatory and racial.

In some cases, incitement to violence can lead to real and dangerous threats to people's safety (Amnesty International, 2019), which is also determined by fighting and protecting freedom of expression.

The fight against hate speech is a complex and evolving issue, with challenges related to defining, moderating and balancing freedom of expression and protecting people's rights.

With measures to counter hate speech, an attempt is being made to regulate online social hatred through various strategies, including a cultural change aimed at the correct use of online platforms. Systems that Meta itself, for example, is also trying to examine and study. The questions they are trying to answer are what systems can be used to counter online hate through content moderation policies and by banning hate speech not only in presence but also online. One trend change is the incitement of awareness-raising campaigns and educational programmes that can help prevent and counter online hate speech. Digital platforms encourage users to report offensive content, helping detect and remove hate speech, a global problem that requires international collaboration to tackle effectively. Cyberviolence as well as hate speech dynamics, as will be extensively outlined in the research and in the following section, involve different social groups involve different social groups, which prove to be more vulnerable than others in experiencing the same digital environments (Livingstone et al., 2018; Aroldi & Mascheroni, 2019).

The Council of Europe (2018, p. 5) defines cyberviolence as

«the use of computer systems to cause, facilitate, or threaten violence against individuals that results, or is likely to result, in physical, sexual, psychological, or economic harm or suffering and may include exploitation of the individual's circumstances, characteristics, or vulnerabilities».

Definition specifying how gender-based violence dynamics are also prevalent in digital and online platforms.

Countering the dynamics of violence and hatred in digital platforms entails an educated and correct social representation, a reading of identity and the dynamics of social belonging: vulnerable individuals may identify more strongly with their own groups in response to external hatred.

It is therefore relevant to promote digital citizenship education (Marinelli, 2021) based on the progressive acquisition of digital/transmedia skills that contemplate the growing complexity and diversity of the communication ecosystem. Educating not only in the use of digital platforms, but also in their role as social influencers and generators of vulnerability and social exclusion.

In this contribution, through the research presented in the following paragraphs, we examine how the concept of cyberviolence is broad and multidimensional, explicating how this phenomenon and the different forms of violence such as hate speech are known, perceived and imagined in the experience of the younger generations. In the analysis presented, the social groups that are most vulnerable and most affected by cyberviolence are shown.

#### 4. METHOD

The analysis presented in this chapter is based on a broader research project entitled 'Cyberviolence and social aspects of online violence' in which online hate speech was investigated within a broader conceptual framework relating to the multiple forms of online violence. The main aim of this study was to gather national data on cyberviolence phenomenon in order to explore better how young people understand and experience it and what maybe the most suitable actions to limit it.

From a methodological point of view, considering the main issue and the target population, a web-based questionnaire was used as the primary data collection technique for the study. The non-probability method known as 'snowball sampling' was used to develop the sample, starting from university students, association of youth people, student associations and social platform groups. The sample surveyed was limited to young people (18-25) residents in Italy. The survey was conducted between February 2021 and May 2021. A total of 2365 valid questionnaires were collected.

A majority of respondents were women (a nearly 80/20 ratio). The ages 21, 22 and 23 are heavily represented. Regarding educational attainment, most had completed high school (80.6%) and very few had higher education. The sample was well distributed throughout the nation, with the exception of the islands, which are less explored (3.3%).

The researchers developed the questionnaire to focus specifically on the cyber-violence issue and its social aspects. The questions were divided into seven thematic areas: socio-personal profile, cyberviolence, use of social networks and risks, victims and social groups, experiences, help, and interventions. Based on the data collected, this article analyses some specific issues related to cyberviolence phenomenon, new generations and hate speech practices, topics covered in the first part of the chapter.

## 5. ONLINE HATE SPEECH AND CYBERVIOLENCE: KNOWLEDGE OF THE PHENOMENON AND RISK PERCEPTION AMONG YOUNG PEOPLE

The evolution of the digital media ecosystem concerning social changes, from the advent of social media to the rise of “platform society”, represents the most recent phase of this transformation. The concept of the platform society highlights the centrality of digital platforms in people’s lives and the ability of technologies to structure daily activities (van Djick et. al., 2018).

Specifically, as observed in several research studies, they exert a considerable impact on the construction of youth’ values, imaginaries, behaviours and expectations, but also on the way they interact with their peers, relate to each other in social settings and in shaping generations (Aroldi & Colombo, 2020). All these aspects seem to be inextricably linked to the use of digital technologies: «They recreate, learn and work in a digital world, which will ultimately continue to influence their home, educational, and work environments» (Luttrell & McGrath, 2021, p. 36). An issue confirmed also by our respondents who stated using the platforms mainly for ‘chatting’ (very much = 57.5%), ‘talking’ (very much = 41.6%) and ‘finding information’ (very much = 31.8%). In digital media platforms, young people socialise, creatively construct and reconstruct their identity, but this frequent use of social media also exposes them to risks and negative phenomena, such as cyberviolence in its various forms.

Considering the multidimensional definition of cyberviolence presented in the previous paragraphs, which includes not only gender-based violence, but also the more recent manifestations of online violence, in our group of young respondents, the percentage of those who stated that they had already heard of cyberviolence is very high. The percentage, 96.8% of the total respondents do not appear to be gender sensitive. This is probably related to the fact that the new generations as “digital natives” are accustomed to using the Internet and due to personal experience, training courses and media coverage of the topic, including social media, they are somehow familiar with potential risks.

In this research the phenomenon was further explored in its different forms such as hate speech. Indeed, as the data show, the level of knowledge is very variable in relation to the specific type of online violence considered. Awareness of cyberviolence is widespread among the young people interviewed, particularly with regard to cyberbullying (99.2%), identity theft (93.4%), cyberstalking (92.9%), revenge porn (90.7%) and also hate speech (83.1%). The figure is similar for males and females. Some specific forms of cyber violence, such as vamping, cyber harassment and malicious code are instead less well known.

Knowledge, however, must not be confused with or directly linked to risk awareness. The types of acts considered the worst by respondents are revenge porn (49.3%), cyberbullying (14%) and self-harming challenges (13.45%). In this case, there are significant gender differences: boys believe cyber bullying is worse, while girls believe revenge porn and self-harming challenges are worse. At 5% hate speech is not particularly targeted by young people.

Considering the dangerous and risk aspects of these forms of online violence perceived by youth some changes and, above all, some gender-related differences emerge (Table 1).

Table 1. Perception of dangerousness (only response mode “very much”) by gender (percentage values)

	<b>Total</b>	<b>Males</b>	<b>Females</b>	<b>Delta</b>
Revenge porn	<b>79.6</b>	73.7	80.9	7.2
Self-harming challenges	<b>76.8</b>	66.1	79.3	13.2
Cyberbullying	<b>60.8</b>	44.5	64.6	<b>20.1</b>
Cyberstalking	<b>58.5</b>	40.1	62.8	<b>22.7</b>
Substitution of person/identity theft	<b>51.3</b>	51.0	51.4	0.4
Hate Speech	<b>45.0</b>	30.7	48.3	<b>17.6</b>
Flaming (denigration)	37.5	22.7	41.0	<b>18.3</b>
Sexting	36.3	20.3	40.1	<b>19.8</b>
Exclusion	33.1	21.4	35.9	14.5
Cyber-harassment	30.6	17.4	33.7	16.3
Revelation and deception	29.7	19.6	32.0	12.4
Malicious Code	22.5	14.5	24.4	9.9
Vamping (nocturnal social use)	19.1	11.8	20.8	9.0

Source: re-elaboration of data collected within the research ‘Cyberviolence and social aspects of online violence’. N = 2365.

First of all, revenge porn, self-harm challenges, cyberbullying, cyberstalking, identity theft and also hate speech are considered the most dangerous forms of cyberviolence by young people. If in general, girls tend to perceive these phenomena as more dangerous than their male peers, the gap increases in some particular cases such as cyberstalking, cyberbullying, sexting, hate speech and flaming. Therefore, hate speech is considered by young people to be among the most dangerous but not the worst forms of online violence.

In this regard, it is interesting to highlight how one of the main characteristics of digital platforms and social media, namely freedom of expression, is also considered their most dangerous aspect (Impossibility of blocking content 74.9%) and easily associated with both cyber-violence and hate speech, according to 62.1% of respondents indeed this relationship is not exclusive. As explained in the previous paragraphs, the balance between freedom of

expression and the protection of rights is an issue that has received particular attention in recent social research, but especially by policy-makers and media platforms themselves.

## 6. INDIVIDUAL AND SOCIAL GROUP VULNERABILITIES IN ONLINE FORMS OF VIOLENCE

The different forms of online violence may have direct and significant consequences on the individual and certain social groups that are more vulnerable than others. The phenomenon can therefore be observed from different perspectives. As already mentioned, online forms of hatred can lead to the isolation and social exclusion of individuals and groups, the targets of discrimination and hatred, also causing a reinforcement of negative stereotypes that lead to strong social intolerance.

The perception of risk and danger in relation to a phenomenon is also somehow related to individuals' fear of it. Among those surveyed, 46% are very afraid, 26.3% extremely afraid of becoming the victim of online violence by an anonymous and invisible attacker. This last figure shows a huge gender difference for which girls (30.6%) are much more scared than boys (8.0%) are. While so far the data analysis has concerned the imaginary, the perception and thus also the narration that the new generations provide us with about online violence, real-life contact and direct experiences seem to bring out some findings.

The experience of cyber-violence was therefore investigated considering that the phenomena may have involved the respondents as victims and indirectly.

Data show that youths were mainly victims of exclusion (18.1%), cyber bullying (14.1%), sexting (12.9%) and hate speech (12.2%). In this case the gender differences are significant: exclusion and cyber bullying concern girls in a small percentage, sexting affects girls almost twice as much as boys, while hate speech concerned boys in a higher percentage than girls. While more than half of the sample has witnessed cyberviolence (59.1%).

Considering that violence, in real life as in virtual life, is inextricably linked to stereotypes, dynamics of discrimination and power between social groups, respondents were asked to outline some characteristics that can make some individuals or social groups particularly vulnerable to cyberviolence. Although the gender variable was influential in some of the analyses discussed, it does not emerge as a discriminating factor in this case. For 63.3% of respondents, boys and girls can be victims of online violence. Furthermore, the respondents' gender does not denote any major variations in the percentages surveyed.

Age, instead, is a factor that outlines some differences. The most vulnerable groups are young people between 14 and 19 years (59.2%), followed by young people between 11 and 13 years (25.4%).

A final area explored in the research, going beyond specific individual characteristics, concerned social groups (Table 2). Among the total sample, ‘Transsexuals/bisexuals/queer/intersexuals’ (76.2%) appear to be most vulnerable to such phenomena, followed by ‘homosexuals’ (71.8%) and ‘Lesbians’ (62.7%). Sexual orientation is therefore a relevant factor determining vulnerability. In addition, cultural differences and non-compliance with physical standards can expose one more easily to cyberviolence in its many forms. Indeed, ‘migrants/foreigners’ are considered to be particularly exposed by 63.8% of respondents while ‘Overweight girls’ in 60%. Analysis of these factors is gender sensitive.

Table 2. Cyber-violence risk and social groups (only response mode “very much”) by gender of respondents (percentage values)

Social group	Total	M	F	Delta
Transsexuals/bisexuals/queer/intersexuals	<b>76.2</b>	59.0	80.2	21.2
Homosexuals	<b>71.8</b>	53.7	76.0	<b>22.3</b>
Migrants/foreigners	<b>63.8</b>	49.0	67.3	18.3
Lesbians	<b>62.7</b>	38.1	68.4	<b>30.3</b>
Overweight girls	<b>60.0</b>	40.0	64.5	<b>24.5</b>
Overweight boys	47.6	33.0	51.0	18.0
who are not sure of self-confident	20.9	13.8	22.6	8.8
who have different opinions from the masses	16.5	12.9	17.4	4.5
Nerds	14.8	8.9	16.2	7.3
who do not follow fashion	8.4	3.3	9.6	6.3
who do not use social networks	7.0	2.9	7.9	5.0

Source: re-elaboration of data collected within the research ‘Cyberviolence and social aspects of online violence’. N = 2365.

A confirmation of how the phenomenon of cyberviolence must be known and explained in its complexity, avoiding associating it with gender-based violence alone. In fact, there seem to be many different dimensions that may influence the potential vulnerabilities associated with it.

## 7. CONCLUSIONS

This contribution analysed with an empirical approach the online hate speech as part of a broader and not yet fully understood current phenomenon of cyber violence. These issues are topical considering the development

and pervasiveness digital platforms have reached, especially in the younger generations. The proposed data discussion specifically allowed us to observe these dynamics using the perspective of young people, fully involved in these processes, through their experiences and their imagination.

Youth surveyed demonstrate a general knowledge of online violence that, if explored, appears to be partial and not completely thorough since not all forms of violence are known and perceived in the same way. Although almost all participants claim to know what cyberviolence is, its complex meaning and many dimensions are not so clear.

Moreover, knowledge is not synonymous with awareness. As the data showed, not all forms of online violence are also considered particularly risky or presented as the worst.

Hate speech in particular is a demonstration of this, because despite being among the most well-known forms of online violence, it is not considered on the same level as the worst. Concerning this issue, the gender variable appears to have an important influence, which has also been highlighted regarding young people's use of different platforms.

Particularly interesting was the evidence that emerged when the young people were asked to reflect on online violence not as a current phenomenon but as an experience that may affect them or has already affected them directly or indirectly in their own experiences. In this case, the prevailing feeling is that of fear, which most respondents feel with respect to negative phenomena, beyond their actual knowledge of them. A sense of fear linked both to narratives received from external sources (for instance, awareness-raising campaigns promoted in educational contexts) but above all proven by experience, in real life. Again, changing the paradigm of reflection also changes the dimensions of the phenomenon that are more experienced than known by young people. In this regard, cyberbullying and hate speech are the forms of online violence that embody both of these aspects.

On the other hand, the vulnerability issue made it clear how young people have a real perception and a defined imaginary concerning the fact that these phenomena may affect some individuals more than others. In this case, the gender variable is not relevant while membership of particular social groups (such as migrants or the LGBTQ+ community) is.

Considering these results, the need to promote digital citizenship education in the younger generations is confirmed as necessary. Moreover, within the research the participants themselves suggested in the effective educational path a real strategy to combat these phenomena. The meaning of citizenship does not concern only young people, rather it widens the gaze and takes into account society and all the actors that together with the new generations are part of it.

Finally, starting from the sensitivities expressed in the research on vulnerabilities' topic, this citizenship is not only about digital skills or tools but rather about the protection of rights and social activism that each of us can put in place to prevent and counter risks and dangers of such digital phenomena but no less real.

## REFERENCES

- Alietti, A. & Padovan, D. (Eds.) (2005). *Metamorfosi del razzismo. Antologia di testi su distanza sociale, pregiudizio e discriminazione*. FrancoAngeli.
- Allport, G. W. (1963). Prejudice: Is it societal or personal?. *Pastoral Psychol*, 14, 33–45. <https://doi.org/10.1007/BF01768375>
- Amnesty International (2019). *Hate speech conoscerlo e contrastarlo. Guida breve per combattere i discorsi d'odio online*. Amnesty International Sezione Italiana.
- Aroldi, P., & Colombo, F. (2020). Media, generations, and the platform society. In Q. Gao & J. Zhou (Eds.), *Human aspects of IT for the aged population. Healthy and active aging. HCII 2020. Lecture notes in computer science*, vol. 12208 (pp. 567–578). Springer. [https://doi.org/10.1007/978-3-030-50249-2\\_40](https://doi.org/10.1007/978-3-030-50249-2_40)
- Aroldi, P. & Mascheroni, G. (2019). Oltre il cyberbullismo. Il contributo di eu kids online alla ricerca sul benessere sociale dei giovani utenti della rete. *Media Education*, 10(2), 181–193.
- Assimakoupoulos, S., Baider, F. H. & Millar, S. (2017). *Online Hate Speech in the European Union*. Springer. <https://doi.org/10.1007/978-3-319-72604-5>
- Bacarani, N. & Porta, L. (eds) (1999). *Il pregiudizio antisemitico, Conoscenza, comunicazione, cooperazione per rielaborare e superare i pregiudizi*. FrancoAngeli.
- Backe, E. L., Lilleston, P. & McCleary-Sills, J. (2018). Networked individuals, gendered violence: A literature review of cyberviolence. *Violence and Gender*, 5(3), 135–146. <https://doi.org/10.1089/vio.2017.0056>
- Barendt, E. (2019). What is the Harm of Hate Speech?. *Ethic Theory Moral Prac*, 22, 539–553. <https://doi.org/10.1007/s10677-019-10002-0>
- Bentivegna, S. & Rega, R. (2020). I discorsi d'odio online in una prospettiva comunicativa: un'agenda per la ricerca. *Mediascapes journal*, 16, 151-157.
- Boccia Artieri, G. (2021). Odio online, verso nuovi equilibri in rete: il rapporto della task force governativa. *Agenda Digitale online*. <https://www.agendadigitale.eu/cultura-digitale/odio-online-verso-nuovi-equilibri-in-rete-il-rapporto-della-task-force-governativa/>.
- Boyd, D. (2014). *It's Complicated. The social life of networked teens*. Yale University Press.
- Brown, A. (2015). *Hate Speech Law A Philosophical Examination*. Routledge. <https://doi.org/10.4324/9781315714899>
- Brudholm, T. & Johansen, S. B. (eds) (2018). *Hate Politics Law*. Oxford University Press. <https://doi.org/10.1093/oso/9780190465544.001.0001>
- Carbado, D.W. & Gulati, M. (2003). What Exactly is Racial Diversity?. *California Law Review*, 91, 1149–1165. <https://doi.org/10.2307/3481413>.
- Castells, M. (2001). *The Internet Galaxy: Reflections on the Internet*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199255771.001.0001>

- CEDU (1950). *Convenzione Europea dei Diritti dell'Uomo*. Consiglio di Europa.
- Chiais, M. (2007). *Menzogna. Strumento di informazione globale*, Lupetti.
- Consiglio di Europa (2014). *Guida dei diritti umani per gli utenti di internet*. Consiglio di Europa.
- Council of Europe. (2018). *Mapping study on cyberviolence with recommendations adopted by the T-CY on 9 July 2018*, Cybercrime Convention Committee (T-CY). Working Group on cyberbullying and other forms of online violence, especially against women and children (T-CY (2017)10). <https://rm.coe.int/t-cy-mapping-study-on-cyberviolence-final/1680a1307c>.
- Delgado, R. & Stefancic, J. (2017). *Critical race theory: An introduction*. NYU Press.
- European Commission. (2019). *Council of the European Union. Information note. Assessment of the Code of Conduct on Hate Speech online*. Consiglio di Europa.
- European Commission. (2021). *Communication from the Commission to the European Parliament and the Council. A more inclusive and protective Europe. Extending the list of EU crimes to hate speech and hate crime*. Consiglio di Europa.
- European Parliament. (2016). *European Parliament Resolution. On strengthening the fight against racism, xenophobia and hate crime*. Consiglio di Europa.
- Ferrini, C. & Paris, O. (2019). *I discorsi dell'odio. Razzismo e retoriche xenofobe sui social network*. Carocci.
- Harell, A. (2010). The Limits of Tolerance in Diverse Societies: Hate Speech and Political Tolerance Norms Among Youth. *Canadian Journal of Political Science / Revue Canadienne de Science Politique*, 43(2): 407-432.
- Hellsten, L. M., Crespi, I., Hendry, B. & Fermani, A. (2021). Extending the Current teorization on Cyberbullying: Importance of Including Socio-Psychological Perspectives. *Italian Journal of Sociology of Education*, 13(3), 85-110.
- Hietanen, M., Eddebo, J. (2022). Towards a Definition of Hate Speech-With a Focus on Online Contexts. *Journal of Communication Inquiry*, 0(0).
- Livingstone, S. & Haddon, L. (eds) (2009). *Kids online: Opportunities and risks for children*. Policy Press.
- Luttrell, R., & McGrath, K. (2021). *Gen Z: The superhero generation*. Rowman & Littlefield Publishers.
- MacKinnon, C. (1996). *Only Words*. Harvard University Press.
- Marinelli, A. (2021). Educare alla cittadinanza digitale nell'era della platform society. *Scuola democratica*, 12, 121-133.
- Mchangama, J. (2022). *Free Speech: A History from Socrates to Social Media*. Basic Books.
- Özbilgin, M. F., Beaugard, T. A., Tatli, A. & Bell, M. P. (2011). Work-life, diversity and intersectionality: a critical review and research agenda. *International Journal of Management Reviews*, 13(2), 177-198.
- Parekh, B. (2012). Hate speech: Is there a case for banning? *Politics, Philosophy & Economics*, 11(2), 151-175.
- Peterson, J. & Densley, J. (2017). Cyber violence: What do we know and where do we go from here?. *Aggression and Violent Behavior*, 34, 193-200.

- Petrilli, R. (eds) (2019). *Hate speech. L'odio nel discorso nel pubblico: politica, media, società*. Round Robin Editrice.
- Santerini, M. (2022). Il contrasto al discorso d'odio nella rete sociale. In Fotia L. (eds). *Discorso d'odio e politiche dell'odio tra passato e presente* (pp. 87-96). Roma Tre Press.
- Sękowska-Kozłowska, K., Baranowska, G. & Gliszczyńska-Grabias, A. (2022). Sexist Hate Speech and the International Human Rights Law: Towards Legal Recognition of the Phenomenon by the United Nations and the Council of Europe. *Int J Semiot Law*, 35, 2323–2345. <https://doi.org/10.1007/s11196-022-09884-8>
- Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S. & Hasebrink, U. (2020). *EU Kids Online 2020: Survey results from 19 countries*. London School of Economics and Political Science. <https://doi.org/10.21953/lse.47fdeqj01of0>
- Symposium (1994). Critical race theory. *California Law Review*, 82(4).
- United Nations (2019). *United Nations Strategy and Plan of Action on Hate Speech*. Nazioni Unite.
- van Dijk, T. (1987). *Communicating racism*. Sage Publication.
- van Dijk, J., Poell, T., & de Waal, M. (2018). *The platform society. Public values in a connective world*. Oxford University Press.
- Waldron, J. (2012). *The harm in hate speech*. Harvard University Press.
- Weber, A., (2009). *Manual on Hate Speech*. Counsile Europe: Strasbourg.
- Whillock, R. K. & Slayden, D. (1995). *Hate Speech*. SAGE Publications.
- Young, I.M. (1990). *Justice and the Politics of Difference*. Princeton University Press.
- Zuberi, T. (2011). Critical Race Theory of Society. *Connecticut Law Review*, 43(5), 1573–1591.



# LA CULTURA DE LA CANCELACIÓN Y SUS CONEXIONES CON LAS TEORÍAS DE LA CONSPIRACIÓN Y LOS DISCURSOS DE ODIO EN LÍNEA<sup>1</sup>

LAURA CABRERA ÁLVAREZ

*Doctoranda en Ciencias Sociales y de la Educación. Programa Investigativo  
Universidad de Huelva, ESEIS, España*

DOI: 10.14679/2762

## **Sumario:**

1. Introducción.
  2. La cultura de la cancelación y la relación con los discursos de odio y las teorías de la conspiración.
  3. El fenómeno de la cultura de la cancelación: de la práctica social de “cancelar” a sus consecuencias.
- Referencias.

