La toxicidad política en redes sociales

El caso de las elecciones del #28M

Uxía Carral

Cátedra Ideograma-UPF
de Comunicación Política y Democracia

Cátedra Ideograma-UPF
de Comunicación Política y Democracia

La **Cátedra Ideograma-UPF de Comunicación Política y Democracia** formaliza una colaboración cualificada, amplia y duradera entre la UPF y la empresa de comunicación pública **ideo**grama, potenciando las relaciones de la universidad con su entorno socioeconómico y tecnológico e incrementando, con la colaboración de la empresa, la oferta de actividades y de investigación en este campo del conocimiento.

La Cátedra plantea el desarrollo de diferentes iniciativas de **investigación, transferencia y difusión del conocimiento generado** en el ámbito de la **Comunicación Política e Institucional**. Su finalidad es hacer llegar sus reflexiones y propuestas a través de la organización de actividades de índole diversa dirigidas tanto a público especializado como al público en general (congresos, seminarios y/o jornadas; publicaciones propias —cuadernos, monográficos, etc.— y artículos e intervenciones en los medios de comunicación).

Este informe es el resultado de la tercera edición del **Premio Cátedra Ideograma – UPF**, que se concede de forma anual al mejor trabajo de fin de máster en el ámbito de la Comunicación Política e Institucional. El premio consiste en una remuneración económica y la posibilidad de colaborar con la Cátedra para expandir el trabajo realizado, permitiendo así potenciar y ampliar la visibilidad de las investigaciones en el ámbito de la Comunicación Política.

Dirección científica:
Dr. Carles Pont-Sorribes
Dr. Reinald Besalú
Sr. Aleix Martí-Danés

Diseño y maquetación:
Sra. Gemma Canalda

Corrección del texto:
Sr. Antonio Marcelo Beltrán

Contacto:
Cátedra Ideograma-UPF de Comunicación Política y Democracia
Universitat Pompeu Fabra
Departament de Comunicació
Roc Boronat, 138
08018 Barcelona
Tel.: +34 935 421 219
catedra.ideograma@upf.edu

BIOGRAFÍA DE LA AUTORA

Uxía Carral es Doctora en Comunicación por la Universidad Carlos III de Madrid (UC3M). Sus líneas de interés versan sobre la innovación digital en comunicación, la educomunicación y el humanismo tecnológico. Es miembro del grupo de investigación *'Analytics, Media and Public Engagement: Communication, Journalism and Technology Lab'* (UC3MediaLab) y participa en varios proyectos internacionales y nacionales. Colabora también con el proyecto educativo y divulgativo Learn to Check. Anteriormente, ha trabajado como docente en el departamento de Comunicación (UC3M), como periodista en medios de comunicación (Cadena SER) y plataformas de verificación (Newtral) y como asistente técnico en varios proyectos europeos. Autora de diversos artículos científicos, capítulos de libros y co-autora del libro '*Comunicación Europea: ¿A quién doy like para hablar con Europa?*' (Ed. Dykinson, 2019).

https://orcid.org/0000-0002-2329-3331

ABSTRACT

In addition to reinforcing polarisation, the habit of informing oneself through social media encourages the normalisation of aggressive rhetoric. Thus, emotional confrontation and deliberate misinformation overshadow the complexity of political issues, causing the conversation on social media to be saturated with toxicity. For the practical study of this phenomenon, the case of the regional elections in the Community of Madrid has been selected. The conversations held by the six main candidates and their parties (PP, PSOE, MásMadrid, Vox, UnidasPodemos and Cs) with their respective audiences on the social network X during the election campaign between 12 and 27 May 2023 have been chosen. Using a hybrid methodology composed of computational data mining and manual procedures, a total of 73,351 tweets were obtained. From this sample, we examined the possible trends that could predict toxic audience feedback to the messages issued by politicians. We also explored the struggle between electoral candidates and their parties in relation to the exercise of their leadership. And, finally, we have investigated the new dynamics of influence that have arisen on social media between political candidates and opinion leaders, looking for who induces audiences more in the generation of toxicity in the conversation.