## 1. INTRODUCCIÓN

La cultura de la cancelación se ha convertido en un fenómeno cada vez más habitual en nuestra sociedad (Chiou, 2020). Diversas autoras y autores manifiestan que la cultura de la cancelación está vinculada con la presente sociedad digital (Sádaba y Herrero, 2022). Internet y las TIC (Tecnologías de la Información y la Comunicación) revolucionaron el mundo, creando espacios en la red que funcionan como lugares donde las personas interaccionan, socializan y se organizan (Sola-Morales et al., 2022). Las redes sociales forman parte de esos espacios ubicados en la red y, se consideran plataformas digitales donde cualquier institución pública o privada puede crear una comunidad de internautas con objetivos e intereses comunes (Sádaba y Barranquero, 2019). Según Castells (2009) las redes sociales son una forma de “autocomunicación de masas”, un instrumento que otorga poder a quien lo usa (Martínez y Mayagoitia, 2021). Los movimientos sociales han hallado en las redes un espacio de movilización

---

<sup>1</sup> Este capítulo forma parte del Proyecto I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGBTI [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDER A way of making Europe”. También agradecemos el apoyo de los centros de investigación de la Universidad de Huelva, el grupo de investigación “Estudios Sociales e Intervención Social, ESEIS” y el centro “Pensamiento Contemporáneo e Innovación para el Desarrollo Social, COIDESO. Igualmente, se ve beneficiado por el Programa Investigativo. Mecanismo de Recuperación y Resiliencia (Expte. HU/INV/0004/2022).

donde volcar su lucha, visibilizarla e instaurar en las personas un sentido crítico del discurso público tradicional (Sádaba y Herrero, 2022). El movimiento feminista, el movimiento antirracista o el colectivo LGTB son algunos ejemplos de movimientos sociales que se han incorporado a la red y han reivindicado su lugar en la sociedad, luchando por sus derechos a través de las redes sociales (Mazón, 2021). Las plataformas digitales han incrementado la posibilidad de declarar indignación y alianza entre internautas (Norris, 2021). Un ejemplo de ello es la movilización feminista que tuvo lugar el 8 de marzo de 2018 el ‘Día Internacional de la Mujer’. Esta movilización marcó un hito en la sociedad española e impulsó la presencia del feminismo en las instituciones públicas y privadas de forma física y virtual (Caballero-Gálvez et al., 2022). Pero conforme el movimiento feminista ganó visibilidad en el ámbito digital, también ganó visibilidad el antifeminismo ya existente en la sociedad (Herrero-Diz et al., 2020).

Las redes sociales no solo se han transformado en un espacio donde poder participar activamente en cuestiones sociales e intercambiar información con otras personas de la red, como pasa con el movimiento feminista o el colectivo LGTB, sino que también se han convertido en un foco de propagación de desinformación (Ching, 2020). Esta desinformación está influida por las “noticias falsas” que se consideran verdaderas y se difunden con el propósito de estafar y falsear para el beneficio de ciertos intereses (Pérez, 2022). Bernal-Triviño y Clares-Gavilán (2019) determinan que la violencia de género, el movimiento feminista, el racismo y la inmigración son el epicentro de difusión de noticias falsas. La desinformación sobre temáticas sociales se difunde con el objetivo de minar su credibilidad y desprestigiar su lucha por la justicia social (Herrero-Diz et al., 2020).

Este capítulo que aquí presento es una revisión bibliográfica breve y exploratoria sobre las teorías de la conspiración, los discursos de odio en línea que reciben las mujeres, las personas inmigrantes y el colectivo LGTB y, cómo esos mensajes de odio se canalizan en ellas a través de la cultura de la cancelación. Esta investigación es importante por varios motivos. El primero está relacionado con la escasez de estudios que investigan acerca de las consecuencias que sufren las mujeres que reciben discursos de odio en línea (Fox et al., 2015). El segundo motivo hace alusión a que, a pesar de la escasa investigación en la materia, estudios sobre acoso online como el de Pew Research Center, revelan que las mujeres tienen un 11 % más de probabilidad de sufrir odio en internet que los hombres cuya probabilidad es del 5% (Duggan, 2017). Y el tercer motivo, que la cultura de la cancelación es un fenómeno recientemente nuevo (Vukcevic, 2022) y necesita ser investigado por su repercusión en la sociedad. Con mi trabajo me gustaría avanzar en el estudio de la cultura de la cancelación, comprender el impacto que tiene en la sociedad y conocer las consecuencias negativas que pueda llegar a provocar.

El objetivo de este capítulo es, específicamente, explorar la relación que existe entre la cultura de la cancelación, las teorías conspirativas y los discursos de odio y, por ello, en primer lugar, expondré brevemente qué son las teorías de la conspiración y su vinculación con la cultura de la cancelación. En segundo lugar, mostraré que los discursos de odio son consecuencia de la cultura de la cancelación y, por último, me adentraré en el fenómeno en sí de la cultura de la cancelación.

## 2. LA CULTURA DE LA CANCELACIÓN Y SUS CONEXIONES CON LOS DISCURSOS DE ODIO Y LAS TEORÍAS DE LA CONSPIRACIÓN

Las redes sociales son un eje central de propagación de desinformación y de teorías de conspiración (Enders et al., 2023). Las noticias falsas, la desinformación y las teorías de la conspiración navegan por las plataformas de redes sociales (Allcott y Gentzkow, 2017). Sin embargo, algunos autores y autoras indican que no se ha demostrado empíricamente que en la era digital hayan aumentado las teorías de la conspiración (Enders et al., 2023). Algunas investigaciones sugieren que el hecho de creer en teorías conspirativas está relacionado con la utilización de las redes sociales (Allington et al., 2020). En algunos estudios se ha observado que creer en determinada información está influenciado por nuestras preferencias individuales, más que por el tipo de información en sí misma (Finkel, 1993). Algunas investigaciones muestran que las personas que creen en teorías conspirativas son más propensas a indagar en este tipo de contenido en redes sociales (Cassese et al., 2020). Por lo que las personas que prefieren no creer en este tipo de ideas, pueden ver limitada la propagación de desinformación y teorías conspirativas en sus redes (Enders et al., 2023). Según Douglas et al. (2017) las personas que no comprenden determinados sucesos o no disponen de información sobre un asunto específico, son más proclives a creer en noticias falsas y teorías conspirativas. En todo caso, algunos estudios muestran que la relación entre la era digital y la creencia en teorías de la conspiración es más compleja de lo que a simple vista parece (Stroud et al., 2017).

Las redes sociales se presentan también como un espacio de debate donde profundizar sobre política, economía, movimientos sociales y temas de actualidad entre otros (Sádaba y Herrero, 2022). Según Bilewicz & Soral (2020) las nuevas generaciones utilizan las redes como medio de información principal. Sin embargo, las redes sociales se emplean más allá de la difusión de información o desinformación mencionada anteriormente, utilizándose también como un espacio donde arrojar odio hacia otras personas (Blanco-Alfonso et al., 2022). Las plataformas digitales están cada vez más presentes en nuestro día a día y esto puede ocasionar que la propagación de mensajes de odio en línea sea cada vez más una realidad (Martínez y Mayagoitia, 2021). Esto sucede porque las redes sociales facilitan la creación de perfiles falsos, originando en

las personas el sentimiento de impunidad ante determinados sucesos, así como también la seguridad de sentirte oculta tras una pantalla (Trejo, 2015). Definir el concepto de discurso de odio presenta limitaciones, ya que hay personas que justifican difundir odio como una forma de libertad de expresión (Gelber, 2002). No existe una explicación general sobre lo que significa el discurso de odio, ya que apela a cuestiones morales, sociales y legales (Gagliardone et al., 2015). No obstante, existen definiciones que reúnen las características que considero imprescindibles a la hora de definirlo. Por ejemplo, la Comisión Europea contra el Racismo y la Intolerancia del Consejo de Europa [ECRI] (2016) define el *discurso de odio* como:

“Fomento, promoción o instigación, en cualquiera de sus formas, del odio, la humillación o el menosprecio de una persona o grupo de personas, así como el acoso, descrédito, difusión de estereotipos negativos, estigmatización o amenaza con respecto a dicha persona o grupos de personas y la justificación de esas manifestaciones por razones de raza, color, ascendencia, origen nacional o étnico, edad, discapacidad, lengua, religión o creencias, sexo, género, identidad de género, orientación sexual y otras características o condiciones personales” (p. 18).

Según Gagliardone et al. (2015), los discursos de odio se consideran declaraciones que promueven actos de violencia y discriminación, así como también contribuyen a promover prejuicios hacia determinadas personas. Parekh (2006) considera que los discursos de odio se dirigen a personas que poseen determinadas características, con el objetivo de estigmatizarlas a través comportamientos que pueden considerarse inofensivos a simple vista como las bromas, los chistes, etc. El discurso de odio se utiliza como estrategia para fomentar la agenda de partidos con discursos populares (Ramírez-Plascencia et al., 2022). Según Blanco-Alfonso et al. (2022) los partidos de extrema derecha en Europa, canalizan sus discursos de odio hacia el movimiento feminista, las personas inmigrantes, racializadas y el colectivo LGTB. En España, durante la campaña de las elecciones generales de 2019, parte de la propaganda electoral que se difundió por WhatsApp fue discurso de odio dirigido a personas inmigrantes (14%), al movimiento feminista y al colectivo LGTB (10%) (Avaaz, 2019). Spring (2021) comprobó a través de un experimento social que las mujeres, junto con situaciones de racismo y homofobia, tenían más posibilidades de ser víctima de discurso de odio en redes que los hombres.

La ciudadanía que justifica los mensajes de odio apelando a la libertad de expresión sigue las mismas pautas que ocurre con la llamada “cultura de la cancelación”. La difusión de odio es apoyada en redes sociales cuando se trata de cancelar a un personaje público que ha roto una norma inaceptable (Ching, 2020). La cultura de la cancelación se presenta como un fenómeno que castiga a personas con poder y busca la justicia social (Romano, 2020). Esta práctica

forma parte de nuestra sociedad digital y aunque pueda resultar inofensiva, hay que tener en cuenta que en las redes sociales las personas siguen cuentas que les resultan afines a sus ideologías y/o pensamientos, y cuanto más convicción tengan en sus ideas, más difícil será entender ideas contrarias (Ching, 2020). Esto puede dar lugar a creer en teorías de la conspiración. Según Seide (2021) las teorías de la conspiración son más probables cuando no se conoce información sobre algún tema o se perciben contradicciones. Según Popoli y Longus (2021), cuando las personas no tienen información sobre algún tema específico, son más propensas a buscar información en fuentes no fiables.

### 3. EL FENÓMENO DE LA CULTURA DE LA CANCELACIÓN: DE LA PRÁCTICA SOCIAL DE “CANCELAR” A SUS CONSECUENCIAS

Según Norris (2021) no existe una definición consensuada sobre la cultura de la cancelación, pero una diversidad de autoras y autores la definen. Para Saint-Louis (2021) la cultura de la cancelación es un fenómeno empleado en redes sociales para justificar el boicot a una figura pública debido a sus acciones o expresiones consideradas reprochables. Ng (2020) la define como un fenómeno donde las personas unen sus voces discriminadas y manifiestan su desacuerdo a un personaje público con poder. A la cultura de la cancelación también se le denominó “Culture Call-out”, impulsada por mujeres con el movimiento #MeToo que consistió en denunciar públicamente a sus abusadores sexuales. Esta cultura también fue impulsada por la comunidad negra a través de la etiqueta #BlackLivesMatter (Sádaba, 2021). Según Vukcevic (2020), la cultura de la cancelación busca defender a una parte de la sociedad relegada a la otredad que durante mucho tiempo ha estado silenciada, como es el caso de las mujeres o de la comunidad negra. La cancelación surge con el objetivo de señalar hechos injustos que son llevados a cabo por personas con influencia y poder (Romano, 2020). Saint-Louis (2021) diferencia entre “cancelar” y “cultura de la cancelación”. Yar & Bromwich (2019) define la cancelación como la acción de eliminar a alguien del espectro público por haber hecho o dicho algo considerado inaceptable y, mediante la vergüenza pública, el boicot o la humillación, se le exigen consecuencias o responsabilidades ante el suceso. Y la cultura de la cancelación es definida igualmente como el fenómeno que impulsa a cancelar (Saint-Louis, 2021).

Según algunas investigaciones, la cultura de la cancelación se lleva a cabo de la siguiente forma. En primer lugar, señala y comparte a través de redes sociales el comportamiento de alguien o de alguna institución por considerarlo inaceptable, ocasionándole humillación pública. En segundo lugar, las personas que promueven la cultura de la cancelación incitan al boicot de la persona que ha mostrado el comportamiento inaceptable, generándole con

ello presión social para que se disculpe o se retire de redes sociales por su conducta inaceptable. En tercer lugar, cuando ya han conseguido hacer viral el acto inadmisibles, promueven que la persona objetivo sufra algún tipo de castigo permanente, como por ejemplo perder seguidores en redes sociales. También crean etiquetas virales para promover la cancelación (Tandoc et al., 2022). En definitiva, la cultura de la cancelación opera a través de la denuncia, el acoso, la deshumanización y la vergüenza entre otras acciones (Tandoc et al., 2022). El objetivo final de la cancelación es que la persona se vuelva invisible para la sociedad (Zhang & Laiu, 2020).

La cultura de la cancelación crea debates en torno a la libertad de expresión. Algunas personas consideran que la inclinación a la cancelación está minando la libertad de expresión e incrementando la autocensura entre personajes públicos (Luu, 2020). Disponer de libertad de expresión es la base sobre la que debe sustentarse una sociedad democrática (Yar y Bromwich, 2019). Sin embargo, la libertad de expresión es limitada, basta con vulnerar cualquier derecho humano básico para que la libertad de expresión quede bajo sospecha de violencia (Vukcevic, 2022).

Natalie Pang, perteneciente a la Universidad Nacional de Singapur, considera que la cultura de la cancelación ha hecho posible que el discurso de la ciudadanía que se encuentra en situación marginal se escuche con más frecuencia y, no esté influenciado por personas en situaciones de poder (Lim, 2020). No obstante, algunas personas sostienen que la cultura de la cancelación tiene como objetivo subestimar a personas de ideologías derechistas (Fox News, 2020). Sin embargo, se ha demostrado a través de casos como el de James Gunn que esa afirmación es errónea. El director de cine fue despedido por haber escrito tiempo atrás una serie de tuits que no agradaron demasiado a personas de ideología conservadora (Knoop, 2018). Existen también algunos ejemplos de personajes públicos que han sido acosados durante estos años. Una de ellas es J.K. Rowling, la autora de los libros de Harry Potter. En el año 2020, Rowling tuiteó sobre la existencia del sexo biológico y su importancia tanto para las personas homosexuales porque la atracción que sienten es hacia personas de su mismo sexo, como para las mujeres por sufrir discriminación en base a su sexo. Esto ofendió al colectivo internacional transgénero y, como consecuencia, originó la cancelación de J.K. Rowling por algunos de sus fans, como, por ejemplo, suspendiendo relación con la autora en sitios dedicados a Harry Potter como Mugglenet (Luu, 2020).

No obstante, la cultura de la cancelación puede resultar perjudicial para las personas que la sufren por incitar a la otredad y a la marginación (Saint-Louis, 2021). Algunas autoras y autores reflejan las posibles consecuencias que puede ocasionar la cultura de la cancelación. Según Vukcevic (2022), las consecuencias de la cancelación pueden ser, entre otras, de índole económica y/o laboral. Así,

las personas que ejercen la cancelación suelen exigir el despido laboral de la persona a la que están boicoteando. También incitan a las marcas a abandonar cualquier tipo de colaboración con la persona cancelada y, motivan a sus seguidores a que la dejen de seguir. Otra de las consecuencias es la pérdida de relaciones sociales que, posteriormente, conduce al aislamiento. La persona cancelada es expuesta públicamente. Esto hace que las personas que quieren brindar apoyo no se atrevan por miedo al linchamiento público, por lo que ya no es solo cancelada la persona que ha hecho o dicho algo inadecuado, sino que cualquier persona que le brinde apoyo también lo será, porque se considerará que opina lo mismo que la persona a la que han cancelado. Otra consecuencia son los problemas psicológicos que provoca el hecho de haber sido objeto de cancelación, como la ansiedad que provoca el haber perdido seguidores o el apoyo de ciertas marcas.

En definitiva, la cultura de la cancelación ha emergido exponencialmente en los últimos años en el contexto del desarrollo tecnológico descrito previamente, como un mecanismo que se conecta a veces con la expresión del odio y la divulgación de teorías conspiratorias que pueden provocar como resultado la cancelación. Por ello, para comprender en mayor profundidad las implicaciones de este fenómeno, resulta imprescindible seguir investigando sobre la cultura de la cancelación y sus conexiones con las redes sociales, las teorías de la conspiración y los discursos de odio, un camino que apenas ha comenzado.

## REFERENCIAS

- Allcott, H. & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2). 211-236. 10.1257/jep.31.2.211
- Allington, D., Duffy, B., Wessely, S., Dhavan, N. & Rubin, J. (2020). Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychological Medicine*, 51(10). 1763-1769. <https://doi.org/10.1017/S003329172000224X>
- Avaaz. (2019). *WhatsApp Social Media's dark web. How the messaging service is being flooded with lies and hate speech ahead of the Spanish elections*. [https://avaazimages.avaaz.org/Avaaz\\_SpanishWhatsApp\\_FINAL.pdf](https://avaazimages.avaaz.org/Avaaz_SpanishWhatsApp_FINAL.pdf)
- Bernal-Triviño, A. & Clares-Gavilán, J. (2019). Uso del móvil y las redes sociales como canales de verificación de fake news. El caso de Maldita.es. *Profesional de la información*, 28(3). <https://doi.org/10.3145/epi.2019.may.12>
- Bilewicz, M. & Soral, W. (2020). Hate speech epidemic. The dynamic effects of derogatory language on intergroup relations and political radicalization. *Political Psychology*, 41(1), 3-33. <https://doi.org/10.1111/pops.12670>
- Blanco-Alfonso, I., Rodríguez-Fernández, L. y Arce-García, S. (2022). Polarización y discurso de odio con sesgo de género asociado a la política: análisis de las interacciones en Twitter. *Revista de Comunicación*, 21(2), 33-50. <https://doi.org/10.26441/RC21.2-2022-A2>

- Caballero-Gálvez, A. A., Zaera, A., Tortajada, I. y Willem, C. (2022). Fe-MI-nismo en las redes sociales: riesgos y oportunidades para el feminismo contemporáneo. *Atlánticas Revista Internacional de Estudios Feministas*, 7(1), 62-87. <https://doi.org/10.17979/arief.2022.7.1.7049>
- Cassese, E. C., Farhart, C. E. & Miller, J. M. (2020). Gender differences in COVID-19 conspiracy theory beliefs. *Politics & Gender*, 16(4), 1009-1018. <https://doi.org/10.1017/S1743923X20000409>
- Castells, M. (2009). *Comunicación poder*. Alianza Editorial.
- Ching, J. (2020). You are cancelled: Virtual collective and the emergence of cancel cultures as ideological purging. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 12(5), 1-7. <https://doi.org/10.21659/rupkatha.v12n5.rioc1s21n2>
- Chiou, R. (2020). We need deeper understanding about the neurocognitive mechanisms of moral righteousness in an Era of online vigilantism and cancel culture. *AJOB Neuroscience*, 11(4), 297-299. <https://doi.org/10.1080/21507740.2020.1830872>
- Comisión Europea contra el Racismo y la Intolerancia del Consejo de Europa [ECRI] (2016). *Recomendación general N° 15 relativa a la lucha contra el discurso de odio y memorandum explicativo. Adoptada el 8 de noviembre de 2015*. <https://rm.coe.int/ecri-general-policy-recommendation-n-15-on-combating-hate-speech-adopt/16808b7904>
- Douglas, K. M., Sutton, R. M. & Cichocka, A. (2017). The psychology of conspiracy theories. *Current Directions in Psychological Science*, 26(6), 538-542. <https://doi.org/10.1177/0963721417718261>
- Duggan, M. (2017). Online Harassment 2017. *Pew Research Center*. <https://www.pewresearch.org/internet/2017/07/11/online-harassment-2017/>
- Enders, A. M., Uscinski, J. E., Seeligs, M. I., Klofstad, C. A., Wuchty, S., Funchion, J. R., Nurthi, M. N., Premarantne, K. & Stoler J. (2023). The relationship between social media use and beliefs in conspiracy theories and misinformation. *Political Behavior*, 45, 781-804. <https://doi.org/10.1007/s11109-021-09734-6>
- Finkel, S. E. (1993). Reexamining the “minimal effects” model in recent presidential campaigns. *The Journal of Politics*, 55(1), 1-21. <https://doi.org/10.2307/2132225>
- Fox, J., Cruz, C. y Lee, J.Y. (2015). Perpetuating online sexism offline: Anonymity, interactivity and the effects of sexist hashtags on social media. *Computers in Human Behavior*, 52, 436-442. <https://doi.org/10.1016/j.chb.2015.06.024>
- Fox News Night (2020, 10 de junio). Cancel culture takes aim at movies, statues, and outspoken public figures. *Fox News*. <https://www.foxnews.com/video/6163227649001#sp=show-clips>
- Gagliardone, I., Gal, D., Alves, T. & Martínez, G. (2015). *Countering Online Hate Speech. Programme in Comparative Media Law and Policy*. University of Oxford.
- Gelber, K. (2002). *Speaking back: The free speech versus hate speech debate*. John Benjamins Publishing Company.

- Herrero-Diz, P., Pérez-Escolar, M. y Plaza, J. F. (2020). Desinformación de género: análisis de los bulos de Maldito Feminismo. *Icono 14*, 18(2), 188-216. <https://icono14.net/ojs/index.php/icono14/article/view/1509>
- Knoop, J. (18 de julio). Disney appeased the alt-right by firing James Gunn- here's why that's a big problema. *Daily dot*. <https://www.dailydot.com/parsec/james-gunn-firing-alt-right/>
- Lim, K. (2020). Cancelar la cultura: cómo la 'brigada despierta' de Asia se convirtió en un político de fuerza. *Correo de la mañana del sur de china*.
- Luu, C. (2020). El fandom de Harry Potter canceló oficialmente a JK Rowling en K. Cho (ed.), *Respondiendo al racismo en el campus: analizando el activismo estudiantil y respuestas institucionales*. Nayar.
- Martínez, L. y Mayagoitia, A. M. (2021). Influencers y mensajes de odio: jóvenes y consumo de contenidos autocensurados. *Revista Prisma Social*, (34), 4-39.
- Mazón, A. (2021). Mujeres en redes de lucha: ciberfeminismo como movimiento social contemporáneo. *Estudios Políticos*, (53), 31-43. <https://doi.org/10.22201/fcpys.24484903e.2021.53.79429>
- Norris, P. (2021). Cancel Culture: Myth or Reality? *Political Studies*, 71(1), 145-174.
- Ng, E. (2020). No grand pronouncements here: Reflections on cancel culture and digital media participation. *Article reuse guidelines*, 21(6), 621-627. <https://doi.org/10.1177/1527476420918828>
- 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>
- Pérez, M. (2022). El discurso antifeminista en el tiempo de la posverdad. *Dilemata* (38), 147-162.
- Popoli, G. & Longus, A. (2021). Gender differences and the five facets of conspiracy theory. *International Journal of Psychological Studies*, 13(3). <https://doi.org/10.5539/ijps.v13n3p64>
- Ramírez-Plascencia, D., Alonzo-González, R. M. & Ochoa-Amezquita, A. (2022). Odio, polarización social y clase media en Las Mañaneras de López Obrador. *Revista Interdisciplinaria de Estudios de Comunicación y Ciencias Sociales*, (35), 83-96. <https://doi.org/10.31921/doxacom.n35a1505>
- Romano, A. (2020, 25 de agosto). Why we can't stop fighting about cancel culture? *Vox*. <https://www.vox.com/culture/2019/12/30/20879720/what-is-cancel-culture-explainedhistory-debate>
- Saint-Louis, H. (2021). Understanding cancel culture: Normative and unequal sanctioning. *First Monday*, 26(7). <https://doi.org/10.5210/fm.v26i7.10891>.
- Sádaba, I. & Barranquero, A. (2019). Las redes sociales del ciberfeminismo en España: identidad y repertorios de acción. *Athenea Digital*, 19(1), 1-24.
- Sádaba, T. (2021). Framing as modus operandi for social movements: The case of Black Lives Mattes. *Review of Nationalities*, 11(1). <https://doi.org/10.2478/pn-2021-0001>
- Sádaba, T & Herrero, M. (2022). Cancel Culture in the Academia: The hispanic perspective. *Methados revista de ciencias sociales*, 10(2), 312-321. <https://doi.org/10.17502/mrcs.v10i2.594>

- Sola-Morales, S., Hernández-Conde, M., Arencón-Beltrán, S. & Sierra, F. (2022). Mitos e imaginarios del activismo digital feminista. Análisis de memes de la cibercampaña #FuckGenderRoles. *Teknokultura. Revista de Cultura Digital y Movimientos Sociales*, 19(1), 43-54. <http://dx.doi.org/10.5209/TEKN.76690>
- Spring, M. (2021). La periodista de la BBC brutalmente troleada en redes sociales (y las impactantes revelaciones de su experimento). *BBC.com*. <https://bbc.in/3qfMqfN>
- Stroud, N. J., Thorson, E. y Young, D. (2017). Dar sentido a la información y juzgar su credibilidad. Comprender y abordar el ecosistema de desinformación. 45-50.
- Tandoc, E. C., Ru, B. T., Huei, G. L., Charlyn, N., Chua, R. A. & Goh, Z. H. (2022). #CancelCulture: Examining definitions and motivations. *New Media & Society*, 1-19. <https://doi.org/10.1177/14614448221077977>
- Trejo, R. (2015). *Las redes frente al espejo. En mensajes de odio y discriminación en las redes sociales*. Consejo Nacional para Prevenir la Discriminación, 37-48.
- Vukcevic, M. (2022). Conceptualization of the culture of cancellation. *Multidisciplinary Research Journal*, (20), 63-76.
- Yar, S. & Bromwich, J. E. (2019). Cuentos de la cultura de cancelación adolescentes. *El New York Times*.
- Zhang, J. & Laiu, D. (2020, 20 de agosto). Cancelling & cancel culture: Are they relevant to us in S'pore & why should we care? *Mothership*. <https://mothership.sg/2020/08/cancel-culture-explainer/>

**V.**  
**NARRATIVAS, DISCURSOS**  
**E IMAGINARIOS SOCIALES**  
**SOBRE PERSONAS LGTBIQ+**



# BLOGS LESBIANOS: UNA EXPLORACIÓN DE SUS USOS Y CONTENIDO POR EL ACTIVISMO LGTBIQ+ A MEDIADOS DE LOS DOS 2000S EN ARGENTINA

AILÍN CANILLO

*Becaria Doctoral INHUS - CONICET - CEHIS  
Universidad Nacional de Mar del Plata, Argentina*

**DOI:** 10.14679/2763

## **Sumario:**

1. Introducción.
2. Antecedentes.
3. Metodología.
4. Resultados y análisis.
5. Discusión y conclusiones.
6. Bibliografía.

## 1. INTRODUCCIÓN

El siguiente trabajo desarrolla una exploración inicial de una selección de blogs creados y activos desde principios y mediados de los 2000s que se definen como blogs de lesbianas, lesbianas feministas, feministas lesbianas, o de activismo LGTBIQ+ lésbico en Argentina. Los blogs son: Mujeres a Bordo (<https://mujeresabordo.blogspot.com/>) creado en el año 2004, Lesbianas Fugitivas (<https://lesbianasfugitivas.blogspot.com/>) creado en el 2005, Baruyera entre Meses (<https://baruyoaldia.blogspot.com/>) del año 2005, Ultra Violetas (<https://mendozamujeresgenero.blogspot.com/>) del año 2007 y Malas como las Arañas (<https://malascomolasa.blogspot.com/>) del año 2008. . El contenido de texto de los blogs se obtuvo a partir de webscraping con Python y se visualizó con las herramientas de la aplicación web Voyant Tools (<https://voyant-tools.org/>) para realizar la lectura y análisis de corpus de texto. Combinando una lectura distante y cercana de los mismos se identifican las temáticas y contenidos que se tratan en estas páginas webs para clasificaciones iniciales que combinan perspectivas nativas y analíticas.

Por un lado, este ejercicio resulta interesante para conocer sobre las prácticas digitales del activismo LGTBIQ+ en la blogósfera, tema poco explorado en el ámbito académico en Argentina. Proponemos pensar los blogs como dispositivos de visibilidad y memoria locales de la disidencia lesbiana (Cano, 2017). Creemos que este análisis contribuirá a responder interrogantes en dos líneas ¿Qué tipos de reclamos se articularon en el activismo de estos blogs? ¿Puede

brindarnos esta mirada el criterio para una clasificación de esta blogosfera? ¿Quién era el contradestinatario en las narrativas políticas que se articulaban en sus posts? En otro plano, nos parece relevante aportar visibilidad al activismo lésbico y conocer su propia historia.

## 2. ANTECEDENTES

Nos servimos de los aportes teóricos de feministas lesbianas como Monique Wittig (1992) y Adriane Rich (1996) para pensar a la heterosexualidad como discurso y régimen político, más que una orientación sexual, que produce un sistema de diferencias y otredades. Para las autoras los discursos de la heterosexualidad oprimen en la medida que no permiten la posibilidad de hablar (o existir) si no es en sus propios términos. El carácter obligatorio se sintetiza en la frase «tú-serás-heterosexual-o-no-serás». De esta manera la heterosexualidad se presenta como un núcleo de “naturaleza” que se resiste al examen de las ciencias sociales y se formula como verdad universal mediante leyes generales para todas las sociedades, en todas las épocas y en todos los individuos. Judith Butler (2007) habla de Matriz Heterosexual para pensarlo como un régimen de poder discursivo hegemónico cuyas categorías fundadoras (hombre y mujer) son normativas y excluyentes. Todo cuerpo que quede por fuera de estas categorías, todas las prácticas y deseos sexuales que escapen a la lógica heterosexual caen por fuera de la matriz. Estar por fuera implica ser rechazados, excluidos y patologizados.

En este sentido la homosexualidad ha sido patologizada por la medicina y psicología (Preciado, 2008) como prohibida y perseguida por los estados nación a partir de legislaciones utilizadas para arrestar a homosexuales, personas trans y trabajadoras sexuales (Perlonger, 2016; Rubin, 1989). Esta falta de reconocimiento institucionalizada, siguiendo a Fraser (2000), en términos generales ha impedido la participación en igualdad de condiciones de la vida social y el acceso a la ciudadanía de las personas que forman parte de los colectivos LGBTIQ+.

El activismo LGBTIQ+ ha luchado históricamente contra la discriminación, patologización y la criminalización de la homosexualidad de diversas maneras según el momento histórico y espacio geográfico. Teniendo en cuenta los debates planteados por Monaco y Zuviría (2017) para fines analíticos podemos clasificar en tres tipos a las estrategias políticas del activismo LGBTIQ+. Las primeras, asociadas a las luchas por la ampliación de derechos civiles y las demandas al Estado. Las segundas, más radicales, que buscan la transformación política, tras consignas como la abolición del hetero-patriarcado y el fin del capitalismo, y otras que están vinculadas a procesos de despatologización y visibilización en diversos medios para disputar otros sentidos en la búsqueda por el reconocimiento. Esta búsqueda de reconocimiento, plantean los autores, también puede darse al interior del movimiento LGBTIQ+ (como en el feminista)

como los cuestionamientos a sus lógicas jerárquicas y la inclusión-exclusión de sujetos diversos y de agendas de demanda. Simultáneamente, es necesario para los autores no asumir como opuestas estas estrategias, para poder reconocer la heterogeneidad dentro del activismo y, en este caso específico, del activismo lésbico con sus particularidades.

Siguiendo a Adriane Rich (1996), la heterosexualidad obligatoria tiende a percibir la experiencia lesbiana, en particular, desde lo desviado hasta lo abominable. Sin embargo, uno de sus efectos más significativos es la invisibilización de la existencia lesbiana en la sociedad. En el contexto del activismo LGTBIQ+ en Argentina, uno de los hitos pioneros en la politización y visibilización del lesbianismo fueron los Cuadernos de Existencia Lesbiana, publicados en Buenos Aires entre 1988 y 1996. Los Cuadernos utilizaron las tecnologías de comunicación disponibles de la época para enfrentar discursivamente el imaginario social heteronormativo y homofóbico que prevalecía en la sociedad argentina. Así, contribuyeron a combatir la invisibilidad y la negación social del deseo sexo-afectivo entre mujeres (Vespucci, 2015).

En este sentido, actualmente las tecnologías y particularmente internet han tenido un lugar central para la difusión de talleres, conocimientos, vivencias, jornadas, reuniones y convocatorias a marchas. Monaco y Zuviria (2017) resaltan el lugar de las tecnologías debido a la frecuente dificultad de generar lugares de encuentro en torno a ciertas prácticas (“patologizadas”) en el espacio público. La extensión de internet durante los 2000s junto con la web 2.0 favoreció la posibilidad de establecer lazos en medios virtuales y generó un aumento significativo de usuarios en blogs, páginas webs y chats que permitieron dar cuenta de la potencialidad de ciertas herramientas tecnológicas (Van Dijck, 2016).

Al mismo tiempo, esta potencialidad de las tecnologías de internet para el activismo se ven debilitadas por comentarios, blogs, foros o comunidades virtuales que circulan y promueven mensajes contra la comunidad LGTBIQ+. Si bien son pocos los estudios realizados sobre esta temática en Argentina como en América Latina estos se encuentran centrados en el análisis crítico del discurso de este tipo de comentarios en blogs argentinos (Pérez-Riedel, 2018) o en blogs personales en contra de la comunidad LGTBIQ+ en Brasil (de Moraes Carvalho y Lionço, 2018). En cuanto a estudios sobre las producciones de contenido de blogs lesbianos encontramos los aportes de Fanese (2014) quien da cuenta de las tensiones entre la necesidad de crear imágenes para identificarse por parte del activismo y el reconocimiento del carácter opresivo del concepto de identidad. Y en particular de las tácticas de resistencia que los blogs involucran para interpelar la heteronormatividad y plantear representaciones de género divergentes. En otras latitudes encontramos trabajos similares que estudian el contenido de blogs LGTBIQ+ en países como Polonia y Letonia atendiendo a las contraestrategias empleadas por los grupos disidentes para desafiar estas representaciones hegemónicas de género como la heterónoma (Chojnicka, 2015).

En consonancia con la perspectiva teórica de Eliseo Verón (1987), concebimos las producciones online de los activismos LGBTIQ+ en blogs como discursos políticos de naturaleza significativa. Dentro de esta estructura discursiva, la enunciación política se entrelaza inextricablemente con la construcción de dos categorías esenciales: el “adversario” u “otro negativo” y el “otro positivo” hacia el cual se dirige. En este contexto, resulta esencial analizar detenidamente cómo se forjan y se delinearán estos “contradestinatarios” en los discursos que se manifiestan en los blogs.

Asimismo, autores como Vir Cano (2017) proponen que estas formas de producción de sentido, centrándose en el blog-archivo Potencia Tortillera, son un ejercicio de resistencia, modos de disputar los sentidos del mundo y de hacer comunidad. De ahí su importancia como la tarea que encuentra Cano en repensar los usos y apropiaciones disidentes de las técnicas de la representación y sus insumos tecnológicos.

En esta misma línea, explorar los usos de las tecnologías por parte del activismo lésbico en Argentina forma parte de conocer la propia historia del mismo. Para estos estudios, destrucción de registros, recuerdos y cartas que documentan las realidades de la existencia lesbiana debe tomarse con mucha seriedad, como un medio de conservar la heterosexualidad obligatoria (Rich, 1996). De esta manera se vuelve importante poder conocer y reconstruir los debates y discursos del activismo LGBTIQ+ en internet como parte de esos registros. Particularmente explorar blogs de mediados de los 2000s nos permite pensar los vínculos con internet y como contra que ideas negativas, estereotipos o prejuicios sobre la comunidad LGTBIQ+ han luchado o resistido mediante producciones de contenido en la web para desafiar estas representaciones hegemónicas tradicionales.

### 3. METODOLOGÍA

El siguiente trabajo adoptó un diseño exploratorio del contenido de cinco blogs de activismo lésbico de Argentina creados a principios y mediados de los 2000s. La extracción de datos y la construcción del corpus consistió en “raspar” el contenido de texto de los blogs seleccionados, técnica conocida también como web scraping, con Python utilizado las librerías BeautifulSoup, requests, re e itertools. De esta manera obtuvimos de cada blog una carpeta que en su interior contenía las publicaciones ordenadas por mes y año de actividad. Al mismo tiempo el texto obtenido fue limpiado de signos de puntuación como de stopwords (artículos, pronombres, preposiciones o aquellas palabras que individualmente carecen de sentido) para ser interpretable a partir de las herramientas digitales. Asimismo, empleamos la técnica de lematización para normalizar términos particulares y sus derivados en el corpus de cada blog

explorado. Para facilitar la identificación de patrones y tendencias en el contenido, ya que elimina las diferencias de flexión y conjugación de las palabras.

Para el procesamiento y análisis del contenido, me he servido de la aplicación web Voyant Tools (<https://voyant-tools.org/>) para realizar la lectura y análisis de corpus de texto. Con esta herramienta, realice un análisis de lectura distante de los textos presentes en los blogs seleccionados. La lectura distante, según la perspectiva de Franco Moretti (2015), implica un enfoque que va más allá de la lectura tradicional y se centra en patrones, tendencias y relaciones en un conjunto de textos amplio. Voyant Tools nos permitió identificar patrones de palabras clave, frecuencia de términos específicos y relaciones textuales dentro del corpus. A través de esta exploración digital, obtuvimos una visión más amplia de los temas recurrentes, las narrativas predominantes y las preocupaciones centrales en los blogs lesbianos. Dentro de la variedad de herramientas que nos ofrece Voyant Tools he utilizado Tendencias, Cirrus, Términos, Enlaces y Contexto para aproximarme al contenido del corpus de cada blog y realizar clasificaciones iniciales.

En un principio visualizar las palabras de esta manera me permitió dar cuenta de los principales ejes de discusión dentro de estos blogs sin la necesidad de una lectura cercana. En este sentido, me parecía interesante incluir estas búsquedas palabras asociadas a demandas del activismo LGBTIQ+ y feminista. En particular la del matrimonio ya que en Argentina la Ley del matrimonio igualitario fue sancionada en 2010, momento en el que algunos de estos blogs estaban activos. Lo mismo fue para el término aborto, que ha sido dentro del activismo feminista argentino una de las demandas al estado más fuertes en los últimos veinte años (Tarducci y Rifkin, 2010). Al mismo tiempo, a partir de las nubes de palabras, explore la presencia o ausencia de algunos términos frecuentes a lo largo de las publicaciones mensuales de los blogs en comparación con los demás.

Para explorar los términos que más se repetían en el corpus también utilice la herramienta Tendencias que muestra las frecuencias de los términos en los documentos que componen al corpus en gráficos de línea y barra apilada. En nuestro caso esos documentos eran los meses junto con el año en lo que hubo publicaciones en los blogs. Estos gráficos muestran la variación en los valores de los términos seleccionados, como al mismo tiempo que representan puntos en el gráfico como subgrupos. El eje Y muestra la frecuencia relativa de los términos en el corpus en porcentajes, es decir la frecuencia del término seleccionado en relación con todos los términos del documento. Mientras que el eje X contiene los documentos por meses y año con las publicaciones del blog.

A su vez, utilice la herramienta Enlaces para explorar la relación entre determinados términos del corpus. Esta herramienta nos permite visualizar en un gráfico de red los términos por proximidad. Las palabras clave, que

seleccionamos, se muestran en azul y las palabras que tienen mayor proximidad dentro del corpus están en naranja. También me serví de la herramienta “Palabras Clave en Contexto” la cual muestra cada ocurrencia de la palabra o término seleccionado con un poco de texto circundante. Esta función resultó útil para comprender de qué manera se utilizaron estos términos en contextos particulares dentro de nuestro análisis. Pudiendo hacer una lectura cercana desde la herramienta Lector, que proporciona una vista del texto del corpus, cuando fue necesario. Además, utilicé la herramienta Líneas de Burbuja que visualiza la repetición y la frecuencia del uso de un término en el corpus. Cada documento del corpus está representado por una línea horizontal y dividido en segmentos de igual longitud. Los términos son representados por burbujas. El tamaño de estas indica la frecuencia en el diámetro de texto correspondiente. La longitud del diámetro de la burbuja indica la mayor frecuencia del término.

#### 4. RESULTADOS Y ANÁLISIS

Comenzaremos con el blog Mujeres a Bordo (<https://mujeresabordo.blogspot.com/>) creado en el 2004 y activo hasta el año 2011. Este blog surge como iniciativa de la Red Informativa de Mujeres de Argentina (RIMA) para crear un nuevo canal de debate en la web para las feministas por fuera de las listas de correo electrónico. Esta herramienta fue utilizada por RIMA para la transmisión de información y el intercambio entre feministas, periodistas, organizaciones del movimiento de mujeres, investigadoras, estudiantes, y activistas del país y también de países de Latinoamérica. Si bien la organización ya poseía una página web desde el año 2000, en 2004 suma también su dirección en Blogger para socializar información, como materiales de producción propia sobre feminismo y género.

Para comenzar a explorar el contenido del presente blog visualice en una nube de palabras<sup>1</sup> aquellos términos que tenían mayor frecuencia en el blog. A partir de visualizar las palabras más utilizadas en el cuerpo de los posts del blog Mujeres a Bordo en una nube n°1 con la herramienta cirrus del Voyant Tools vemos que las palabras que más se utilizan son ‘aborto’ (252), ‘derechos’ (201), ‘salud’ (186), ‘violencia’ (152), le siguen en menor tamaño, que expresa una

---

<sup>1</sup> Las nubes de palabras se han mostrado especialmente útiles para visualizar campos semánticos (Khusro et al., 2018). En la nube de descriptores semánticos se observan los tamaños de letra proporcionales importancia numérica (frecuencia) de cada palabra (Hernández Fernández, 2015). Los colores y la distribución se dieron de manera aleatoria. En las nubes de palabras se eliminaron los términos mujer y lesbianas/lesbiana ya que eran los que mayor frecuencia poseían y opacaban al resto de los términos que pretendíamos explorar. A mayor tamaño de las palabras en la nube mayor frecuencia posee en el corpus y a menor tamaño menor frecuencia. Por otro lado, eliminamos del corpus las palabras “no”, “mas”, “mujeres” que eran las que mayor frecuencia tenían, como podemos ver en la nube N°1.

menor frecuencia, ‘contra’ (141), ‘derecho’ (140), ‘sexual’ (138) ‘argentina’ (126) ‘feministas’ (119) y continuados por genero (112), ‘todas’ (109), ‘ley’ (108).

Al indagar en la herramienta Términos, en menor medida aparece ‘lesbianas’, unas 79 veces, y sumado sus variaciones contamos 116 frente a ‘mujeres’ que aparece más del doble. Al explorar el contexto con de este término con la herramienta Palabras Clave en Contexto y Lector vemos que aparece asociado a la militancia por la visibilidad lésbica como convocatoria a jornadas y talleres, a escritos de pensadoras e intelectuales lesbianas como Adriane Rich y Val Flores como sobre salud sexual lesbiana, convocatoria al Encuentro Lésbico Feminista Latinoamericano Caribe realizado en el país en 2007. Al mismo tiempo se utiliza el termino lesbiana/s como un término que diversifica al sujeto mujeres junto con otros como indígenas, bisexuales, negras y trans. A partir del 2010 el termino se vincula al crimen de odio a Natalia Gaitán<sup>2</sup> en la provincia de Córdoba y se suman publicaciones que denuncian la discriminación y los estigmas que sufren las mujeres lesbianas. Las últimas publicaciones que contienen este término hacen referencia al lanzamiento del archivo documental digitalizado del activismo lésbico en Argentina en el blog Potencia Tortillera.

Figura 1. Nube de palabras n°1

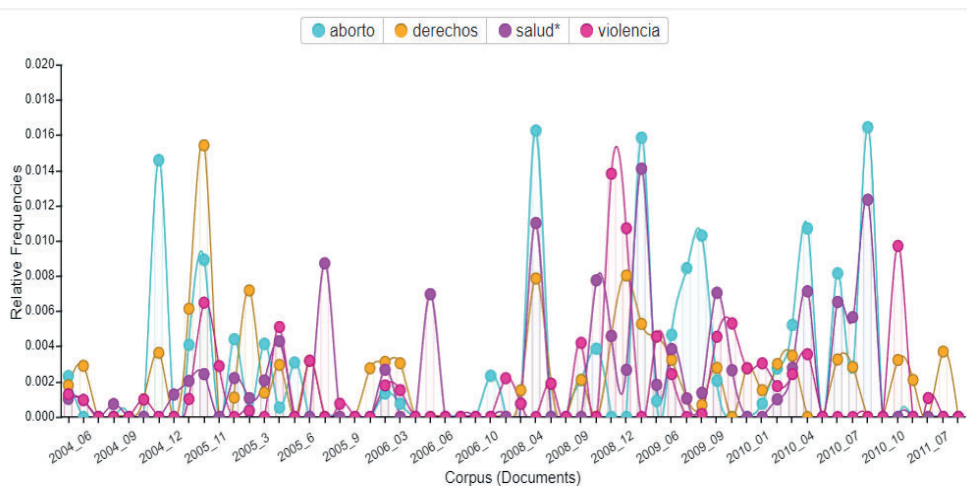


Fuente: Elaboración propia con Voyant Tools y Python.

<sup>2</sup> <https://agenciapresentes.org/2022/03/07/crimen-de-pepa-gaitan-el-dia-que-la-lesbofobia-fue-un-arma-asesina/>

Podemos ver que en las palabras con más frecuencia están vinculadas con demandas de la agenda feminista argentina de comienzos de los 2000s (Tarducci y Rifkin, 2010; Burton, 2013) como ‘aborto’, ‘ley’, ‘derechos’, ‘salud’, ‘violencia’. Cuando exploramos a partir de un gráfico de frecuencias la distribución de estos términos observamos que a lo largo de los todos los meses (y años) de actividad del blog una gran parte de los posts contenían este tipo de palabras. Destaca el término ‘aborto’ como aquel que tiene mayores picos de frecuencia y continuidad a lo largo del tiempo en relación del resto de los términos. El cual suele encontrarse acompañado de los términos ‘derechos’ y ‘salud’, pero en menor medida.