RESUMEN

Además de reforzar la polarización, el hábito de informarse a través de las redes sociales propicia la normalización de la retórica agresiva. Así, la confrontación emocional y la desinformación deliberada eclipsan la complejidad de los problemas políticos, provocando que la conversación en redes sociales se encuentre saturada de toxicidad. Para el estudio práctico de este fenómeno se ha seleccionado el caso de las elecciones autonómicas de la Comunidad de Madrid. Se han escogido las conversaciones que mantienen los/as seis principales candidatos/as y sus partidos (PP, PSOE, MásMadrid, Vox, UnidasPodemos y Cs) con sus respectivas audiencias en la red social X durante la campaña electoral entre el 12 y el 27 de mayo de 2023. A través de una metodología híbrida compuesta por procedimientos computacionales de minería de datos y manuales se ha obtenido un total de tuits n = 73.351. De esta muestra, se han examinado las posibles tendencias que puedan prever una retroalimentación tóxica de la audiencia a los mensajes que emiten los políticos. Asimismo, se ha explorado la pugna entre las candidaturas electorales y sus partidos en relación con el ejercicio de su liderazgo. Y, finalmente, se ha indagado sobre las nuevas dinámicas de influencia surgidas en las redes sociales entre los candidatos políticos y los líderes de opinión, buscando quién induce más a las audiencias en la generación de toxicidad en la conversación.

01 Introducción p. 05

02 Metodología p. 07

03 Resultados p. 11

04 Conclusiones p. 21

Bibliografía p. 23
Anexo p. 27

01
Introducción

El presente informe analiza la toxicidad en la conversación política en España en la red social X, conocida anteriormente como Twitter. La investigación nace del trabajo de fin de máster «**¿Cuán tóxica es la conversación sobre política española en Twitter? Estudio de este fenómeno a través de la IA**», el cual fue galardonado con el Premio Cátedra Ideograma - UPF al mejor trabajo de fin de máster en comunicación política e institucional.

La toxicidad puede entenderse como todo aquel contenido descontextualizado, manipulado o de autor impostado creado deliberadamente para confundir o hacer daño. Este fenómeno, cuyas consecuencias damnifican al individuo y al colectivo más allá de la dimensión virtual, se propaga preferiblemente por las redes sociales y mensajería instantánea, pues son los canales más efectivos para la difusión en masa (Casero-Ripollés *et al.*, 2016; Sánchez-Duarte & Magallón-Rosa, 2023). Por ello, plataformas sociales como X (Twitter), concebidas como los grandes canales de la democratización comunicativa, han pasado a ser consideradas, una década después de su nacimiento, escenarios vinculantes a la creación de todo tipo de amenazas para los sistemas democráticos occidentales.

En ese contexto, la política se ha erigido como una de las temáticas que suscita mayor crispación a la hora de discutir en redes sociales. El debate político siempre ofrece un sustrato fértil para la difusión de diversos tipos de toxicidad, ya que la polarización política actúa como un catalizador para la desinformación y su beneficio económico, impulsando narrativas engañosas con el objetivo de movilizar a su base de seguidores o desacreditar a sus oponentes (Renee, 2018). Los votantes pueden llegar a percibir a sus oponentes políticos como enemigos, lo que dificulta la cooperación entre distintos grupos políticos y amplifica los discursos de odio. La desinformación influye también en las preferencias de voto, altera la percepción de la legitimidad de los resultados y, por ende, socava la confianza en el sistema electoral, en los representantes y en las instituciones democráticas (Benkler *et al.*, 2018).

En el marco de un escenario mediático de tal calado, el nivel de infoxicación se incrementa aún en mayor magnitud durante la época de campaña electoral. Surge entonces una amalgama de actores (líderes políticos, partidos, medios de comunicación, líderes de opinión y usuarios prosumidores) que contribuye a la creación de un tejido discursivo cargado de agresividad. El uso recurrente de lenguaje ofensivo por parte de estos agentes no solo busca imponer sus puntos de vista, sino que también pretende desestabilizar el terreno del debate público, exacerbando la saturación informativa con contenido contradictorio y polarizado, que crea confusión y desalienta la participación *online* y *offline* (Miller *et al.*, 2022). En respuesta a este ambiente tóxico *online*, se han manifestado una serie de iniciativas que pretenden contrarrestar el impacto de diversas tipologías de toxicidad (desinformación, discursos de odio o acoso, entre otros).

Por una parte, existe una corriente de verificación de hechos, principalmente representada por el *fact-checking*, para identificar y desacreditar afir-

Los votantes pueden percibir a sus oponentes políticos como enemigos, lo que dificulta la cooperación entre grupos políticos y amplifica los discursos de odio

maciones falsas o engañosas; y por la colaboración entre organizaciones de verificación y plataformas en la etiquetación de contenido dudoso, para la promoción de la credibilidad informativa (Graves & Anderson, 2020). Por otra parte, la implementación de políticas más estrictas en redes sociales —como puede ser la censura de contenidos ofensivos, o la expulsión de usuarios reincidentes en sus comportamientos negativos (*deplatforming*)— también ha jugado un papel fundamental en la mitigación de la toxicidad *online*, buscando equilibrar la libertad de expresión con la responsabilidad en la difusión de información (Rogers, 2020).