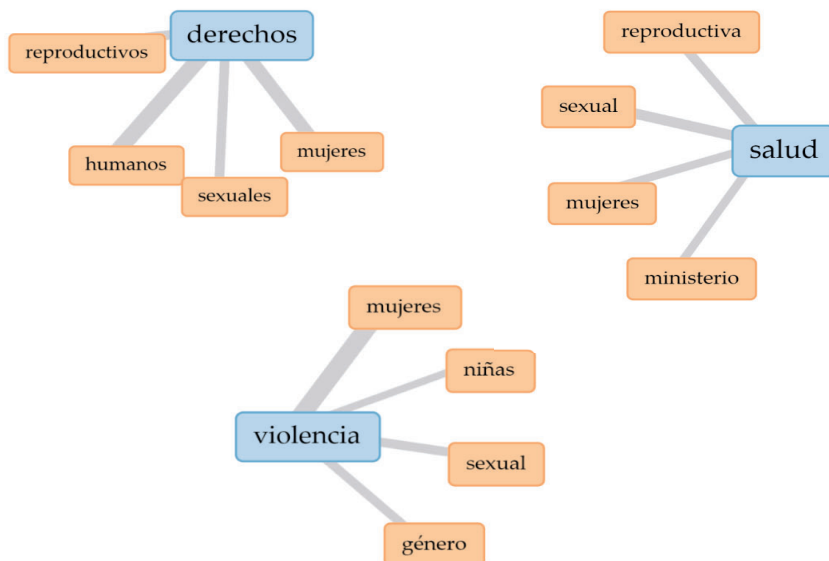
Figura 2. Gráfico de frecuencias n°1



Fuente: Elaboración propia con Voyant Tools.

Siguiendo a estas palabras con mayor frecuencia, explore los términos que se encontraban más cercanos a las palabras 'derechos', 'salud', y 'violencia' con la herramienta Enlaces para poder entender a qué hacían referencia dentro del contenido del blog. Vemos que 'derechos' y 'salud' están asociados a palabras similares como 'reproductivos', 'reproductiva', 'sexuales', 'sexual' y comparten el término 'mujeres'. Derechos se vincula a 'humanos', haciendo referencia a derechos humanos y 'salud' a 'ministerio'. Por otro lado, 'violencia' hace referencia a sujetos como 'mujeres' y 'niñas' y a 'sexual'. A partir de esta visualización podríamos decir que los grandes temas de este blog se vinculan con los derechos sexuales y reproductivos de las mujeres, entendidos también como parte de la salud reproductiva y como un derecho humano. Y también con la temática sobre las violencias, en particular de género, a mujeres y niñas. En este sentido, podemos decir que son demandas dirigidas al estado como a sus instituciones como lo reflejan los términos ‘ministerio’ y ‘derechos’.

Figura 3. Gráficos de enlaces n°1



Fuente: Elaboración propia Voyant Tools.

Con respecto a los términos que presentan baja frecuencia, cuando buscamos heterosexualidad en la herramienta Términos aparece 3 veces. En cuanto a matrimonio solo aparece 10 veces en el documento, vinculado a una convocatoria a una marcha por la ley de unión civil en la Ciudad Autónoma de Buenos Aires, capital de Argentina. Otras veces asociado a las demandas por la ley de matrimonio igualitario de manera crítica. Ya que esta no es la única demanda de las personas de la comunidad LGBTIQ+:

«el matrimonio no arregla los prejuicios». Y a los “discursos progresistas que no pueden nombrar la palabra lesbiana” (gabrielaa, 13 de marzo de 2010) los entienden como parte de la lesbofobia de la sociedad. Sin embargo, reconocen la importancia de la ley “pero nos ayuda a vivir” (gabrielaa, 25 de junio de 2010) .

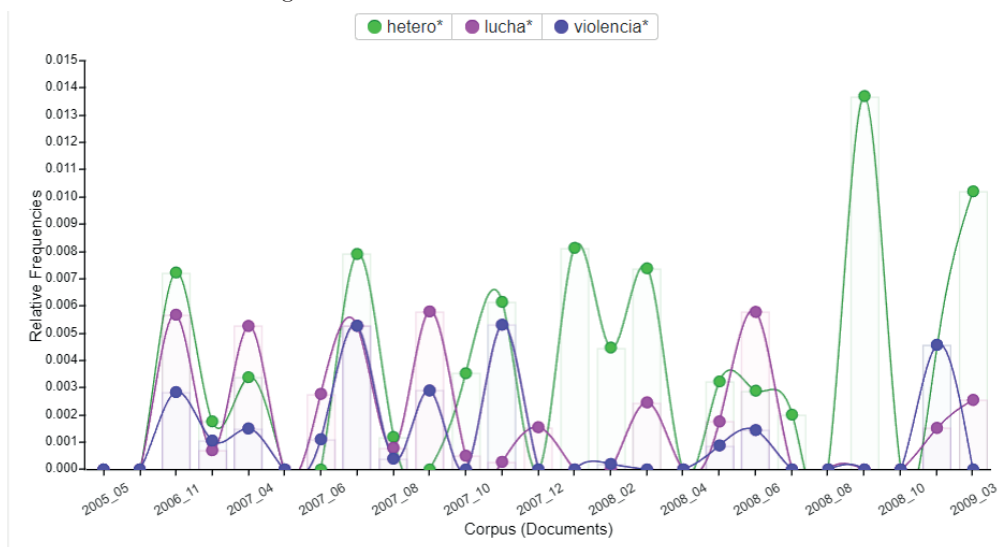
También aparece como parte de los temas tratados durante el Encuentro Nacional de Mujeres en 2009.

El blog Lesbianas Fugitivas (<https://lesbianasfugitivas.blogspot.com/>) comienza con post en el 2005 hasta el año 2009, es un blog de activistas lesbianas feministas de Neuquén, de la Patagonia Argentina. Para visualizar las frecuencias de las palabras eliminamos los términos que más se repetían como mujeres y lesbianas ya que competían con el resto de los términos. Las palabras más utilizadas fueron ‘política’ (113), ‘cuerpos’ (97), seguida de ‘día’ (83), ‘contra’ (76) ‘sexual’



Asimismo, elabore un gráfico de frecuencias relativas con las palabras violencia (50 veces aparece en total) y hetero que reunía todas las variaciones del término y el termino lucha (que se repite 45 veces y sumándole términos que derivan de esta palabra es de 68) para explorar su relación con los otros términos.

Figura 6. Gráfico de frecuencias relativas n°2



Fuente: Elaboración propia con Voyant Tools.

Al examinar la presencia de estos términos en el corpus utilizando la herramienta Línea Burbuja, y al agregar el término 'contra', observamos que estas palabras son recurrentes a lo largo de las publicaciones del blog. Notamos que su uso se superpone durante ciertos meses, lo que sugiere una asociación entre ellos en el contenido del blog. Se puede notar que la lucha contra la violencia y la resistencia a la heterosexualidad son temas frecuentes en las publicaciones. Además, al introducir el término 'lesbianas' en el análisis, observamos que también se superpone en algunos contextos con los otros términos explorados, en particular cuando 'hetero\*' y 'contra' o 'violencia' y 'contra' aparecen juntos.

Si bien estas palabras no son de las más utilizadas en el blog, nos arrojan información sobre contra que es lo que luchan en su producción de contenido y activismo el blog Lesbianas Fugitivas. En este sentido, podemos pensar que la construcción de contradestinatarios (Verón, 1987) en este caso está construida a partir del término 'hetero\*'.

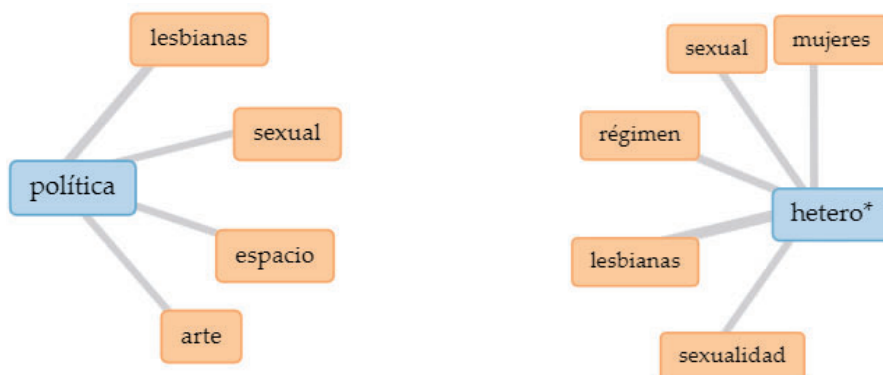
Figura 7. Gráfico de línea de burbuja n°1



Fuente: Elaboración propia con Voyant Tools.

Referencias: en azul el término contra, en verde hetero, en rosa el término violencias y en violeta el término lesbianas.

Figura 8. Gráfico de enlaces n°1

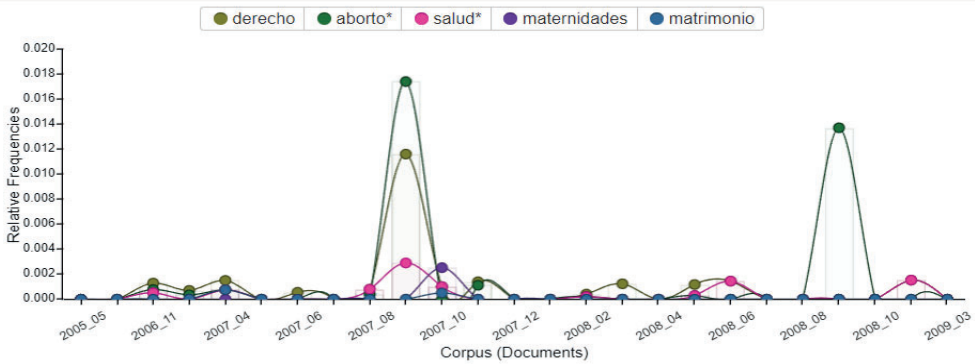


Fuente: Elaboración propia con Voyant Tools.

En este sentido, explore los términos que se encontraban más cercanos a las palabras con mayor frecuencia en el corpus, que son ‘Política’ y ‘Hetero\*’. ‘Política’ se asocia con términos como ‘sexual’ y ‘lesbianas’, lo que sugiere una vinculación con la temática de la identidad y la visibilidad. Además, encontramos términos como ‘espacio’ y ‘arte’ relacionados con ‘política’. Por otro lado, en el caso de ‘Hetero\*’, los términos ‘sexual’ y ‘sexualidad’ hacen referencia a la heterosexualidad y los conceptos relacionados. Además, el término ‘régimen’ se vincula con una serie de normas que parecen estar asociadas con la idea de heterosexualidad.

Con respecto a los términos que “faltan” como ‘salud’ (13), ‘matrimonio’ (4), ‘maternidad(es)’ (24) ‘derechos’ (32) y ‘aborto’ (18) si están presentes en el corpus, pero con una frecuencia muy baja en relación a los términos que más se repiten. Contrastando estas palabras con las de la nube podemos decir que las primeras, y por ende el contenido del blog, están vinculadas más con una política de la identidad. Es decir, a ciertas estrategias del activismo LGBTQI+ asociadas con la transformación de un sistema político vinculado con la heterosexualidad y con la visibilización identitaria. Más que realizar reclamos por derechos al Estado como pueden indicar los términos salud, aborto, matrimonio y derechos. Diferencia que podemos notar con el blog Mujeres Abordo, donde derechos, salud y aborto si eran palabras con mayor frecuencia en el corpus.

Figura 9. Gráfico de frecuencias relativas n°4



Fuente: Elaboración propia con Voyant Tools.

El blog Baruyera entre Meses (<https://baruyoaldia.blogspot.com/>), activo desde el año 2005 al 2013, es el blog de una revista homónima de lesbianas feministas de la ciudad de Buenos Aires con el fin de divulgar pensamiento disidente. Las palabras que más frecuencia tienen dentro del blog son ‘Baruyera’ (93) que hace referencia a la revista, ‘aborto’ (92), ‘cuerpo’ (89), ‘otra’ (88), ‘mundo’ (88), ‘feminista’ (88), ‘cuerpos’ (85), ‘personas’ (84), ‘genero’ (83), ‘sexual’ (82). Seguido por ‘vida’ (72), ‘política’ (61), ‘derechos’ (58), ‘sexualidad’ (54). Y con tamaños más pequeños ‘identidad’ (51), ‘derecho’ (50), ‘deseo’ (49), ‘heterosexual’ (45). Podemos notar que este blog comparte términos con Mujeres Abordo como lo son ‘aborto’ y ‘derecho/s’, pero también comparte con Lesbianas Fugitivas con ‘sexual’ y ‘cuerpos’.

Figura 10. Nube de palabras n°3



Fuente: Elaboración propia Voyant Tools.

De manera similar a lo observado en el blog Lesbianas Fugitivas, cuando sumamos todas las variaciones del término 'heterosexual' agrupándolas bajo 'hetero', el número total asciende a 181, mientras que, en la nube de palabras, este término ocupa un lugar menos importante.

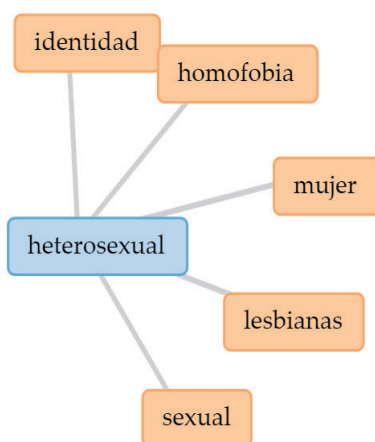
Figura 11. Imagen n°2

hetero
hetero* (181)
heterosexual (45)
heterosexualidad (39)
heterosexuales (20)
heteropatriarcal (12)
hetero (12)
heterosexista (6)
heterosexismo (5)
heteronormatividad (5)
heteronormativo (4)

Fuente: Captura del Voyant Tools.

Al explorar las palabras asociadas al término 'heterosexual', observamos nuevamente la presencia de los términos 'mujeres' y 'lesbianas'. Sin embargo, es interesante notar que también aparece el término 'homofobia' en relación con 'heterosexual'.

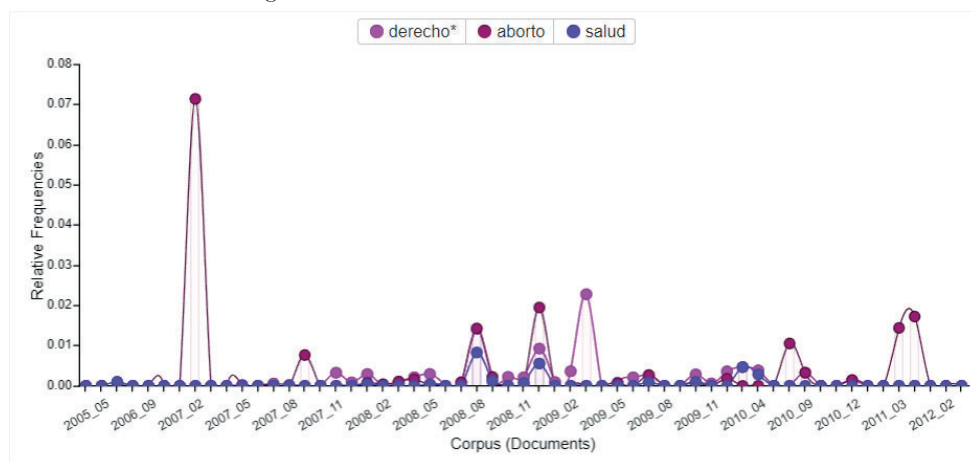
Figura 12. Gráfico de enlaces n°2



Fuente: Elaboración propia con Voyant Tools.

En relación a los términos que tienen una baja frecuencia en el corpus, nuevamente encontramos que ‘matrimonio’ (11), ‘salud’ (28) y ‘maternidad’ (19) aparecen con poca frecuencia, similar a lo observado en el blog *Lesbianas Fugitivas*. Sin embargo, en contraste con el blog anterior, los términos ‘derechos’ y ‘aborto’ se presentan con mayor frecuencia en este corpus. Al visualizar estos términos en un gráfico de frecuencias relativas, notamos que su uso es limitado en varios meses. Destaca especialmente la aparición de la palabra ‘aborto’ en los posts de febrero de 2007, pero no se observan aumentos significativos en otros momentos.

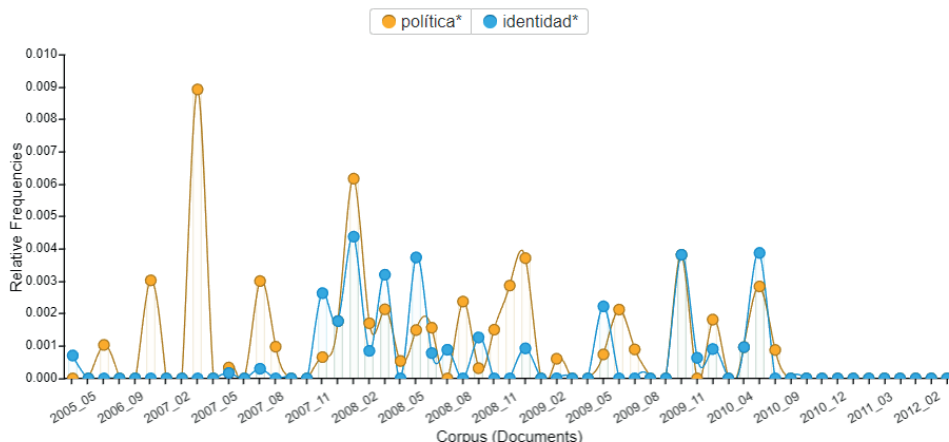
Figura 13. Gráfico de frecuencias relativas n°5



Fuente: Elaboración propia con Voyant Tools.

En este sentido, busqué la distribución en los documentos de términos vinculados con lo que llamamos estrategias de transformación política o de reconocimiento (Del Monaco y Zuviría, 2017) según su frecuencia. En particular, nos centramos en los términos ‘política’ e ‘identidad’. Observamos que estos términos coinciden en su uso en la mayoría de los meses de actividad del blog, y presentan frecuencias similares desde 2008 hasta 2010. En síntesis, a partir de los términos más frecuentes, encontramos demandas asociadas a derechos, pero también vinculadas con la temática de la heterosexualidad como parte de las violencias hacia las identidades lésbicas.

Figura 14. Gráfico de frecuencias relativas n°6



Fuente: Elaboración propia con Voyant Tools.

El blog UltraVioletas (<https://mendozamujeresgenero.blogspot.com/>), funcionó desde el año 2007 al 2009, se define a sí mismo en su presentación como una "colectiva de mujeres feministas convocadas con el objetivo de crear y sostener un espacio de reflexión e intervención en torno a la crítica de la heterosexualidad obligatoria" (ultravioletas, septiembre de 2007) .

Figura 15. Nube de palabras n°4

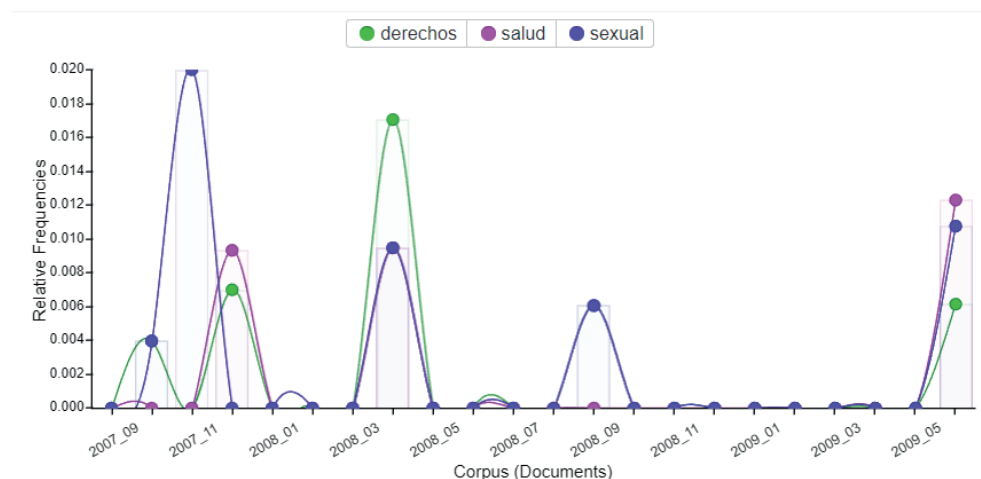


Fuente: Elaboracion propia con Voyant Tools.

En este caso la palabra con mayor frecuencia es ‘derechos’ (18), seguida por ‘salud’ (17), ‘sexual’ (14), ‘contra’ (14), ‘día’ (14), ‘sexuales’ (14), seguido de ‘contra’ (11) y ‘nuestro’ (12) y en menor medida ‘hetero’ (8).

Continuando con la exploración, visualice los tres términos más frecuentemente utilizados en relación a su aparición a lo largo de los meses de actividad del blog. Notamos que estos términos ocupan un lugar central en el contenido de los meses de actividad del blog, y con frecuencia comparten el mismo mes de publicación. Además, es interesante destacar que estos tres términos también se encuentran con mayor frecuencia en otros blogs, como Mujeres a Bordo, donde comparten el término ‘aborto’, así como en el blog Baruyera entre Meses, donde se utiliza el término ‘salud’. Asimismo, se observa que el término ‘sexual’ es compartido con el blog Lesbianas Fugitivas siendo uno de los más frecuentes.

Figura 16. Gráfico de frecuencias relativas n°7



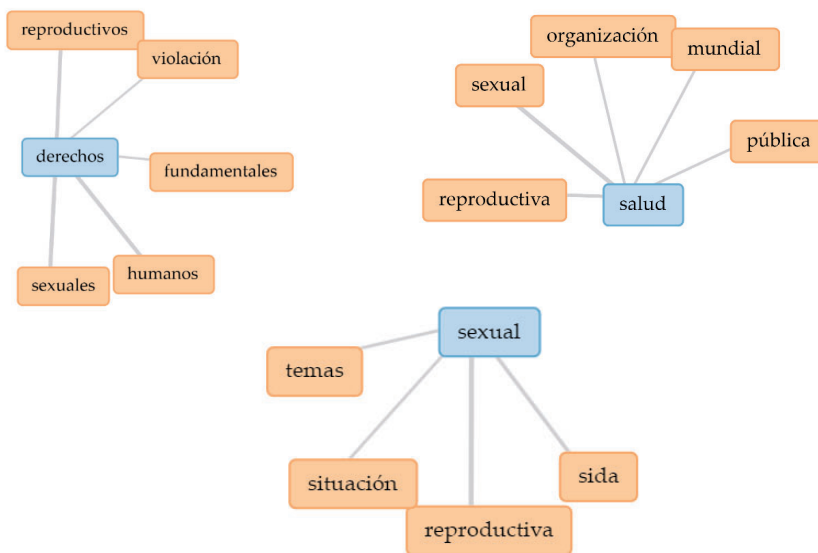
Fuente: Elaboración propia con Voyant Tools.

Al examinar las palabras que están más asociadas con los términos más utilizados, se observan conexiones interesantes. El término ‘derechos’ se relaciona con ‘derechos humanos’ y ‘derechos sexuales y reproductivos’, a pesar de que no aparezca el término ‘aborto’. De manera similar, ‘salud’ está vinculada con ‘salud sexual y reproductiva’, así como con ‘pública’, ‘organización’ y ‘mundial’. En cuanto al término ‘sexual’, se asocia con palabras como ‘temas’, ‘situación’, ‘reproductiva’ y ‘sida’.