Estas medidas, aunque no exentas de controversia, buscan promover una atmósfera *online* más segura, informada y propicia para un debate constructivo. Igualmente, desde el mundo académico se ha tratado de innovar en la manera de investigar estos nuevos acontecimientos digitales. De esta manera, se examinan las novedades en el campo periodístico, como la implementación de IA en las diferentes etapas del proceso informativo; pero también se aplica la innovación en el estudio del marco metodológico. Además de los procedimientos puramente cualitativos o cuantitativos convencionales (encuestas, entrevistas, análisis de contenido, de discurso), hoy en día existen nuevas técnicas de análisis de redes y, especialmente, de minería de datos, que permiten la investigación de nuevos fenómenos comunicativos como la toxicidad conversacional en plataformas *online*.

Concretamente, en el ámbito español, los medios de comunicación tradicionales han desempeñado un papel fundamental en la perpetuación de la polarización política en la opinión pública a través de tácticas como la segmentación de audiencias, el sensacionalismo o el sesgo mediático (Berrocal-Gonzalo *et al.*, 2023). No obstante, los propios efectos de la polarización afectan a la sociedad en su consumo, aumentando la desconfianza en los medios convencionales y acrecentando el hábito de informarse a través de las redes sociales. De hecho, en 2023, el 50 % de las noticias consultadas en España provenían de fuentes encontradas en las redes sociales (*Digital News Report*, 2023: 99). En este ambiente virtual, los usuarios se sumergen en argumentos que refuerzan sus propias perspectivas, propiciando además la normalización de la retórica agresiva. Así, la confrontación emocional y la desinformación deliberada eclipsan la complejidad de los problemas políticos, provocando que la conversación política en redes sociales se encuentre saturada de toxicidad.

Por ello, para el estudio práctico de este fenómeno se ha seleccionado el caso de las elecciones autonómicas de la Comunidad de Madrid, celebradas el 28 de mayo de 2023. Se han escogido las conversaciones que mantienen los/as seis principales candidatos/as y sus partidos con sus respectivas audiencias en la red social X. De este modo, la presente investigación pretende examinar las posibles tendencias que puedan prever una retroalimentación tóxica de la audiencia a los mensajes que emiten los políticos. Asimismo, se explora la pugna entre los candidatos electorales y sus partidos en relación con el ejercicio de su liderazgo. Y, finalmente, se indaga sobre las nuevas dinámicas de influencia surgidas en las redes sociales entre los candidatos políticos y los líderes de opinión, examinando quién induce más a las audiencias en la generación de toxicidad en la conversación.

02
Metodología

Este estudio analiza la calidad de la conversación *online* de los candidatos políticos en la Comunidad de Madrid con sus audiencias durante el período de la campaña electoral #28M. Para ello, se recurre a un análisis de los comportamientos en redes sociales a partir de las siguientes preguntas de investigación:

PI 1: ¿Existe alguna tendencia que pueda prever una retroalimentación tóxica de la audiencia a los mensajes que emiten los candidatos políticos?

A través de un detallado análisis de diferentes atributos de toxicidad, los resultados permiten estudiar la relación entre los políticos y sus audiencias (sean o no seguidores). Se presta especial atención a las posibles conexiones tóxicas que existen entre los líderes políticos y sus audiencias según factores temáticos, de género e ideológicos, con el objetivo de encontrar algún patrón común que alumbre los motivos de la producción de toxicidad.

PI 2: ¿Quiénes generan mayores niveles de toxicidad al ejercer su liderazgo político?, ¿los candidatos o los partidos?, ¿los candidatos o los líderes de opinión?

Se realizan dos análisis comparativos para con los datos de la PI 1: los candidatos contra los partidos políticos, y contra los líderes de opinión.

Por una parte, se analizan en términos de toxicidad las interacciones de las organizaciones políticas con sus usuarios de modo que, posteriormente, se produce la comparación con los datos de sus respectivos candidatos. De esta manera, se pretende dar a conocer el poder de liderazgo que ejerce cada agente durante la campaña electoral y su implicación en las situaciones de mayor toxicidad.