Por otro lado, algunos términos que faltan o tienen una menor frecuencia de aparición en este blog son ‘Matrimonio’ (0), ‘feminismo’ (0) y ‘maternidad’ (0), mientras que ‘política’ (3), ‘heterosexualidad’ (4), ‘heterosexuales’ (2) y

‘heterosexual’ (2) son mencionados con menos frecuencia en comparación con otros blogs.

Figura 17. Gráficos de enlaces n°3

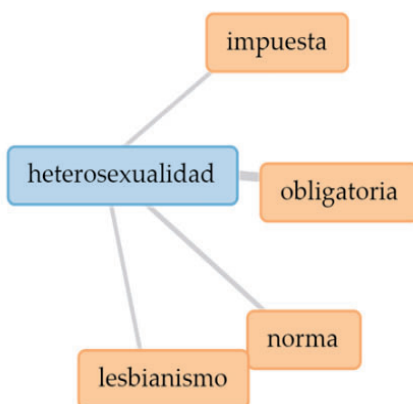


Fuente: Elaboración propia con Voyant Tools.

El blog Malas como las Arañas (<https://malascomolasa.blogspot.com/>) posee funcionamiento desde el año 2008 al 2021, se define como un blog de lesbianas feministas, de la ciudad de la Plata capital de la provincia de Buenos Aires. Podemos ver que las palabras con mayor frecuencia en el corpus de los post son ‘cuerpo’ (25), ‘heterosexualidad’ (22), ‘plata’ (20) ‘nuestra’ (20), ‘domingo’ (18), ‘todas’ (17), ‘hs’ (15) ‘travestis’ (14) ‘violencia’ (13). Podemos notar que el término ‘cuerpo’, siendo el de mayor frecuencia, también coincide con los blogs Lesbianas fugitivas y Baruyera entre Meses. Además, estos blogs comparten el término heterosexualidad y sus derivados como uno de los de mayor frecuencia en su contenido.



Figura 19. Gráfico de enlaces n°4



Fuente: Elaboración propia con Voyant Tools.

Al explorarlo con la herramienta Contextos, donde podemos visualizar las palabras que anteceden y preceden al término, se observa que esta palabra es frecuentemente asociada al término obligatoria, impuesta y a norma. Vemos que nuevamente el término ‘heterosexualidad’ posee un poder de explicación de la opresión, definiéndola como norma obligatoria.

En cuanto a los términos que no están o aparecen con una frecuencia muy baja, se observa que ‘matrimonio’ (0) no se encuentra en el corpus, mientras que ‘derechos’ (2), ‘salud’ (5) y ‘aborto’ (7) aparecen con una frecuencia mínima. Estos términos se pueden asociar con demandas específicas dirigidas al estado. En contraste, los términos con mayor frecuencia dentro del corpus se vinculan con la identidad lésbica y con la visibilización de una norma opresiva como la heterosexualidad.

## 5. DISCUSIÓN Y CONCLUSIONES

A partir de esta exploración pudimos identificar términos relevantes que nos aproximaron al contenido de los blogs de activismo lésbico. Como también a elementos discursivos de los blogs que nos permiten identificar contra qué luchan mediante su producción de contenido. En esta primera exploración, a partir de las visualizaciones de los términos de cada blog con las herramientas de Voyant Tools resultó posible indagar sobre su frecuencia o ausencia en cada uno de los blogs analizados. Esto nos permitió caracterizar el tipo de activismo que representan estos blogs como para determinar si están más enfocados en la defensa de derechos, en la promoción de cambios políticos o en el reconocimiento de identidades.

Teniendo en cuenta el análisis anterior sobre los términos más frecuentes podemos clasificar dos tipos de contenido en los blogs. Uno más ligado al reclamo por derechos y demandas al estado. En particular por el derecho al aborto, derechos sexuales y reproductivos y salud Y otro relacionado a la visibilización de la identidad lesbiana vinculado con los términos ‘cuerpo’/ ‘cuerpos’ y ‘sexual’.

Al mismo tiempo el término ‘hetero’, que incluía las múltiples variaciones de la palabra heterosexual, se volvía uno de los términos más utilizados en los blogs Lesbianas Fugitivas y Baruyera entre Meses. Mientras que en Malas como las Arañas y Baruyera entre Meses. el término heterosexualidad aparecía dentro de los términos más frecuentes. En los tres casos cuando buscábamos los términos más cercanos a esta palabra aparecían ‘norma’, ‘régimen’, ‘impuesta’, ‘obligatoria’, ‘homofobia’. Palabras que hacen referencia a características de este término, no solo como una orientación sexual, si no como un régimen o norma impuesta y obligatoria. Al mismo tiempo, que se encuentre asociada al termino homofobia también resulta interesante, ya que heterosexualidad como norma incluye, ya no como una orientación sexual, las conductas discriminatorias u ofensivas hacia la homosexualidad y en este caso particular hacia las identidades lesbianas.

De esta manera, podemos identificar que el blog Mujeres a Bordo se relaciona más con el primer tipo de activismo, centrado en temas como la demanda por el aborto legal y los derechos reproductivos. Por otro lado, los blogs Lesbianas Fugitivas, Malas como las Arañas y Baruyera entre Meses comparten términos como ‘cuerpo’ o ‘cuerpos’, y una lectura más detallada revela que abordan principalmente la identidad lesbiana. Al mismo tiempo, el blog El Ultra Violetas, junto con Baruyeras, comparten términos relacionados con la salud, el aborto y los derechos reproductivos, similar a Mujeres a Bordo. Sin embargo, también comparten temas con los blogs más enfocados en la identidad lesbiana, como ‘cuerpos’ y ‘sexual’.

De esta manera resultan dos tipos de blogs, unos que producen contenido más vinculado a demandas al estado y derechos particulares como Mujeres a Bordo. Y otros que se encargan de disputar representaciones y sentidos sobre la sexualidad y la identidad lesbiana como sucede con Lesbianas Fugitivas y Baruyeras entre Meses. Y otros como Baruyeras entre Meses y Ultra Violetas comparten temáticas con ambos tipos de blogs. Temáticas que entendemos como parte de las estrategias políticas del activismo LGBTIQ+. Cabe resaltar que las demandas al estado se encuentran mayormente relacionadas con la agenda feminista de los años 2000s que con reclamos específicos del activismo lesbiano. Esto se hace evidente en la ausencia o la poco significativa frecuencia del término matrimonio en todos los blogs. Ya que una de las demandas que se

esperaba encontrar era alguna relacionada con la ley de matrimonio igualitario sancionada en 2010 en la Argentina.

En este sentido podemos ver que la producción de contenido de estos blogs se orientó en colocar en la esfera pública a partir del uso de Internet las demandas particulares del activismo lesbiano feminista. Como también en denunciar a la heterosexualidad como un sistema de opresión de estas identidades y al mismo tiempo como una norma. De esta manera, entiendo que los blogs dieron la posibilidad a estos grupos de activismo para difundir ideas, demandas de formas que hubieran sido inasequibles en la prensa escrita o medios de comunicación hegemónicos como la televisión. Si bien en nuestro país los Cuadernos de Existencia Lesbiana marcaron un inicio del uso de las tecnologías disponibles para el activismo, con Internet y en particular con los blogs vemos como se potenció el desarrollo de este tipo de producciones. Dando espacio a discursos que denunciaron las violencias experimentadas por el colectivo, pero también para difundir sus propias ideas. Como indicamos, la temática en común que comparten los blogs que en su contenido se vincula con la transformación política es el término 'heterosexualidad', entendida como obligatoria, norma o régimen que afecta a las identidades lesbianas. Al mismo tiempo se la asocia también al término 'homofobia' en el blog Baruyeras entre Meses. También encontramos que se utilizaba junto a los términos homofobia, 'lucha', 'contra' y 'violencia' en el blog Lesbianas Fugitivas. Considerando el contenido de los blogs como discursos políticos, resulta posible identificar la construcción de contradestinatarios, siguiendo el enfoque de Eliseo Verón (1987), a través del uso del término 'hetero\*' y sus derivados. Este término no solo se construye como un contradestinatario, sino que también adquiere un poder explicativo en relación con la opresión y la lucha en el activismo lésbico.

## REFERENCIAS

- Burler, J. (2007). *El género en disputa: el feminismo y la subversión de la identidad*. Paidós.
- Cano, V. (2017). *Políticas del archivo y memorias tortilleras: Una lectura de los Cuadernos de existencia lesbiana y Potencia tortillera*. Centro de Investigaciones y Estudios sobre Cultura y Sociedad. Programa de Estudios de Acción Colectiva y Conflicto Social; Onteakien; 24, pp. 11-39.
- Chojnicka, J. (2015). Contesting Hegemonic Gender and Sexuality Discourses on the Web: Latvian and Polish Discourses of Gender Dissidents. *Critical Approaches to Discourse Analysis across Disciplines*, 7(2).
- de Moraes Carvalho, K. J., & Lionço, T. (2018). Pânico moral, homofobia e discurso de ódio em um blog na Internet. *Revista Periódicus*, 1(10), pp.163-183.

- del Mónaco, R. L., Zuviría, R. L. *Activismos LGBTIQ: procesos de organización, reconocimiento y modos de vida*. Centro de Investigaciones y Estudios sobre Cultura y Sociedad; Oteaiken; 24, pp. 1-13.
- Fanese, G. (2014). *Estereotipos de género y blogs: Ethos en disidencia*. La aljaba, 18, pp.189-212.
- Fraser, N. (2000). Heterosexismo, falta de reconocimiento y capitalismo: una respuesta a Judith Butler. *New left review*, 2(1), pp. 123-134.
- gabrielaa (13 de marzo de 2010). Natalia Gaitán. *Mujeres a Bordo*. <https://mujeresabordo.blogspot.com/>
- gabrielaa (25 de junio de 2010). Que vengan los que nos quieren a la marcha por el matrimonio gay, por Patricia Kolesnikov. *Mujeres a Bordo*. <https://mujeresabordo.blogspot.com/>
- Hernández Fernández, C. (2015). Nuevos recursos para la investigación cualitativa: Software gratuito y herramientas colaborativas. *Opción*, 31(5), pp. 453-471.
- Khusro, S., Jabeen, F. & Khan, A. (2018). *Tag Clouds: Past, Present and Future*. Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci.
- Moretti, F. (2015). El matadero de la literatura. *Lectura distante*. Fondo de Cultura Económica.
- Pérez-Riedel, M. D. (2018). Espacio público digital, disputas y violencias: comentarios en dos blogs argentinos LGBTIQ. *Comunicación y medios*, 27(38), pp. 125-137.
- Perlongher, N. (2016) *Devenires Minoritarios (recopilación)*. DIACLASA.
- Preciado, P. B. (2008). *Testo yonqui*. Espasa.
- Rich, A. (1996). Heterosexualidad obligatoria y existencia lesbiana. *DUODA: estudios de la diferencia sexual*, pp. 15-48.
- Rubin, G. (1989). Reflexionando sobre el sexo: notas para una teoría radical de la sexualidad. *Placer y peligro. Explorando la sexualidad femenina*, 113, p. 190.
- Tarducci, M. & Rifkin, D. (2010). Fragmentos de historia del feminismo en Argentina". En Chaher, S. & Santoro, S. *Las palabras tienen sexo II: herramientas para un periodismo de género*. Artemisa Comunicación Ediciones.
- ultravioletas (septiembre de 2007) Sobre nosotras. *Ultra Violetas*. <https://mendezmujeresgenero.blogspot.com/>
- Van Dijck, J. (2016). *La cultura de la conectividad. Una historia crítica de las redes sociales*. Siglo Veintiuno.
- Verón, E. (1987). La palabra adversativa. En *El discurso político: lenguajes y acontecimientos*. Hachette.
- Vespucci, G. (2015). Identificaciones sexuales politizadas y modos de vida lesbicos: un análisis sobre Cuadernos de Existencia Lesbiana (Buenos Aires, 1987-1996). *Question/Cuestión*, 1(47), pp. 432-448.
- Wittig, M. (1992). *El pensamiento heterosexual y otros ensayos*. Beacon Press.

# DISCRIMINATION AND HATE SPEECH AMONG TRANSGENDER INDIVIDUALS IN AN ONLINE COMMUNITY: BETWEEN PASSING AND TRANS- NORMATIVITY

MARIANNA COPPOLA

*Research Fellow in General Sociology. DISUFF  
University of Salerno, Italy*

GIUSEPPE MASULLO

*Associate Professor in General Sociology. DISUFF  
University of Salerno, Italy*

**DOI:** 10.14679/2764

## **Summary:**

1. Introduction.
  2. Hate speech towards and among LGBTQ+ individuals.
  3. Sexual normativity and trans-normativity.
  4. The research.
    - 4.1. The Informa Trans Italia community.
  5. Results and discussion.
    - 5.1. Defining and passing: the main forms of intra-categorical discrimination within the Informa Trans Italia community.
    - 5.2. Discriminated users and targets of hate speech among binary and non-binary transgender individuals: a typology.
  6. Conclusions.
- References.

## 1. INTRODUCTION

Since its inception, the Internet has proven to be a valuable opportunity for emancipation, particularly for those who, discriminated against and stigmatised in the real world because of their differences, have found ways and possibilities to express themselves in the spaces of digital society. This is the case of homosexual, bisexual, and transgender individuals who exchange information, build connections, and experience their identities beyond heteronormative norms and standards on digital platforms<sup>1</sup>.

Wakeford (1997) emphasised how cyberspace is characterised as cyberqueer, *i.e.*, a context in which one can experiment with and express their gender and

---

<sup>1</sup> The chapter was devised and written jointly by the authors. However, for the sake of authorship, the sections 2. 4. 5.1 are attributed to Giuseppe Masullo, the sections 3. 4.1.5.1 are attributed to Marianna Coppola. The authors co-wrote the introduction and conclusions.

sexual fluidity. It is an ideal space where individuals can, on the one hand, “escape” from heteronormative canons and, on the other hand, “narrate” their identities using performative language that contributes to the redefinition of self-expectations and social expectations (Butler, 2004).

However, recent scholars (Farci, Scarcelli, 2022) have also highlighted how the digital space, far from being solely a place of liberation, serves as a context where old and new forms of discrimination are experienced. These forms of discrimination often manifest through offensive and derogatory language aimed at belittling individuals due to their differences.

While, in recent years, the scientific community has devoted considerable attention to the issue of hate speech, partly because of its alarming spread in social media, there has also been a prevailing rhetoric that views hate speech as a characteristic feature of conflicts between social groups, between the “us” and the “other”, the “included” and the “excluded”, in-group and out-group. This perspective fails to grasp the transversal and intra-categorical nature of the phenomenon. Hate speech is not a distinctive trait of the “oppressors” but can also be perpetrated by the “oppressed” within their reference communities.

As highlighted in recent studies by Masullo and Coppola (2021), rainbow communities replicate dynamics typical of social groups both in the real world and in virtual spaces. These dynamics, based on physical, psychological, and social parameters, determine inclusivity and exclusivity, defining who can be part of the community and who is excluded.

The application of intersectionality theory to hate speech studies thus allows us to shed light on the differences that emerge within the LGBTQ community in terms of languages, imaginaries, and power positions, thereby expanding the scope for discussing intra-categorical discrimination.

Based on this premise, this research paper aims to analyse the phenomenon of *social discrimination* and *hate speech* within the transgender community, considering internal hate speech as a way of identifying and describing those physical, psychological, social, and identity-related characteristics that constitute an ideal and normative reference model for transgender individuals.

The first part of the paper will present the theoretical framework. It will address the main theories concerning hate speech, with a particular focus on marginalised categories within the LGBTQ+ spectrum. Additionally, it will introduce the concept of sexual normativity and, more specifically, that of *trans-normativity*, which is the dominant and elective identity construction among transgender individuals (Johnson, 2016; Coppola, 2023).

Finally, to address the research questions, we will present and discuss the results of a netnographic study involving one of the main Italian Transgender communities on the social network Facebook. The study was conducted between

March and October 2022, and it aimed to investigate the discrimination experienced by transgender individuals within the community through the analysis of posts and interactions among users.

## 2. HATE SPEECH TOWARDS AND AMONG LGBTQ+ INDIVIDUALS

The proliferation of social media has represented an opportunity for emancipation and social visibility for many marginalised groups, offering new spaces and communication modalities that are distinct from traditional face-to-face interaction (Kendell, 2002). However, the unique characteristics of online communication, such as immediacy, diversification of audiences and recipients, and the potential for anonymity and covert behaviour, have also provided fertile ground for the spread of hate messages.

Hate speech refers to offensive content (which can be verbal/textual, visual, symbolic, or non-verbal) aimed at belittling, denigrating, or dehumanising one or more individuals belonging to specific categories or social groups (Simpson, 2013). The online variant of hate speech involves denigrating or expressing hatred towards a specific target through multimedia content disseminated in cyberspace. Hate speech can be motivated by racial, sexual, religious, political, sporting, or other factors (Hawdon, Oksanen & Räsänen, 2017).

In 2017, Brown conducted a comparative study between offline and online hate speech and emphasised the greater negative impact on the psychological health of victims of online hate speech. The vast online audience and the resulting “systems collapse”, typical of online communication, lead victims to experience greater shame and social anxiety compared to hate speech perpetrated in the offline world.

Mondal *et al.* (2017) further highlighted a correlation between anonymity and hate speech. Some users adopt invisibility as a disinhibitory strategy that allows them to vent their anger, resentment, and hatred toward specific social categories. The authors also pointed out that the social groups most targeted by online hate speech are immigrants and individuals with non-heteronormative sexual orientations. Indeed, the LGBTQ+ population appears to be one of the primary targets of online hate speech. This situation contributes to the reinforcement of stereotypes and prejudices within mainstream society and harms the psycho-physical and social well-being of LGBTQ+ individuals.

Friedrisen-Goldsen *et al.* (2014), in a comprehensive survey on the mental health and psycho-social well-being of LGBTQ+ victims of online hate speech, highlighted a greater risk of health-related deviant behaviours (such as social withdrawal, panic attacks, episodes of anger, etc.) compared to the cisgender and heterosexual population also subjected to hate speech. In particular, the

transgender population experiences depressive symptoms and social withdrawal twice as often as the other non-normative sexual identities and four times as often as the cisgender heterosexual population. Furthermore, Habbeard (2020) has emphasized that the transgender population is particularly vulnerable to online hate speech compared to the LGB population. Hate speech mainly targets transgender women, as they are still associated with stereotypes that place them on the fringes of society and the world of prostitution (Coppola, 2023).

In a study aimed at analysing the effects of internalising negative aspects and social stigmatisation among LGBTQ+ individuals who are victims of hate speech, Ghafoori *et al.* (2019) highlighted a correlation between internalised homophobia and transphobia and the experience of online verbal violence. This contributes to the phenomenon of *intra-categorical stigmatisation*. According to the authors, exposure to hate speech can lead LGBTQ+ individuals to engage in discriminatory and intolerant behaviour towards other members of the rainbow community, resulting in a form of intra-categorical desensitisation (Seglow, 2016).

### 3. SEXUAL NORMATIVITY AND TRANS-NORMATIVITY

In the study of social inclusion and exclusion processes among diverse sexual identities, the concept of *sexual normativity* plays a central role. Sexual normativity refers to the construction of identity considered normative and accepted by members of a specific society. This identity construction encompasses a combination of individual aspects (physical and psychological) and social aspects (related to status, culture, and personal experience) that constitute the parameters of normalcy for a specific sexual identity. Different instances of sexual normalcy exist within all sexual identity constructions, and there are interdependencies among them. As we will see, not all types of sexual normalcy hold the same hierarchical position in society; some are more visible and have more power than others.

Several authors, including Butler (2004) and Connell (2009), have extensively argued about the alignment of the cis-heteronormative model with the dominant sexual normativity in society. This supremacy has been nurtured and fuelled by the assumptions of the heterosexual model, which involves a one-to-one correspondence between biological sex and gender identity, the existence of only two possible genders (gender binary), and the existence of only one sexual orientation considered normative as a prerequisite for generativity (heteronormativity). This model is firmly rooted in the essentialist conception shared and supported even within scientific communities (such as medicine, genetics, and jurisprudence). Over time, other non-normative sexual identities have developed their identity constructions, which are considered

standard and bearers of authenticity and conformity of the social and collective identity of that specific sexual instance.

While there has been extensive discussion of the historical and conceptual evolution of homosexual identity within mainstream society and the LGBTQ+ community, scientific and scholarly interest in the analysis and study of transgender identity in terms of dominant, prototypical, and legitimised models within the transgender community has been almost absent.

A recent work on the concept of trans-normativity, thus far unique in its comprehensiveness, was proposed by Austin H. Johnson in 2016 in an article published in *Sociological Inquiry* titled “Transnormativity: a new concept and its validation through documentary film about transgender man”. The author analysed the dominant gender models and gender expression within the transgender community, particularly among transgender men, through the study of various documentaries, films, podcasts, and social media videos in the American landscape. Johnson specifically analysed 9 multimedia products featuring a total of 57 transgender men and highlighted that almost all the individuals in the analysed multimedia products were “medicalised” transgender people. This means that they had, to varying degrees, undergone a gender affirmation process following official medical and healthcare protocols. Only 1% of individuals were non-medicalised, outside the gender affirmation process, but the author noted that the profiles within this percentage were “about to begin” the transition process.

Johnson defines trans-normativity as a condition of a transgender person that tends to meet several parameters considered prototypical of the normative model. Specifically, a transgender person is considered normative if they have phenotypically adjusted their appearance following their desired identity and, therefore, do not exhibit “residual signs of sexual ambiguity”. Additionally, a normative condition completes the legal process with the change of documents and personal data in line with the laws and regulations. Furthermore, through the narratives analysed in the multimedia products, the scholar emphasised how non-binary and genderfluid identities are demonized and deemed to deviate from the normative model of a transgender person, namely the dominant medicalized model. The new expressions and gender identities do not aim to “restore the norm and the binary gender order” but rather to challenge and redefine that order. Thus, the spread of non-binary perspectives is viewed as a “revolutionary and political” act rather than an authentic expression of the process of self-determination.

Johnson’s analysis excludes the condition of non-binary, and trans-normativity appears to be the exclusive domain of only binary transgender identities. Non-binary refers to identities that do not recognise the male/female gender binary. Non-binary individuals do not exclusively identify with

either male or female gender identity but have fluid, open, oscillating, and situational identity constructions. Genderfluid identities oscillate between the two genders or neither, and their gender expression changes according to parameters experienced as episodic and contextual. Agender identities reject closed and immutable gender affiliations, constructing new gender expressions without reducing them to known and dominant genders (Scandurra, Mezza, Bochicchio, 2019).

Coppola (2023), in her book titled “Trans-normativity: Gender models and processes of normalisation among transgender and non-binary individuals”, sought to outline the dominant identity construction within the transgender community. She highlighted how non-binary transgender individuals propose a more fluid and less rigid view of sexual orientations, thereby challenging the dominant models that only envisaged three possible sexual orientations: heterosexual, homosexual, and bisexual.

#### 4. THE RESEARCH

As observed in the previous sections, every sexual identity (normative or non-normative) tends to define a normative identity construction shared by most of its members, establishing physical, psychological, and social criteria considered as ideals to attain. Within the transgender universe, these parameters are constantly negotiated and performed through online interactions, with digital platforms being particularly used by transgender individuals for their identity self-determination.

Based on this premise, the present research aimed to analyse the phenomenon of *social discrimination and hate speech* within the transgender community, considering internal hate speech as a means to identify and describe those physical, psychological, social, and identity-related characteristics that constitute an ideal and normative reference model for transgender individuals.

We hypothesised that hate speech within the transgender community manifests itself in the form of often offensive and derogatory language directed at those who deviate from that ideal reference model. To better understand and analyse the aspects presented so far, we formulated the following research questions:

RQ1. What are the phenotypical, psychological, and social characteristics that the transgender community ideally look at for the purpose of normalisation?

RQ2. To what extent do these characteristics deviate from the cis-heteronormative model?

RQ3. What argumentative strategies do transgender individuals use to praise or devalue individuals who do not conform to dominant models?

To address the research questions formulated above, we conducted a netnographic study choosing one of the transgender-themed Facebook groups in Italy as the virtual context of analysis. The selection was based on theoretical sampling considerations and criteria, allowing for the choice of cases that could provide quantitative and qualitative information necessary and useful for the study, thereby obtaining a relatively close sample to the target population (albeit not representative). The research was conducted covertly to respect the setting, community rules, and user interactions, and avoid guided, untruthful, or highly socially desirable responses. The analysis focused on a six-month timeframe, from January 10th to May 10th, 2023, and examined posts of introductions, responses, comments, user interactions, and life stories that had topics related to identity processes, social legitimacy processes, social inclusion and/or exclusion, gender models, and gender-related aspects.

We analysed approximately 789 posts, interactions, introductions, and life stories. The interpretation followed the criteria and principles of the hermeneutic approach. We aimed to emphasize the identity dimensions, the perspectives of the individuals, their experiences, emotions, and relational aspects, to identify recurring traits and points of differentiation among the members of the community, whenever possible.

#### **4.1. The Informa Trans Italia community**

As previously stated, the community examined in this research is the Facebook group dedicated to transgender individuals “Informa Trans Italia”. The Informa Trans group was created with the primary purpose of providing information to transgender individuals about Gender Incongruence, and more specifically, serving as a reference point to guide and assist individuals in their process of gender self-determination and/or gender transition. Over time, the group has taken on the role of a “virtual town square” where, in addition to exchanging information, it is possible to share one’s life experience, form relationships and friendships, meet partners, and organise “in-person” events aimed at informing, educating, and raising awareness about transgender issues.

The group was founded in 2013 by a transgender man who had started a YouTube channel a few years earlier, where he regularly uploaded videos and shared experiences related to his gender transition journey. Over time, the group administrators have organised annual gatherings and “in-person” meetings in various Italian cities, providing users with the opportunity to meet offline and thus form friendships and connections.

The Informa Trans Italia group numbers approximately 4,400 members from all over Italy. Thanks to the group administrators’ help, we were able to gather some socio-biographical information about the members. It emerged

that a higher percentage of transgender individuals identify with binary gender (approximately 67%) compared to various forms of non-binary identity (33%). There is a geographical concentration around major Italian cities or capitals such as Rome, Milan, Naples, and Bologna. The average age of the users is 29 years old. 50% of the members are over 30, 40% are between 29 and 20, and 10% are between 20 and 18 years old.

## 5. RESULTS AND DISCUSSION.

### 5.1. Defining and passing: the main forms of intra-categorical discrimination within the Informa Trans Italia community.

Observing the interactions among users, the different relational modalities, and the various ways of including and/or excluding people who are part of the examined community, a specific and clear dichotomy in self-definition and self-determination emerged: binary or non-binary. Whether one identifies as binary or non-binary would seem to be a private and intimate matter that primarily concerns intrapsychic dimensions. However, this view is partial and underestimates the importance of interpersonal, social, and cultural dynamics that gender-related issues involve on various levels.

As noted by Johnson (2016) and Coppola (2023), the transgender community would be extremely attentive to how they represent, narrate, and perform gender-related matters, particularly elective gender acquired through a multidimensional process that encompasses psychological, medical, social, and legal aspects. The authors have emphasised the presence of trans-normativity within the T community and how it is strongly and inescapably linked to adherence to the dominant model of sexual normativity, namely, cisnormativity and heteronormativity.

The observation of the analysed community has confirmed the tendency to discriminate harshly, and in some cases, through forms of outright hate speech, against individuals who identify as non-binary. The rest of the T community and binary transgender people impute various contradictory aspects to these individuals, which are seen as non-functional to the community's cause. Often, these aspects are overlapped with psychological and private issues that cannot be extended to the entire transgender community.

The following passage provides a good example to illustrate the point just discussed:

«In my opinion, the non-binary issue is political. It is ideology. It is a way of conceiving the world, it is not real compared to the issues of trans people, who need to feel male or female, period! Before, we had to fight with others,

now even among ourselves (Marco Michele, 34 years old, AFAB on a medical transgender care pathway).»

It is also possible to observe various forms of intra-category discrimination among binary transgender individuals. In this case, however, it is not the *defining* process that is being questioned but rather an aspect considered fundamental among transgender individuals who undergo a gender affirmation process through a medical care protocol: namely, achieving a high level of passing<sup>2</sup>.

The word *passing* first appeared in the United States about the phenomenon of black people who, due to their lighter phenotypic characteristics, appeared as white. This allowed them to assume roles, positions, and privileges typical of the social category considered “normal” and, consequently, “dominant”. Within the LGBT community, the term passing takes on the dynamic of attributing normative and heteronormative characteristics by an external observer. Specifically, it involves attributing heterosexual characteristics to a homosexual person or cis-sexual characteristics to a transgender person.

Physical appearance, a fundamental prerequisite for achieving a high level of passing, becomes the main parameter of social differentiation between “normative” binary transgender individuals and those who, for various reasons, appear to be “non-conforming”. Most of the discrimination and hate speech experienced by binary transgender individuals relate to their physical appearance and phenotypic characteristics, insofar as these do not guarantee passability as “cis-sexual” individuals, as highlighted in the following post:

«Good morning, here is my photo at six months of hormone therapy, what do you think? (Federica, 22 years old, AMAB on a medical transgender care pathway).

- Answer: and you even take pictures? With that man’s nose where do you think you are going? Before you take another picture go get your nose done! You’re so ugly! (Ines, 32 years old, AMAB on a medical transgender care pathway).

- Just the nose? And that skirt? You look like my grandma, terrible... (Marta, 22 years old, AMAB on a medical transgender care pathway).»

---

<sup>2</sup> Some factors condition and influence the process of passing for normal:

- Biological, physical, and biochemical factors: genetic predisposition, biological and somatic characteristics, and anthropometric aspects influence the process of passing for normal; for example, an MtoF transgender person with a height above female standard).

- Gender expression factors: the way one represents and expresses the desired gender influences the process of passing for normal; for example, an FtoM transgender person who tends to wear a shaved beard would present a low passability compared to an FtoM who presents a male gender expression with a beard.

- Social and relational factors: associating with non-LGBT individuals or those with high passability would facilitate the passing for normal process for transgender individuals, as they would fit into a heteronormative social and contextual construct.

There are also derogatory comments related to psychological, social, and role dimensions. From an analysis of posts and relative interactions, a form of discrimination emerges, albeit less frequent than that related to phenotypic passing, which refers to adherence (or lack thereof) to the gender role ascribed to the chosen gender. There is, therefore, a tendency to also attack binary transgender people who do not “behave like cis-heterosexual people”, thus contravening the main criterion of passing for normal, namely, emulating in all characteristics the normative and dominant identity.

The following posts clarify the dilemma of adherence to the dominant elective gender role for binary transgender individuals:

«Good morning, I would like to share my absolute dislike for children. I believe I don't want to think about motherhood at all. I'm not interested, just as I'm not very interested in things that make me look like cisgender people. Am I the only one who thinks this way? (Hester, 23 years old, AMAB on a medical transgender care pathway).

- Answer: And why did you become a woman? To look like a transvestite? Or a faggot? How can we be credible if we make ourselves into pariahs? My goodness, what a drag, all of it! (Valeria, 23 years old, AMAB on a medical transgender care pathway).»

Cyberspace is often considered by non-heterosexual transgender individuals as a place where they can find experiences, connections, and information to legitimize their identity. It is also seen as a space to seek potential friends and/or partners who are more in line with their emotional and sexual needs. The choice of non-heterosexual transgender individuals to join transgender-themed groups is motivated by an ongoing search for digital spaces where they can experience less invisibility and find forms of communication and interaction that better suit their needs. However, the condition of double normative misalignment exposes non-heterosexual transgender individuals to various forms of discrimination, both in mainstream society and within the transgender community.

The following passage describes the experience of a homosexual transgender woman who recounts the bitterness she experienced during an interactive exchange following her introduction on the day of joining the group:

«When I say that I'm an MtoF lesbian, the other straight trans women attack me! It's as if they believe I must like men just because I've become a woman, and hearing this from other trans people is truly disheartening! It's not easy to make them understand that gender identity is separate from sexual orientation. (Christel, 28 years old, AFAB on a medical transgender care pathway).»

Non-heterosexual transgender individuals appear to face greater relational and socialisation difficulties compared to heterosexual transgender individuals

due to a general confusion in the defining processes of gender identity, sexual orientation, and gender expression, even within the LGBTQ+ community. It is challenging to navigate a gender transition journey, both cognitively and socially, while having a homosexual orientation. This aspect places homosexual transgender individuals in a double stigmatisation: first, as a transgender person who challenges the gender binary, which views somatic and ideological connotation as the basis for gender identity, and as a homosexual person after the transition process, experienced as a fundamental inconsistency, making them doubly non-normative (Masullo & Coppola, 2021).

## 5.2. Discriminated users and targets of hate speech among binary and non-binary transgender individuals: a typology

The research experience described in the previous sections has allowed for the construction of a typology of discriminated users and targets of hate speech within the online transgender community. To determine the four types that make up this typology, we identified two fundamental dimensions through the analysis of posts. The first dimension is the level of phenotypic passing, which refers to the extent to which individuals can pass as cisgender people. The second dimension is adherence to cis-heteronormative models in terms of affirming their chosen gender (see Figure 1).

Figure 1. Types of people discriminated against in the transgender community

Adherence to cis-heteronormative models	Phenotypic passing	
	Low	High
High	<b>Impostors</b> Adhere to cis-heteronormative models but do not achieve a good level of phenotypic passing, hence they are discriminated against as they are not perceived as legitimately assuming their chosen gender role.	<b>Homogenised</b> Adhere to cis-heteronormative models and exhibit a good level of phenotypic passing, representing individuals with fewer incidents of discrimination and intra-category hate speech.
	<b>Subversives</b> Do not adhere to cis-heteronormative models and do not achieve a good level of phenotypic passing, represent a subversion of the system of sexual and social normativization and thus experience higher levels of discrimination.	<b>Innovators</b> Do not adhere to cis-heteronormative models while still achieving a good level of phenotypic passing, which leads to criticism and discrimination due to their behaviours and social actions challenging the well-established stereotypical construction of genders.
Low		

Source: Authors.

In the first quadrant, here called *Impostors*, the authors of the posts show high adherence to cis-heteronormative models in terms of gender identity but face forms of disapproval from the online community based on what is considered low phenotypic passing. This aspect leads to a “perceptual blackout” with frequent misgendering and deadnaming, as individuals falling into this category are often overlooked and not recognised for their chosen gender. Most of the discrimination and hate speech directed at this type of user revolve around their lack of “passability”, as the person emphasises through their posts and photos that their gender does not conform to the chosen one. The following passage clarifies the discrimination process just explained:

«Good morning, this is me today and I would like to have my first day at the beach. What do you think? Do you think I pass as a guy? Suggestions accepted (Federico, 33 years old, AFAB on a medical transgender care path).

- Answer: Forgive me but are you on T (testosterone)? Because if you are on T, you need to change your endocrinologist! You don't pass at all; you look like a 12-year-old girl... It's not a binder issue; it's your face! (Livia, 30 years old, AMAB)

- But what are these butch lesbians thinking, believing they are FtoM just because they wear their brother's clothes? Ridiculous! (Enrico, 27, AFAB in follow-up).»

The second quadrant represents a type resulting from the intersection of low adherence to cis-heteronormative gender roles and low level of phenotypic passing. We referred to this type as *Subversive*, signifying a person who, in their process of self-determination, seeks to propose a completely different vision of genders and the relationships between them, with the potential for more fluid or changing polarisations than prescribed gender roles. It is precisely this revolutionary and simultaneously subversive nature of the criteria for interpreting the world of gender and sexuality that constitutes the main reason for discrimination and episodes of hate speech within the T community, particularly directed at non-binary transgender individuals. They are seen as a “regression” in credibility and legitimacy compared to the social acceptance achieved by binary transgender individuals, a recognition also obtained through the legal and medical mechanisms in gender affirmation processes. Non-binary or non-heteronormative self-determination processes are thus perceived by the rest of the transgender community (and the broader LGBTQ+ community) as unstable and dysfunctional.

The following post highlights the difficulty many binary transgender individuals face in accepting and assimilating non-binary perspectives:

«Can I say my piece? Yesterday, I saw someone on a Rete4 TV show who identified as “agender” and they only further confused the audience in

attendance and, in my opinion, those watching at home as well. I believe that agenderism is undermining all the work that transgender people have done over the years by presenting a version of themselves that is difficult to understand. And who pays the price? All of us who are referred to as circus freaks. You are ruining us all. It's my opinion, but I'm certain of it. (Kissy, 29 years old, AMAB on a medical transgender care path).

- They are ruining us all; in my view, they shouldn't even be considered transgender. They are confused, confusing, and ridiculous. I just can't stand them. (Ivan, 30 years old, AFAB on a medical transgender care path).»

The third quadrant features a type that we call *Innovator*, with a high level of phenotypic passing but low (or absent) adherence to cis-heteronormative models in terms of gender roles. These users are perceived with the chosen gender but deviate from the associated gender expectations through their posts and photos, consciously (or unconsciously) proposing innovative models where masculinity and femininity do not operate in watertight separate compartments but often blend or reverse. This includes emphatic men (Coppola, 2023), that is, transgender men who do not conform to the dictates of hegemonic and stereotypical cisgender masculinity but instead embrace and support alternative forms of masculinity, where aspects of femininity are integrated into their male identity. This new vision of masculinity (Connel, 1995; Demitou, 2009, Masullo, Coppola, 2021, Coppola, 2023) is often the target of hate speech by transgender women and some transgender men who uphold a more traditionalist view of gender roles, closer to the dominant cis-heteronormative model in mainstream society.

«The following passages are examples of hate speech directed at emphatic men. Do you ever cry while watching a movie? I always do, especially when it's a love story. I think that sensitive men have a hard time in relationships and daily life. I would like to know your opinion. Thank you (Marco, 30 years old, AFAB)

- Answer: I believe that a transgender man will never be a real man[...] maybe that's why I'll never date an FtoM; it doesn't feel like a real man to me! (Jessica, 37 years old, AMAB).»

The last type, which we can define as Homogenised, represents the least discriminated user within the community and can be considered as the closest to the internal trans-normativity of the group. They exhibit a high level of phenotypical passing and adhere to the cis-heteronormative models present in mainstream society. Since they are less discriminated against and less subject to hate speech, we will not provide examples of discrimination involving them. Instead, we will present their introductions, in which the desire to “pass” as cis-hetero or to minimise the gap between the dominant sexual identity (specifically, cis-heteronormativity) and their own is evident.

«Hello, I am Silvia, I am 33 years old, I am from Rome, and I am a transgender girl who has had surgery. I would like to say that I am heterosexual, sorry but I don't like girls. I am attending the 3rd year of nursing studies at La Sapienza University, I enrolled late because I waited to finish the course. I would like a relationship with a cisgender boy, I like animals [...] [...] My dream is to find a job and a boyfriend and start my normal life after so many sacrifices.

This is me in the picture [...] what do you think?»

## 6. CONCLUSIONS

The research showed that in the examined online transgender community, cisnormativity and heteronormativity constitute the horizon to which transitioning individuals look to assess themselves and others according to sexual normativity. Those who do not conform to this standard or who achieve low levels of passing are the ones most exposed to discrimination, often in the form of offensive language aimed at belittling and dehumanizing others, denying them a core aspect of a successful transition process: social recognition. Hate speech within the transgender community thus highlights a widespread intra-categorical desensitisation among transgender individuals, undermining a supposed internal group solidarity, which is often taken for granted in the collective imagination.

In line with the studies of Johnson (2016) and Coppola (2023), the dominant identity construction among transgender individuals is binary, heterosexual, and follows a medical care pathway – that is, those that at the end of the gender affirmation process most closely resemble cis-heterosexual identity. The higher the level of trans-normativity within the T community, the more likely it is to experience forms of discrimination and verbal expressions of hatred for those individuals who, for both subjective and ideological reasons, do not adhere to it. Hate speech would thus demarcate the symbolic and imaginary line between the “true transgender person” and other identities that are not deemed to match this status.

The results of this study, therefore, prove valuable in envisioning future prevention programs for gender and sexual discrimination, not only aimed at individuals external to the transgender community but also within it, to make members more aware of alternative paths to the canons imposed by the dominant trans-normative model today.

## REFERENCES

- Butler, J. (2004). *Undoing gender*. Routledge.
- Brown, A. (2017). What is hate speech? *Law and Philosophy*, 36, 561-613.
- Connell, R. W. (2006). Understanding men: Gender sociology and the new international research on masculinities. *The handbook of gender and education*, 18-30.
- Coppola, M. (2023). *Transnormatività. Modelli di genere e processi di normalizzazione tra le persone transgender e non binarie*, Napoli, Loffredo editore (in press).
- Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, 43, 1241-1299.
- Farci, M., & Scarcelli, M. C. (2021). *Media digitali, genere e sessualità*. Mondadori.
- Fredriksen-Goldsen, K. I., Kim, H. J., Barkan, S. E., Muraco, A. & Hoy-Ellis, C. P. (2013). Health disparities among lesbian, gay, and bisexual older adults: Results from a population-based study. *American journal of public health*, 103(10), 1802-1809.
- Ghafoori, B., Caspi, Y., Salgado, C., Allwood, M., Kreither, J., Tejada, J.L., Hunt, T., Waelde, L.C., Slobod in, O., Failey, M., Gilberg, P., Larrondo, P., Ramos, N., von Haumeder, A. & Nadal, K. (2019). Global Perspectives on the Trauma of Hate-Based Violence: An International Society for Traumatic Stress Studies Briefing Paper. [www.istss.org/hate-based-violence](http://www.istss.org/hate-based-violence)
- Hawdon, J., Oksanen, A., & Räsänen, P. (2017), Exposure to online hate in four nations: A cross-national consideration. *Deviant Behavior*, 38, 254-266.
- Hubbard, L. (2020). *Online Hate Crime Report: Challenging online homophobia, biphobia and transphobia*. Routledge.
- Johnson, A. (2016). Transnormativity: A New Concept and Its Validation through Documentary Film About Transgender Men. *Sociological Inquiry*, 86,4, 465-491.
- Kendall, L. (2002). *Hanging, Masculinities as Homophobia: Fear, Shame and Relationships Online*. University of California Press, Galop the LGBT+ anti-violence charity.
- Masullo, G., & Coppola M. (2022). *Affettività invisibili. Storie e vissuti di persone e famiglie transgender*, Varese, PM edizioni.
- Masullo, G., & Coppola, M. (2021). Scripts and Sexual Markets of Transgender people on online dating Apps: A netnographic study, *Italian Sociological Review*, 11, 4S, 319-341.
- Motschenbacher, H. (2018). "Corpus linguistics in language and sexuality studies: Taking stock and future directions." *Journal of Language and Sexuality*, 7.2: 145-174.
- Scandurra, C., Mazza, F. & Bochicchio, V. (2021). Individui non-binary e genderqueer: Una review critica su salute, stigma e risorse, *La camera blu*, 21, 43-57.
- Seglow, J. (2016). Hate Speech, Dignity and Self-Respect. *Ethical Theory and Moral Practice*, 19(5), 1103-1116.

- Simpson, R. M. (2013). Dignity, harm, and hate speech. *Law and Philosophy*, 32, 701-728.
- Thorne N, Yip AK-T, Bouman WP, Marshall E. & Arcelus, J. (2019). The terminology of identities between, outside and beyond the gender binary - a systematic review. *Int J Transgenderism*, 20, 138-154.
- Wakerford, N. (1997). Cyberqueer. In Medhurst, A., Munt, S.R. *Lesbian and Gay Studies. A Critical Introduction*, London, 20-38.

# HOMOFOBIA 2.0: UNA APROXIMACIÓN TEÓRICA AL DISCURSO DE ODIOS LGTBIFÓBICO EN LÍNEA Y LAS ESTRATEGIAS PARA DESARTICULARLO<sup>1</sup>

FRANCISCO JAVIER SANTOS FERNÁNDEZ  
*Investigador Predoctoral en Formación (FPU20/02848)*  
*Universidad de Huelva, ESEIS/COIDESO, España*

ELENA RUIZ-ÁNGEL  
*Profesora Sustituta Interina.*  
*Universidad de Huelva, ESEIS/COIDESO, España*

PATRICIA RUIZ-ÁNGEL  
*Profesora Sustituta Interina.*  
*Universidad de Huelva y Universidad Pablo de Olavide, Sevilla, España*

**DOI:** 10.14679/2765

## **Sumario:**

1. Introducción.
  2. Homofobia.
  3. La radicalización en línea hacia las personas LGTBQ+.
  4. Consecuencias del discurso de odio homofóbico en línea.
  5. Activismo digital y nuevos movimientos sociales LGTBQ+.
  6. La contranarrativa como estrategia de subversión del discurso de odio en línea.
  7. El uso del humor y la reapropiación de símbolos como contranarrativa hacia la homofobia en línea: el caso de Gaysper.
- Referencias.

## 1. INTRODUCCIÓN

La rápida expansión de internet y la plataformización han tenido como resultado una vertiginosa integración de la digitalización en la esfera social, siendo más acusada tras el desarrollo de las redes sociales, generando una redefinición de los estilos de vida de la ciudadanía. Pero, además, dicho fenómeno ha cambiado la forma en que las personas se comunican entre sí, ya que permite

---

<sup>1</sup> Este capítulo forma parte del Proyecto I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGBTI [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDER A way of making Europe”.

Agradecemos también el apoyo del Ministerio de Educación y Formación Profesional, con la ayuda predoctoral para la Formación de Profesorado Universitario (FPU20/02848) que beneficia a uno de los autores. Nos beneficiamos igualmente del apoyo de los centros de investigación de la Universidad de Huelva, el grupo de investigación “Estudios Sociales e Intervención Social, ESEIS” y el centro “Pensamiento Contemporáneo e Innovación para el Desarrollo Social, COIDESO.

la difusión de información a grandes audiencias de manera instantánea, superándose no solo las barreras geográficas, sino también las sociales y tecnológicas.