Por otra parte, se examina cuál es la dinámica dialéctica (horizontal o vertical) en donde se crea un mayor grado de toxicidad. Conocer este factor concreto permitirá discernir si ciertas conductas negativas son provocadas voluntariamente por cuentas intermediarias (líderes de opinión) que necesitan crear esas dinámicas para conseguir sus objetivos o, por el contrario, si esos entornos de toxicidad son el resultado de los modos de interacción naturales de los usuarios comunes en respuesta a los mensajes de los candidatos políticos.

Con el fin de responder a los anteriores interrogantes, hemos seleccionado una muestra total de tuits (n = 73.351) que se corresponden con el mensaje inicial de cada político y cada partido (tabla 1) y su cascada de respuestas, respectivamente. Esta recopilación pertenece al período de la campaña electoral de los comicios autonómicos, concretamente entre el 12 y el 27 de mayo de 2023. En cada jornada seleccionamos una única publicación de cada uno de los miembros de la muestra (candidatos y partidos); en concreto, aquella con el mayor número de contestaciones recibidas por parte de la audiencia. Decidimos centrar el foco en un tuit diario, y en el que más respuestas obtiene, porque el interés de la investigación no recae sobre el mensaje del político, sino en la naturaleza de las

Este estudio analiza la calidad de la conversación *online* de los candidatos en la Comunidad de Madrid con sus audiencias durante la campaña electoral #28M

TABLA 1. CANDIDATOS Y PARTIDOS POLÍTICOS QUE COMPONEN EL OBJETO DE ESTUDIO

Candidato	Partido político
Isabel Díaz Ayuso @IdiazAyuso	Partido Popular (PP) @ppmadrid
Juan Lobato @juanlobato_es	Partido Socialista Obrero Español (PSOE) @psoe_m
Mónica García @Monica_Garcia_G	Más Madrid @MasMadrid__
Rocío Monasterio @monasterioR	VOX @madrid_vox
Alejandra Jacinto @AleJacintoUrang	Podemos @PodemosCMadrid
Aruca Gómez @ArucaGomez	Ciudadanos @CS_Madrid

Fuente: Elaboración propia.

contestaciones de cada publicación. Estudios previos (Lampe & Johnston, 2005) demuestran que las contestaciones «aumentan las posibilidades de que se inicie una conversación y se cree una comunidad de usuarios comprometidos y motivados para futuras participaciones» (Burke *et al.*, 2010: 4).

Precisamente por la naturaleza y por la profundidad de los datos que se pueden extraer de la muestra, esta investigación utiliza metodología de análisis de redes sociales, más conocida por sus siglas en inglés, Social Network Analysis (SNA), con el fin de facultar una comprensión más completa de los fenómenos que se crean en las plataformas *online*. Esta metodología se centra en el estudio de la estructura y las relaciones dentro de las redes sociales (Teves & Cueto, 2020), utilizando herramientas como *softwares* de análisis de redes para visualizar y analizar las conexiones entre usuarios, identificar actores clave, medir la centralidad y estudiar la difusión de información.

Concretamente, dentro del SNA se pretende destacar la minería de datos (Pérez-López & Santín-González, 2007), puesta en práctica, en este caso, a través de dos herramientas: Communalytic y Perspective API. Esta técnica implica el uso de algoritmos y técnicas de aprendizaje automático para extraer información valiosa de grandes conjuntos de datos de redes sociales, lo cual puede incluir la detección de patrones, la clasificación de contenido y la formulación de predicciones basadas en datos históricos; y, en esta investigación, su valioso poder reside también en la fácil comprensión que se obtiene de sus resultados a simple vista (Lozano-Blasco *et al.*, 2023; Carral & Elías, 2024).

Primeramente, se procedió a compilar el material de estudio a través de **Communalytic**, un instrumento que recopila todos los datos sobre el conjunto de las publicaciones y combina las variables sobre toxicidad y otros elementos de la conversación (conexiones, contenido, frecuencias, etc.). De esta manera, con la previa autorización para la utilización de la API de X (que entonces aún se llamaba Twitter), y mediante esta herramienta de investigación en ciencias sociales computacionales (Gruzd *et al.*, 2022), se recuperaron 180 conjuntos de datos, uno por cada tuit diario de cada político y partido con sus respectivas respuestas.

A continuación se ejecutó el análisis de toxicidad a cada conjunto de datos, a través de la vinculación de Communalytic con **Perspective API**, la segunda herramienta utilizada en este análisis. Surge en el año 2016 como producto de una investigación colaborativa entre Jigsaw y el equipo de Counter Abuse Technology de Google. Este instrumento utiliza modelos de aprendizaje automático que identifican comentarios abusivos, y califican una frase en base a su nivel de «toxicidad», definida como el impacto percibido que un mensaje pueda tener en un individuo como para que decida abandonar la conversación.