Y, si nos remitimos al contexto LGBTQ+ (acrónimo que define al conjunto de lesbianas, gays, personas trans, bisexuales, intersexuales, queers y otras diversidades afectivo-sexuales), las redes sociales se integran directamente en la diversidad afectivo-sexual y de género de las personas, cambiando por completo sus dinámicas y generando nuevos paradigmas. De hecho, el desarrollo de aplicaciones dirigidas específicamente al público LGBTQ+ ha supuesto una revolución y liberación sexual sin precedentes, pasándose de un mundo en el que tenían que reunirse clandestinamente en pubs y bares como Stonewall por miedo a la persecución policial, a otro en el que las minorías sexuales usan el ciberespacio como escenarios en el que expresar sus ideas y opiniones, manifestar disidencias, ampliar su red de contactos, reforzar su identidad, mostrar visibilidad y realizar activismo para combatir la violencia y discriminación mediante la promoción de valores de defensa, orgullo, solidaridad y respeto hacia sus derechos. Además, este tipo de plataformas permiten a quienes no se atreven a “salir del armario”, ya sea porque no lo desean o por miedo a la hostilidad y posibles represalias de su contexto social, a poder vivir su sexualidad, encontrar apoyo y recreación emocional y tener referentes sociales. Sin embargo, la digitalización también tiene una contraparte oscura, favoreciendo la proliferación de discursos de odio que presentan a la homosexualidad de forma muy negativa, encontrándose una diversidad de discursos de odio al respecto. Desde aquellos que consideran la homosexualidad como una amenaza para los valores conservadores y los modelos familiares y culturales hasta una aberración antinatural que pone en riesgo la perpetuación de la especie humana, una desviación sexual que conlleva a practicar la pedofilia, una ofensa moral y pecado mortal, una enfermedad mental y, un vicio y práctica perversa que promueve las infecciones de transmisión sexual.

## 2. HOMOFOBIA

La homofobia puede ser definida como la actitud de rechazo o aversión a la homosexualidad masculina y femenina y a las personas bisexuales y transexuales, la cual se basa en prejuicios y se manifiesta en las esferas públicas y privadas por medio de la incitación al odio, la discriminación, la ridiculización y la violencia verbal, psicológica y física, incluyendo la persecución, el asesinato y las limitaciones injustificadas e irracionales de derechos (Resolución (2012/2657(RSP)) del Parlamento Europeo, de 24 de mayo de 2012). No obstante, dada las discrepancias terminológicas que el concepto implica para algunos académicos y activistas sociales, afirmando que éste alude exclusivamente a la homosexualidad (atracción romántica y sexual entre personas del mismo sexo o género) y, obvia las violencias que sufren otras diversidades

afectivo-sexuales, en el seno de la comunidad LGBTQ+, empezaron a surgir voces que reclamaban el establecimiento de un término que etimológicamente contemplase a todas las expresiones de la sexualidad sin hacer distinciones. Por ello, se creó el concepto de “LGTBIfobia” entendiéndose a éste como una forma de discriminación que afecta a todo un colectivo social y, que nace de la necesidad de aunar en una sola palabra a todas las diferentes manifestaciones de odio de las que son víctimas las personas LGBTQ+ (Observatorio Andaluz contra la Homofobia, Transfobia y Bifobia, 2019).

En cuanto a las formas que asume la homofobia, esta puede manifestarse por medio de ideas, percepciones, conceptos y estereotipos negativos, sentimientos de rechazo en aquellas situaciones en las que un sujeto tiene que relacionarse con una persona homosexual, ataques verbales o agresiones físicas y atentar contra la igualdad y derechos de las personas LGBTQ+ cuando está presente en el conjunto de leyes, organismos públicos y medios de comunicación (COGAM, 2005). Además, puede reproducirse en un contexto digital, consolidándose las redes sociales como el medio más empleado para ejercer ciberbullying LGTBIfóbico y, siendo las prácticas más comunes el envío de mensajes con contenido ofensivo, la difusión de imágenes denigrantes y la suplantación de identidad de la víctima con el fin de humillarla (COGAM, 2016), incluyendo otras formas de violencias virtuales como el acoso laboral, el outing (hacer pública la orientación sexual de una persona sin su consentimiento) y, hacerse pasar por homosexual y contactar con una persona LGBTQ+ fingiendo un interés afectivo para quedar a solas con esta con el propósito de robarle, chantajearle, agredirle físicamente o humillarlo grabando y difundiendo las imágenes en internet (Ojeda Melián, 2020).

### 3. LA RADICALIZACIÓN EN LÍNEA HACIA LAS PERSONAS LGBTQ+

La emergencia de grupos radicales de carácter conservador en España, en especial vinculados con la extrema derecha, ha supuesto un rápido incremento de las prácticas de odio en línea hacia minorías sociales como las personas LGBTQ+. De hecho, en el contexto internacional, los intentos por tratar de luchar contra el incremento del odio en las redes sociales parecen estar fracasando, tal como apuntó el Doctor Fernand de Varennes en el Informe A/77/246 del Relator Especial sobre Protección de los derechos de las minorías en las instituciones (ONU, 2022), conclusión que también constató el anterior Informe A/HRC/46/57 del Relator sobre cuestiones de las minorías (ONU, 2021) al afirmar que el 70 % de los delitos de odio que se cometen en los medios digitales se dirigen específicamente a minorías. Es más, el informe advierte de que estas se encuentran en una posición de especial vulnerabilidad en el uso de los servicios en línea y ello como consecuencia de las representaciones sociales que de ellas

se hace, convirtiéndose en chivos expiatorios de toda clase de teorías conspirativas, información errónea y desinformación que las representan como una amenaza y peligro, contribuyendo a la propaganda de la violencia en su contra. En este sentido, la rápida expansión del ciberespacio unida al vertiginoso auge de las redes sociales ha convertido a internet en la herramienta idónea para la incitación al odio. Y ello como consecuencia del mecanismo de difusión de bajo coste y alta velocidad que ofrecen este tipo de plataformas virtuales (Ștefăniță y Buf, 2021), facilitando así la propagación del discurso de odio, término referido a toda forma de comunicación, ya sea oral, escrita o a través del comportamiento y que implique un ataque o haga uso de un lenguaje claramente peyorativo o discriminatorio hacia una persona o un grupo por motivo de sus creencias religiosas, origen étnico, nacionalidad, etnia, ascendencia, género u otro factor de identidad (ONU, 2019), incluyendo además, según la Comisión Europea contra el Racismo y la Intolerancia en su Recomendación 15 de 8/12/2015, otras manifestaciones del mismo como la humillación, el acoso, descrédito, difusión de estereotipos negativos, estigmatización o amenaza, así como otras formas específicas de odio por razón de lengua, edad, discapacidad, orientación sexual e identidad de género (ECRI, 2016, p.18) y, que se caracteriza por materializarse a través de distintas formas de expresión como imágenes, dibujos animados, ilustraciones, memes, objetos, gestos y símbolos; es claramente discriminatorio y tiene un sentido ofensivo; se difunde tanto en internet como fuera de éste; y, se centra en factores de identidad, ya sean reales o percibidos, de un grupo o individuo (ONU, s.f). Y, en el caso del discurso de odio homofóbico o LGTBIfóbico en línea, el problema radica en que sus precursores tienden a normalizar este tipo de expresiones violentas bajo el pretexto de que todas las opiniones deber ser validadas y respetadas, amparándose en un supuesto derecho a poder expresarse de forma totalmente libre y sin censura. Este hecho es bastante preocupante, ya que el discurso del odio puede convertirse en la antesala de un potencial crimen de odio. Cuestión sobre la que ya reflexionaba el Observatorio Madrileño contra la LGTBIfobia (2021) al señalar que cuando alguien lanza un mensaje que fomenta, promueve o incita al odio contra una parte de la sociedad demonizándola en algún grado, si bien inicialmente puede causar indignación y temor entre la población destinataria, convencéndola de la peligrosidad del colectivo objeto de odio, puede finalmente alentar a ciertos individuos a atacar a sus miembros.

Por otro lado, cabe señalar que internet se ha convertido en un espacio de violencia verbal, amenaza, humillación, hostigamiento y persecución para las personas LGTBIQ+. Según Gámez-Guadix & Incera (2021) uno de cada cuatro jóvenes pertenecientes a minorías sexuales sufre discriminación en línea por no asumir los roles de género hegemónicos. De hecho, si nos remitimos al contexto español, el *Informe sobre Delitos e Incidentes de Odio elaborado por la Federación Estatal de Lesbianas, Gais, Trans, Bisexuales, Intersexuales y más* (2022) en su análisis del

perfil de discurso de odio destaca que el 15,3% de dichas expresiones se cometieron a través de las redes sociales y, recalca el incremento de los delitos de odio de naturaleza homófoba en los entornos digitales. Pero es que el *Informe sobre evolución de los delitos de odio en España (2021)*, elaborado por el Ministerio del Interior, reveló que los delitos de odio por razón de orientación sexual e identidad de género en internet y las redes sociales fueron la segunda causa de hechos delictivos más denunciada, con una variación del 87,5% en el año 2021 respecto al 2020. Por su parte, el Observatorio Madrileño contra la LGTBIfobia (2021) señaló en su *Informe de delitos de odio motivados por LGTBIfobia* que los incidentes ocurridos a través de internet y las redes sociales ascendieron a un 21,1%, consolidándose como el segundo escenario en el que mayor número de ataques homofóbicos se generan tras la calle. Y, si nos extrapolamos al marco internacional, según Cheng y Björkstén (2020), miembros de Acces Now, organización que vela por la defensa de los derechos digitales de los usuarios en situación de riesgo en todo el mundo, los ciberataques contra la comunidad LGTBIQ+ se intensificaron durante la pandemia global del COVID-19, siendo el acoso en línea la principal amenaza a la que debe hacer frente dicho colectivo en la esfera virtual.

#### 4. CONSECUENCIAS DEL DISCURSO DE ODIO HOMOFÓBICO EN LÍNEA

La peligrosidad de los discursos de odio en línea radica en su potencial de daño sobre las víctimas, ya que se difunden a través de grandes audiencias, por lo que los sentimientos de vulnerabilidad, humillación, ridículo y vergüenza pública que experimentan las víctimas tienden a magnificarse. Además, la naturaleza del discurso de odio en línea lo convierte en un fenómeno social descontrolado con un enorme potencial de daño, difusión y audiencia a causa del riesgo de viralización (Gagliardone, 2015). Y, a dichos factores se suman la permanencia de los contenidos, la itinerancia entre distintos sitios webs, el uso de pseudónimos, el anonimato y la transnacionalidad que caracterizan al ciberespacio, configurando una cultura del odio en internet hacia las minorías sociales (Cabo Isasi y García Juanetey, 2017) y, que tiene como resultado la radicalización de las formas de expresión y comunicación (Miró, 2016). Además, la experiencia de los propios usuarios en internet y su exposición a determinadas ideologías nocivas puede derivar en que estos acaben normalizándolas, generando determinadas creencias que son inculcadas en la audiencia y que animan a otros sujetos a reproducir contenidos de odio (Kilvington, 2021). Y, cuando estas expresiones incívicas prevalecen en el contexto digital y son alentadas por figuras públicas dejan de ser percibidas como un tabú, provocando en consecuencia una disminución del nivel de sensibilidad hacia el lenguaje hostil y de los niveles de empatía hacia ciertos grupos sociales (Bilewicz y Soral, 2020), por lo que dicha insensibilización unida

a la tendencia de normalización del abuso podría perpetuar la violencia en las redes sociales (Ștefăniță y Buf, 2021), lo que explicaría el hecho de que resulte más probable que las víctimas sufran más acoso y episodios de odio en línea que en un contexto *offline* (Keum y Miller, 2018). Incluso, características psicológicas como la psicopatía, la impulsividad y el comportamiento de búsqueda de emociones pueden ser el desencadenante de la incitación al odio y los comportamientos incívicos en línea (Sorokowski et al., 2020).

Respecto a las secuelas del discurso de odio en línea sobre la salud mental de las personas LGTBIQ+, los insultos y amenazas que reciben en internet pueden provocarles multitud de efectos psicológicos como sentimientos de ira, tristeza, ansiedad, estrés, vergüenza, aislamiento social, baja autoestima y pensamientos suicidas (Hubbard, 2020); trastornos del sueño y ataques de pánico (Nyman y Provozin, 2019); autculpabilidad por formar parte de la comunidad LGTBIQ+, lo que les lleva a aislarse para reducir dichos discursos o a cerrar sus cuentas de las redes sociales como un mecanismo de defensa (Ștefăniță y Buf, 2021). Y, dado que los individuos con diversidades afectivo-sexuales no normativas son más propensos a recibir mensajes de odio en comparación con las personas heterosexuales, en especial las personas trans, las cuales sufren en mayor medida los efectos más destructivos del odio online (Ștefăniță y Buf, 2021), podrían acabar desarrollando actitudes de homofobia internalizada, es decir, el sujeto, como consecuencia de interiorizar y normalizar la homofobia a la que continuamente se expone en los entornos virtuales, puede adoptar mecanismos de defensa y evitación ocultando su propia condición sexual (Ghafoori et al., 2019). Además, el discurso de odio online puede manifestarse a través de otras violencias virtuales como el cibercoso homofóbico, el cual puede provocarles una distorsión de su autoconcepto, dificultades de interacción social con terceras personas y un uso inadecuado de los dispositivos inteligentes debido a su necesidad de buscar apoyo emocional en internet como refuerzo para aliviar sus emociones negativas (Lee et al., 2019).

## 5. ACTIVISMO DIGITAL Y NUEVOS MOVIMIENTOS SOCIALES LGTBIQ+

Internet se ha convertido en un escenario de continua acción y transformación para los movimientos sociales, teniendo como resultado el nacimiento de una contracultura que trata de hacer frente a los discursos hegemónicos, en especial aquellos que fomentan y difunden el odio, la discriminación y la violencia hacia los grupos más desfavorecidos; el resurgimiento de un activismo digital que se reinventa cada día para atender las cambiantes y exigentes demandas de la sociedad; y, el fomento de una acción colectiva mediante la difusión y persuasión de acciones que han conformado una identidad pública y colectiva (Bernárdez, 2021), construyendo significados sobre situaciones ilegítimas o

injustas y realizando llamadas a la acción para la defensa de ciertos problemas sociales (Rebollo, 2021). De hecho, según Bernárdez (2021) las redes sociales permiten viralizar determinados temas que carecen de visibilidad en los medios de comunicación, convirtiéndose en burbujas de información para algunos sectores de la audiencia que se desconectan de la actualidad mediática. Es más, la autora afirma incluso que este tipo de plataformas crean oportunidades de expresar ideas y opiniones que de otra forma resultarían imposible de difundir, conformando nuevas vías para el derecho a la libertad de expresión y participación haciendo que la ciudadanía cobre un nuevo valor en términos de activismo y deliberación en asuntos públicos (Acosta y Lassi, 2019).

Y, en el caso de las personas LGTBQI+, la esfera digital ha favorecido la construcción de comunidades virtuales que afianzan, refuerzan y visibilizan su identidad. Un estudio realizado por Strand (2019) sobre el uso de Twitter por parte de las minorías sexuales ugandesas puso de relieve que, la exclusión sistemática de dichas comunidades de los lugares tradicionales de producción del discurso, tales como medios de comunicación y espacios físicos, convierte estas plataformas en una alternativa para la defensa de los derechos humanos y la autorrepresentación, exhibiendo su agencia y a los miembros de su red y contribuyendo de esa forma a mejorar las circunstancias de las minorías sexuales en un contexto hostil hacia ellas. También, proporcionan un foro en el que se generan debates en torno a cuestiones relacionadas con la homosexualidad, en concreto la inclusividad en el deporte y la masculinidad y la forma en la que ésta última debe ser entendida y practicada (Sherwood et al., 2020). Además, estos escenarios facilitan el uso del humor o la reapropiación de símbolos originalmente de naturaleza homofoba y sexista como estrategias de desarticulación de los discursos normativos (Zarea Bonfill, Tortajada Giménez y Caballero Gálvez, 2021). Incluso, este tipo de entornos generan cambios de estructura en las formas o patrones de comunicación de los movimientos LGTBQI+, teniendo como resultado la aparición de nuevos movimientos. Ejemplo de ello son las comunidades diverso-afectivas que componen el movimiento LGTBQI+ en Turquía, las cuales, mediante el uso de las redes sociales han creado espacios de comunicación social entre la esfera pública geográfica y su espacio privado, logrando captar seguidores y construir un escenario de acción para alcanzar una mayor visibilidad a la vez que usan estas plataformas para mantener la integridad del movimiento (Caliskan, 2021). Además, en el marco de referencia de la diversidad afectivo-sexual, el activismo online adquiere una gran repercusión, en especial en países donde la homosexualidad se criminaliza y las minorías sexuales sufren represión, ya que a través del lenguaje sus miembros transgreden y desafían los discursos de las estructuras hegemónicas de la sexualidad, visibilizando las formas en que la homofobia se manifiesta en la esfera sociocultural, creando oportunidades para que tanto personas LGTBQI+ como sus simpatizantes se desahoguen y expresen su frustración frente a aquellos que no respetan su condición, evidenciando cómo internet crea un contrapúblico que deconstruye las ideologías heteronormativas dominantes (Otu, 2021).

## 6. LA CONTRANARRATIVA COMO ESTRATEGIA DE SUBVERSIÓN DEL DISCURSO DE ODIOS EN LÍNEA

El discurso de odio en línea se consagra como una amenaza para la ciudadanía y el principio de libertad en el que se fundamentan las sociedades democráticas, ocasionando efectos muy perniciosos sobre la salud de las víctimas. De ahí que surjan iniciativas por parte de organizaciones sociales y plataformas de la sociedad civil que tratan de frenar la hostilidad en línea y, que interpelan a los Estados y las grandes compañías de internet sobre su obligación de adoptar medidas legales y de moderación de contenidos en las plataformas digitales. Sin embargo, el problema radica en que este tipo de políticas pueden producir efectos colaterales como la supresión tanto de las expresiones valiosas como de aquellas ilícitas, teniendo como resultado la dispersión de la incitación al odio en lugar de su reducción (Hangartner, 2021), pudiendo además restringir opiniones y disidencias, lo que ha llevado a la Organización de las Naciones Unidas a promover un discurso más positivo en el que impere la libertad de expresión como norma y, defender la idea de que cualquier restricción debe ser excepcional para evitar posibles daños al tiempo que garantiza la igualdad y participación pública (ONU, 2019). Según Donzelli (2021) el odio en línea puede contrarrestarse mediante tres estrategias principales: las sanciones legales, las normas de retiradas de contenidos y el contradiscurso. No obstante, la autora asegura que las dos primeras pueden tener un carácter censor, entrando en conflicto el valor democrático fundamental de la libertad de expresión, suponiendo en tanto un reto para la sociedad civil, mientras que la contranarrativa se constituye como la solución más atractiva por carecer de un carácter coercitivo.

La contranarrativa se caracteriza por ser un mensaje que ofrece una alternativa positiva a la propaganda extremista con el objeto de deconstruirla y deslegitimarla (Tuck y Silverman, 2016). De acuerdo con Jääskeläinen (2019) este tipo de contrapropaganda no desacredita necesariamente las creencias previamente establecidas, sino que deconstruye las narrativas de odio en las que se basa ofreciendo una forma diferente de pensar sobre los temas, satirizando la propaganda radical y socavando su credibilidad. Por lo que, dado que el discurso de odio en línea puede degradar a ciertos grupos sociales, la intervención de los usuarios resulta esencial para contrarrestarlo (Obermaier, Schmuk y Saleem, 2021; Iganski, 2020). Sin embargo, este tipo de contrarreplicas pueden avivar la polarización del discurso como resultado del enfrentamiento entre los usuarios que difunden el discurso de odio y aquellos que tratan de contrarrestarlo, convirtiendo a las redes sociales en un terreno de batalla para el conflicto partidista (Keller y Askanius, 2020) y la lucha por la imposición ideológica. Ello hace que la contranarrativa deba articularse con cuidado como consecuencia de ese aumento del discurso polarizado (López-Sánchez y Müller, 2021), y más

cuando puede ser malinterpretada por ciertos grupos o sectores de la población, pues, por ejemplo, no todos los usuarios entienden el humor y la ironía de la misma manera. Igualmente, un simple comentario articulado en forma de contranarrativa para promover valores de respeto, igualdad e inclusividad hacia las personas LGTBQ+ y subvertir la homofobia en internet puede ser percibido por determinadas personas como una ofensa.

Por ello, para contrarrestar eficazmente el odio en línea es necesario utilizar un contradiscurso no solo basado en evidencias y datos, sino apelar a emociones como la empatía y valores compartidos como la dignidad humana, la justicia, la solidaridad o la igualdad (Bazzaco et al, 2019), lo que hace que no todos los usuarios estén dispuestos a implicarse en este tipo de discursos contrahegemónicos debido al esfuerzo, dominio del tema y habilidades personales, emocionales y locutivas que exigen (Kalch y Naab, 2017; Jost, Ziegele y Naab, 2020).

## 7. EL USO DEL HUMOR Y LA REAPROPIACIÓN DE SÍMBOLOS COMO CONTRANARRATIVA HACIA LA HOMOFOBIA EN LÍNEA: EL CASO DE GAYSPER

Gaysper (Imagen 1) es un meme de internet y símbolo de la comunidad LGTBQ+ cuyo diseño se basa en el emoji del fantasma de Android 5.0, pero con un fondo que conmemora los colores de la emblemática bandera del arcoíris. Respecto a su nombre, éste surge de la unión de las palabras “gay”, anglicismo empleado para designar la homosexualidad masculina, y “Casper”, un personaje de animación protagonista de la historieta *Casper the Friendly Ghost* (Seymour Reit y Joe Oriolo, 1930) y la película infantil que lleva su nombre por título (Brad Silberling, 1995).

Imagen 1. Gaysper, el fantasma gay



El meme se originó y popularizó en España durante las elecciones generales del 28 de abril de 2019, a raíz de un tuit publicado por el partido político ultraderechista español VOX en su cuenta oficial de Twitter, a través del cual instaban a sus votantes a ejercer el voto mediante la propaganda “¡Qué comience la batalla!” (Imagen 2). Para ello, acompañaron dicho mensaje de una imagen en cuyo primer plano aparecía *Aragorn*, personaje de *El señor de los Anillos* (Peter Jackson, 2001) enfrentándose a una horda de orcos y trasgos, pero cuya apariencia fue intencionadamente modificada y sustituida por símbolos contrarios a la ideología del partido, entre los que aparecía el icónico fantasma como una referencia a los movimientos LGTBIQ+ junto al movimiento feminista, el republicanismo, el movimiento antifascista, el comunismo, el independentismo catalán y diversos medios de comunicación de izquierdas tales como *Cadena Ser*, el periódico *El País* y la cadena de televisión *La Sexta*.

Durante la celebración de las elecciones generales, la estrategia de VOX consistió en apelar a un discurso patriótico que ensalzaba los valores nacionales, promoviendo la idea de una España unida frente a aquellos que considera los indiscutibles enemigos del estado, en este caso, como la formación política los denomina, los grupos “feminazis”, el “lobby LGTBI”, la “dictadura comunista”, los “podemitas” y los “progres”. Por ello, creó al meme (Imagen 2) con el objeto de mostrarlo como una figura tétrica que representaba a la diversidad afectivo-sexual como una amenaza para el modelo de familia tradicional, el matrimonio, los valores conservadores de la sociedad, la moral católica, el sistema heteronormativo y el binarismo hombre-mujer, en un intento de denostar y deslegitimar no solo los derechos de las personas LGTBIQ+, sino también de hacer apología del discurso reaccionario biológico frente al cultural para mostrar a las minorías sexuales como contrarias al orden natural, acusándolas de adoctrinar a los menores de edad y de promover la homosexualidad y transexualidad. Sin embargo, lejos de fomentar la homofobia, tras los acontecimientos, el colectivo LGTBIQ+ se reapropió del meme y comenzó a utilizarlo como un símbolo identitario, difundiéndolo por internet mediante todo tipo de mensajes irónicos de agradecimiento a la formación política y, cuyo propósito era subvertir su discurso homofóbico, convirtiendo así al colorido meme en una poderosa contranarrativa mediante el refuerzo de valores prosociales sobre el respeto a la igualdad y la libertad afectivo-sexual.

Imagen 2. Tuit de la cuenta oficial de VOX



Fuente: Tomado de TN (2019).

Por otra parte, Zarea Bonfill, Tortajada Giménez y Caballero Gálvez (2021) aseguran que Gaysper ilustra cómo a partir de la creatividad empleada por el colectivo LGTBQ+ para difundir una imagen positiva, alegre y festiva de dicho meme se sabotó la imposición hegemónica de VOX. Incluso, según estos autores, el uso de imágenes para deconstruir su significado intrínseco permite a los usuarios copiar, modificar y cambiar su sentido original como un proceso en el que pierden su autoría y significado original. Pero, además, el alegre fantasma es todo un reflejo de la intervención cívica en línea al escenificar un conjunto de acciones emprendidas por los ciberusuarios corrientes para luchar contra los comportamientos disruptivos y de odio online (Porten-Cheé et al., 2021). De hecho, no es la primera vez que la comunidad LGTBQ+ se reapropia de un símbolo, expresión, idea o término de carácter homofóbico y lo dota de un nuevo significado inclusivo. Un ejemplo de ello es la descortesía simulada en contextos Drag Queens, que implica el uso lúdico de expresiones potencialmente descorteses por parte de sus miembros (Días Oliva, Marcelo Antonialli y Gomes, 2020) y, cuyo propósito es hacer frente a la hostilidad y la homofobia. Y lo realmente interesante de ello es que, si bien expresiones como “maricón”, “loca”, “bollera” o “tortillera”, entre otras, son en primera instancia insultos homofóbicos, siendo percibidos como descalificativos grotescos, ofensivos y humillantes fuera de un contexto LGTBQ+ por factores meramente culturales, estos pueden ser reapropiados por parte de los miembros de la comunidad para

autodenominarse, hacer humor de sí mismos y construir una identidad grupal con la que hacer frente a las narrativas estigmatizadoras.

Por último, en relación con el proceso de difusión de Gaysper en las redes sociales, bastaron menos de 24 horas para que Twitter se convirtiese en una amalgama de memes que empleaban al popular fantasma como un símbolo de resistencia frente a VOX. Es más, enseguida surgieron otros diseños alternativos del divertido personaje, y que encarnaban a las diferentes diversidades afectivo-sexuales con sus respectivas banderas representativas. Y es que tal fue su repercusión mediática que, el grupo parlamentario de extrema derecha se vio obligado a retirar el tuit original. Esto se debió igualmente a las presiones que recibió de Warner Bros por haber utilizado el fotograma de *Aragorn* sin su permiso e infringir los derechos de propiedad (*El Español*, 2019). También, comenzó a comercializarse todo tipo de merchandising con la estampa del fantasma, como camisetas, mochilas, gorras, pegatinas o tazas. Incluso, la imagen del icónico meme se trasladó al Congreso de los Diputados cuando dos políticos del Partido Socialista Obrero Español, en concreto Felipe Sicilia y Arnau Ramírez, usaron camisetas con la serigrafía de Gaysper durante la constitución de las Cortes del 21 de mayo de 2019. Pero, además, en Twitter apareció repentinamente el hashtag #gaysper, el cual se convirtió en Trending Topic en España con narrativas que giraban en torno a cuestiones como la solidaridad, la justicia social y el apoyo a las personas LGTBIQ+, la confrontación política contra VOX, la extrema derecha y el fascismo y, la defensa y reivindicación de la libertad sexual y del Movimiento del Orgullo LGTBIQ+, evidenciándose así la capacidad que tienen los hashtags para contribuir a la acción colectiva (Jackson et al., 2020, en Kilic, 2021), así como proyectar voces de grupos oprimidos y desarrollar movimientos políticos y reivindicaciones (Rebollo, 2021).

## REFERENCIAS

- Acosta y Lassi (2020). #8M 2019. La conversación digital durante la Huelga Internacional de Mujeres. *Revista Arbitrada Interdisciplinaria Koinonía*, 5(9). <https://doi.org/10.35381/r.k.v5i9.516>
- Bazzaco, E., Díaz, A., García Juanatev, Palacios, A., y Tarragona, L. (2019). CIBERESPECT. Guía práctica de intervención online para ciberactivistas. [https://www.idhc.org/arxius/recerca/1579268292-CIBERESPECT\\_Guia\\_practica\\_2019.pdf](https://www.idhc.org/arxius/recerca/1579268292-CIBERESPECT_Guia_practica_2019.pdf)
- Bernárdez-Rodal, A., López-Priego, N., Padilla-Castillo, G. (2021). Cultura y movilización social contra la violencia sexual a través de Twitter: el caso del fallo judicial “#LaManada” en España. *Revista Latina de Comunicación Social*, 79, 237-262. <https://www.doi.org/10.4185/RLCS-2021-1502>

- Bilewicz, M. y Soral, W. (2020). Hate speech epidemic. the dynamic effects of derogatory language on intergroup relations and political radicalization. *Political Psychology*, 41, 3-33. <https://www.doi.org/10.1111/pops.12670>
- Cabo Isasi, A. y García Juanatey, A. (2017). El discurso del odio en las redes sociales: Un estado de la cuestión. *Ajuntament de Barcelona progress report*. [https://www.injuve.es/sites/default/files/2019/02/noticias/el\\_discurso\\_del\\_odio\\_en\\_rrss.pdf](https://www.injuve.es/sites/default/files/2019/02/noticias/el_discurso_del_odio_en_rrss.pdf)
- Caliskan Ozgur (2021). Digital pride on the streets of the internet: Facebook and Twitter practices of the LGBTI movement in turkey. *Sexuality & Culture*, 25(4), 1447-1468. <http://dx.doi.org/10.1007/s12119-021-09818-9>
- Cheng, S. y Björkstén (29 de junio de 2020). Digital Security Helpline: In 2020, LGBTQ groups are facing more online harassment than ever. *Acces Now*. <https://www.accessnow.org/digital-security-helpline-in-2020-lgbtq-groups-are-facing-more-online-harassment-than-ever/>
- COGAM (2005). Homofobia en el Sistema Educativo. [https://eprints.ucm.es/35733/1/1437\\_es\\_Homofobia%20en%20el%20Sistema%20Educativo%202005.pdf](https://eprints.ucm.es/35733/1/1437_es_Homofobia%20en%20el%20Sistema%20Educativo%202005.pdf)
- COGAM (2016). Benítez, E. (Coord.). *Cyberbullying LGBT-fóbico. Nuevas formas de intolerancia*. Madrid: Grupo de Educación de COGAM. <https://cogameduca.files.wordpress.com/2016/03/4-ciberbullying-lgbt-fc3b3bico-informe-completo-web.pdf>
- Comisión Europea Contra el Racismo y la Intolerancia (2016, 21 de marzo). *Recomendación General N° 15 relativa a la lucha contra el Discurso de Odio y Memorandum explicativo*. <https://rm.coe.int/ecri-general-policy-recommendation-n-15-on-combating-hate-speech-adopt/16808b7904>
- Dias Oliva, T., Antonialli, D. M., & Gomes, A. (2021). Fighting Hate Speech, Silencing Drag Queens? Artificial Intelligence in Content Moderation and Risks to LGBTQ Voices Online. *Sexuality & Culture*, 25(2), 700-732. <https://doi.org/10.1007/s12119-020-09790-w>
- Donzelli, S. (2021). Countering harmful speech online. (In)effective strategies and the duty to counterspeak. *Phenomenology and Mind*, 20, 76-87. <https://doi.org/10.17454/PAM-2007>
- Corral, M. (2019, 4 de mayo). Éste es Gaysper, el icono homófobo creado por Vox que se le ha vuelto en contra. *El Español*. [https://www.elespanol.com/social/20190504/gaysper-icno-homofobo-creado-vox-vuelto/395460632\\_0.html](https://www.elespanol.com/social/20190504/gaysper-icno-homofobo-creado-vox-vuelto/395460632_0.html)
- Federación Estatal de Lesbianas, Gais, Trans, Bisexuales, Intersexuales (2022). Delitos e Incidentes de Odio. Informe 2022. [https://felgtbi.org/wp-content/uploads/2023/01/delitos-de-odio\\_felgtbi.pdf](https://felgtbi.org/wp-content/uploads/2023/01/delitos-de-odio_felgtbi.pdf)
- Ghafoori, B., Caspi, Y., Salgado, C., Allwood, M., Kreither, J., Tejada, J., Hunt, T., Waelde, L., Slobodin, O., Failey, M., Gilberg, P., Larrondo, P., Ramos, N., von Haumeder, A. y Nadal, K. (2019). *Global Perspectives on the Trauma of Hate-Based Violence: An International Society for Traumatic Stress Studies Briefing*

- Paper*. [https://istss.org/ISTSS\\_Main/media/Documents/ISTSS-Global-Perspectives-on-the-Trauma-of-Hate-Based-Violence-Briefing-Paper\\_1.pdf](https://istss.org/ISTSS_Main/media/Documents/ISTSS-Global-Perspectives-on-the-Trauma-of-Hate-Based-Violence-Briefing-Paper_1.pdf)
- Gamez-Guadix, M., & Incera, D. (2021). Homophobia is online: Sexual victimization and risks on the internet and mental health among bisexual, homosexual, pansexual, asexual, and queer adolescents. *Computers in Human Behavior*, 119, 106728. <https://doi.org/10.1016/j.chb.2021.106728>
- Gagliardone, I. (2015.) Countering Online Hate Speech. Programme in Comparative Media Law and Policy, University of Oxford. [https://www.researchgate.net/publication/284157227\\_Countering\\_Online\\_Hate\\_Speech\\_-\\_UNESCO](https://www.researchgate.net/publication/284157227_Countering_Online_Hate_Speech_-_UNESCO)
- Hangartner, D., Gennaro, G., Alasiri, S., Bahrigh, N., Bornhoft, A., Boucher, J., Demirci, B. B., Derksen, L., Hall, A., Jochum, M., Munoz, M. M., Richter, M., Vogel, F., Wittwer, S., Wuthrich, F., Gilardi, F., y Donnay, K. (2021). Empathy-based counterspeech can reduce racist hate speech in a social media field experiment. *Proceedings of the National Academy of Sciences of the United States of America*, 118(50). *Natl Acad Sciences*. <https://doi.org/10.1073/pnas.2116310118>
- Hubbard, L. (2020). Online Hate Crime Report: Challenging online homophobia, biphobia and transphobia. Galop, the LGBT+ anti-violence charity. [https://galop.org.uk/wp-content/uploads/2021/06/Online-Crime-2020\\_0.pdf](https://galop.org.uk/wp-content/uploads/2021/06/Online-Crime-2020_0.pdf)
- Iganski, P. (2020). Civil courage as a communicative act Countering the harms of hate violence. *Pragmatics and Society*, 11(2), 316-335. <https://doi.org/10.1075/ps.18075.iga>
- Jaaskelainen, T. (2020). Countering hate speech through arts and arts education Addressing intersections and policy implications. *Policy Futures in Education*, 18(3), 344-357. <https://doi.org/10.1177/1478210319848953>
- Jost, P., Ziegele, M., y Naab, T. K. (2020). Klicken oder tippen? Eine analyse verschiedener interventionsstrategien in unzi-vilen online-diskussionen auf Facebook. *Zeitschrift Für Politikwissenschaft*, 105(1), 231. <https://doi.org/10.1007/s41358-020-00212-9>
- Kalch, A., y Naab, T. (2017). Replying, disliking, flagging: How users engage with uncivil and impolite comments on news sites. *Studies in Communication Media*, 6(4), 395-419. <https://doi.org/10.5771/2192-4007-2017-4-395>
- Keum, B.T., y Miller, M.J. (2018). Racism on the Internet: Conceptualization and recommendations for research. *Psychology of Violence*, 8(6), 782-791. <https://doi.org/10.1037/vio0000201>
- Keller, N., y Askanius, T. (2020). Combatting hate and trolling with love and reason? A qualitative analysis of the discursive antagonisms between organized hate speech and counterspeech online. *SCM Studies in Communication and Media*, (4), 540-572. <https://doi.org/10.5771/2192-4007-2020-4-540>
- Kilic, O. (2021). “Every Parade of Ours is a Pride Parade”: Exploring LGBTI+ digital activism in Turkey. *Sexualities*, 0(0), 1-17. <https://doi.org/10.1177/13634607211060510>

- Kilvington, D. (2021). The virtual stages of hate: Using Goffmans work to conceptualise the motivations for online hate. *Media, Culture & Society*, 43(2) 256-272. <http://dx.doi.org/10.1177/0163443720972318>
- Lee, J.I, Yen, C.F, Hsiao, R. C., y Hu, H. F. (2019). Relationships of homophobic bullying during childhood and adolescence with problematic internet and smartphone use in early adulthood among sexual minority men in taiwan. <https://doi.org/10.1590/0101-60830000000203>
- Lopez-Sanchez, M., y Mueller, A. (2021). On Simulating the Propagation and Countermeasures of Hate Speech in Social Networks. *Applied Sciences-Basel*, 11(24). <https://doi.org/10.3390/app112412003>
- Ministerio del Interior de España (2021). Informe de Evolución de los Delitos de Odio en España. [https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/publicaciones-periodicas/informe-sobre-la-evolucion-de-los-delitos-de-odio-en-Espana/Informe\\_evolucion\\_delitos\\_odio\\_Espana\\_2021\\_126200207.pdf](https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/publicaciones-periodicas/informe-sobre-la-evolucion-de-los-delitos-de-odio-en-Espana/Informe_evolucion_delitos_odio_Espana_2021_126200207.pdf)
- Miró, F. (2016). Taxonomía de la comunicación violenta y el discurso de odio en internet. *Revista de internet, derecho y política* (22). <https://doi.org/10.7238/idp.v0i22.2975>
- Nyman, H., y Provozin, A. (2019). The Harmful Effects of Online and Offline Anti LGBTI Hate Speech. <https://lnu.diva-portal.org/smash/get/diva2:1355445/FULLTEXT02.pdf>
- Obermaier, M., Schmuck, D., y Saleem, M. (2023). I'll be there for you? Effects of Islamophobic online hate speech and counter speech on Muslim in-group bystanders' intention to intervene. *Media & Society*, 25(9), 2339-2358. Sage Publications LTD. <https://doi.org/10.1177/14614448211017527>
- Observatorio Andaluz contra la Homofobia, Bifobia y Transfobia. (1 de junio de 2019). *Qué es la LGTBfobia*. <https://observatorioandaluzlgbt.org/que-es-la-lgtbifobia/>
- Observatorio Madrileño contra la homofobia, lesbofobia, transfobia y bifobia (2021). *Informe de incidentes de odio motivados por LGTBfobia en la comunidad de Madrid*. <https://contraelodio.org/wp/wp-content/uploads/2022/11/Informe-ObsMadLGTBI-2021-compressed-1.pdf>
- Ojeda Melián, J. (2020, 25 de junio). *Delitos de odio LGTBfóbico en internet*. <https://dialogando.com.es/delitos-de-odio-lgtbfbicos-en-internet/#>
- ONU (2022). A/77/246: Informe del Relator Especial sobre cuestiones de las minorías, Fernand de Varenes. Protección de los derechos de las minorías en las instituciones, estructuras e iniciativas de las Naciones Unidas. <https://www.ohchr.org/es/documents/thematic-reports/a77246-protection-rights-minorities-institutions-structures-and>
- ONU (2021). A/HRC/46/57: Cuestiones relativas a las minorías Informe del Relator Especial sobre cuestiones de las minorías, Fernand de Varenes. <https://www.ohchr.org/es/documents/thematic-reports/ahrc4657-minority-issues-report-special-rapporteur-minority-issues>

- ONU (2019). *Discurso de odio frente a la libertad de expresión*. <https://www.un.org/es/hate-speech/understanding-hate-speech/hate-speech-versus-freedom-of-speech>
- ONU (2019, mayo). *Estrategia y Plan de Acción de las Naciones Unidas para luchar contra el Discurso de Odio*. [https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action\\_plan\\_on\\_hate\\_speech\\_ES.pdf](https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_ES.pdf)
- ONU (s.f). *¿Qué es el discurso de odio?*. <https://www.un.org/es/hate-speech/understanding-hate-speech/what-is-hate-speech>
- Otu, K. E. (2021). Queer slacktivism as silent activism? the contested politics of queer subjectivities on GhanaWeb. *Sexualities*, 24(1-2), 46-66. <https://doi.org/10.1177/1363460719893620>
- Parlamento Europeo (2012). *Resolución del Parlamento Europeo, de 24 de mayo de 2012, sobre la lucha contra la homofobia en Europa (2012/2657(RSP))*. [https://www.europarl.europa.eu/doceo/document/TA-7-2012-0222\\_ES.pdf](https://www.europarl.europa.eu/doceo/document/TA-7-2012-0222_ES.pdf)
- Porten-Cheé et al. (2021). Do “Good Citizens” fight hate speech online? Effects of solidarity citizenship norms on user responses to hate comments. *Journal of Information Technology & Politics*, 18:3, 258-273. <https://doi.org/10.1080/19331681.2020.1871149>
- Rebollo Díaz, C. (2021). *Tuiteando sobre refugiados: una comparación internacional de discursos, imaginarios y representaciones sociales*. [Tesis Doctoral, Universidad de Huelva, <https://rabida.uhu.es/dspace/handle/10272/20113>].
- Rodríguez-Macarro, M. (29 de abril de 2019). Gaysper, el fantasma gay del meme de Vox. *Diario de Sevilla*. [https://www.diariodesevilla.es/sociedad/gaysper-fantasma-LGTiB-Vox-Meme\\_0\\_1350165368.html](https://www.diariodesevilla.es/sociedad/gaysper-fantasma-LGTiB-Vox-Meme_0_1350165368.html)
- Sherwood, M., Donaldson, A., Dyson, S., Lowden, D., Marjoribanks, T., Shill, J., Bolam, B., Nicholson, M. (2020). Football, media, and homophobia: Public framing of the first pride game in the Australian football league. *Communication and Sport*, 8(4-5), 545-565. <https://rabida.uhu.es/dspace/handle/10272/20113>
- Sorokowski, P., Kowal, M., Zdybek, P., y Oleszkiewicz, A. (2020). Are online haters psychopaths? Psychological predictors of online hating behavior. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00553>
- Ștefăniță, O., y Buf, D. (2021). Hate speech in social media and its effects on the LGTB community: A review of the current research. *Romanian Journal of Communication and Public Relations*, 23(1), 47-55. <https://doi.org/10.21018/rjcp.2021.1.322>
- Strand, C. (2019). Navigating precarious visibility: Ugandan sexual minorities on Twitter. *Journal of African Media Studies*, 11(2), 229-256. [https://intellectdiscover.com/content/journals/10.1386/jams.11.2.229\\_1](https://intellectdiscover.com/content/journals/10.1386/jams.11.2.229_1)
- TN. (21 de mayo de 2019). Gaysper, el meme antigay de la ultraderecha española que se les volvió en contra en el Parlamento. [https://tn.com.ar/internacional/gaysper-el-meme-antigay-de-la-ultraderecha-espanola-que-se-les-volvio-en-contra-en-el-parlamento\\_964133/](https://tn.com.ar/internacional/gaysper-el-meme-antigay-de-la-ultraderecha-espanola-que-se-les-volvio-en-contra-en-el-parlamento_964133/)

- Tuck, H. y Silverman, T. (2016). The counter-narrative handbook. Institute for Strategic Dialogue. [https://www.isdglobal.org/wp-content/uploads/2016/06/Counter-narrative-Handbook\\_1.pdf](https://www.isdglobal.org/wp-content/uploads/2016/06/Counter-narrative-Handbook_1.pdf)
- Zaera Bonfill, A., Tortajada Gimenez, I., y Caballero Galvez, A. A. (2021). The reappropriation of the insult as queer resistance in the digital universe: The Gaysper case. *Investigaciones Feministas*, 12(1), 103-113. <https://doi.org/10.5209/infe.69684>



## GRACIAS POR CONFIAR EN NUESTRAS PUBLICACIONES

Al comprar este libro le damos la posibilidad de consultar gratuitamente la versión ebook.

### Cómo acceder al ebook:

- ☞ **Entre en nuestra página web**, sección Acceso ebook  
([www.dykinson.com/acceso\\_ebook](http://www.dykinson.com/acceso_ebook))
- ☞ **Rellene el formulario** que encontrará insertando el código de acceso que le facilitamos a continuación así como los datos con los que quiere consultar el libro en el futuro (correo electrónico y contraseña de acceso).
- ☞ Si ya es **cliente registrado**, deberá introducir su **correo electrónico y contraseña habitual**.
- ☞ Una vez registrado, **acceda a la sección Mis e-books de su cuenta de cliente**, donde encontrará la versión electrónica de esta obra ya desbloqueada para su uso.
- ☞ Para consultar el libro en el futuro, ya sólo es necesario que se identifique en nuestra web con su correo electrónico y su contraseña, y que se dirija a la sección Mis ebooks de su cuenta de cliente.



### CÓDIGO DE ACCESO

Rasque para ver el código

Nota importante: Sólo está permitido el uso individual y privado de este código de acceso. Está prohibida la puesta a disposición de esta obra a una comunidad de usuarios.



**MANTÉNGASE INFORMADO  
DE LAS NUEVAS PUBLICACIONES**

**Suscríbase gratis  
al boletín informativo  
[www.dykinson.com](http://www.dykinson.com)**

**Y benefíciense de nuestras ofertas semanales**

**E**l libro *Teorías de la conspiración y discursos de odio en línea en la sociedad de las plataformas* reúne un conjunto de textos en torno a las teorías de la conspiración y los discursos de odio en línea desde una perspectiva tanto teórica como aplicada. Sus autores exploran cómo se concretan las narrativas conspirativas y el odio en línea en conexión con la pandemia de COVID-19, los inmigrantes y/o refugiados, el odio y las ciberviolencias dirigidas principalmente a mujeres, o las narrativas, discursos e imaginarios sociales sobre personas LGTBIQ+.

Uno de los aspectos más novedosos de este libro es reunir contribuciones internacionales en torno a estos dos ejes temáticos que conectan entre sí, las teorías de la conspiración y los discursos del odio, textos escritos desde una diversidad de ciencias y de disciplinas que miran interdisciplinariamente su objeto de estudio, propiciando, con la estrategia de conjugar la teoría con los estudios de caso, una mejor comprensión de dos problemáticas que tienen una gran centralidad en las sociedades actuales y que a veces están marcando el ritmo de fenómenos sociales y políticos. El foco principal del libro, como se aprecia en su título, es examinar precisamente las narrativas conspirativas y los discursos de odio en el contexto de la sociedad de las plataformas presentando una atención especial a estos fenómenos cuando se despliegan en plataformas, como Twitter o X, Instagram, Telegram, u otras, así como a través de servicios de mensajería o páginas web tales a Facebook, Youtube, WhatsApp, amén de otras que se encuentran en Internet.

Este libro es uno de los productos que resultan de la ejecución del proyecto de I+D+i titulado “Teorías de la conspiración y discurso del odio online: Comparación de pautas en las narrativas y redes sociales sobre el COVID-19, los inmigrantes y refugiados y personas LGBTI [NON-CONSPIRA-HATE!]”, PID2021-123983OB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por “FEDER A way of making Europe”