En ese sentido, Perspective API ofrece diferentes atributos para medir el lenguaje tóxico dependiendo del idioma al que deba enfrentarse la herramienta. Desde que se comenzó a desarrollar en 2018 la herramienta en español, se pueden clasificar los comentarios de los usuarios en seis atributos, que se definen de la siguiente manera:

- **toxicidad**: publicación con lenguaje ofensivo leve u obsceno;
- **toxicidad severa**: comentario de incitación al odio o que resulta agresivo;
- **ataque identitario**: mensaje dirigido expresamente por razones de identidad;
- **insulto**: comentario incendiario y negativo contra un individuo o un grupo;
- **blasfemia**: contenido que incluye un lenguaje profano o ataque a lo sagrado;
- **amenaza**: respuesta con intención de infligir dolor o violencia contra otro(s).

La metodología Social Network Analysis (SNA) utiliza herramientas como softwares de análisis de redes para visualizar y analizar las conexiones entre usuarios, identificar actores clave, medir la centralidad y estudiar la difusión de información

Para cada atributo, las puntuaciones proporcionadas representan una probabilidad, con un valor entre 0 y 1. Una puntuación más alta indica una mayor probabilidad de que un lector perciba que el comentario contiene el atributo dado. A modo ilustrativo, de un comentario con una puntuación de 0,92 de lenguaje tóxico se extrapola que en torno a nueve de cada diez usuarios percibirán la publicación como tóxica. En consecuencia, este nivel de toxicidad considerablemente alto podría provocar que otro usuario decidiese no participar en esa conversación.

Por tanto, la elección de un umbral adecuado depende de cada uso específico. Esta vez hemos escogido un umbral de 0,7 para aplicar en este caso práctico, puesto que, dado el tamaño de la muestra, cabía la posibilidad de supervisar todo el conjunto de datos y solventar los sesgos más clamorosos que pudiesen resultar de haber escogido una mayor experimentalidad. Además, trabajos previos que han usado la herramienta en Twitter (Hua *et al.*, 2020; Pascual-Ferrá *et al.*, 2021), han seleccionado todos ellos un umbral semejante.

Por último, cabe mencionar que ha existido implicación humana en diferentes etapas del proceso de investigación: por una parte, se ha llevado a cabo una supervisión humana sobre los resulta-

dos obtenidos de las herramientas; y, por otra, se ha realizado un pequeño análisis de contenido y cálculos estadísticos descriptivos sobre los datos proporcionados por los instrumentos. En este contexto, se aprovecha para mencionar que se ha elegido mostrar el valor de la mediana en lugar de la media, puesto que informa mejor del punto central de la distribución al verse esta alterada por los promedios extremos de Aruca Gómez, completamente desiguales a sus adversarios.

LIMITACIONES:

Exponiendo el carácter experimental —fuera de su uso en inglés— que aún presentan algunas de las herramientas utilizadas, es de esperar que cuenten con ciertas limitaciones que detallamos a continuación:

En primer lugar, Communalytic muestra una restricción temporal para su uso de manera gratuita, ya que las descargas de mensajes no se pueden realizar ni anteriormente ni posteriormente a los siete días previos a la fecha actual, lo cual acota la cantidad de respuestas que un tuit puede recibir a la recolección de los comentarios emitidos durante ese rango temporal. Igualmente, las limitaciones logísticas impiden el análisis simultáneo de varias cuentas de Twitter, de manera que se complica que el objeto de estudio pueda tener una mayor envergadura, pues la recopilación de datos es parcialmente manual y el retraso en la descarga de datos implica una demora en el análisis.

Por su parte, Perspective API implementa en el análisis de los mensajes ciertos sesgos mayoritariamente debidos al conjunto de datos del que se alimenta esta herramienta de aprendizaje automático, que, a su vez, contienen las inclinaciones cognitivas que la sociedad carga sobre ellos con su manera de expresarse. En consecuencia, la propia plataforma devuelve esos sesgos a la hora de la categorización y la detección de comentarios tóxicos.

Asimismo, el modelo de aprendizaje automático funciona en base a conductas asociadas, por lo que tampoco se detecta toxicidad en comentarios con errores de deletreado o letras separadas por signos ortográficos, o en mensajes escritos de manera irónica o con ciertas palabras de argot o dialectos. Se debe resaltar que el análisis solamente comprende unidades textuales, por lo que fotografías, gifs, emoticonos y contenido en audio y/o vídeo quedan fuera del alcance de la inspección.

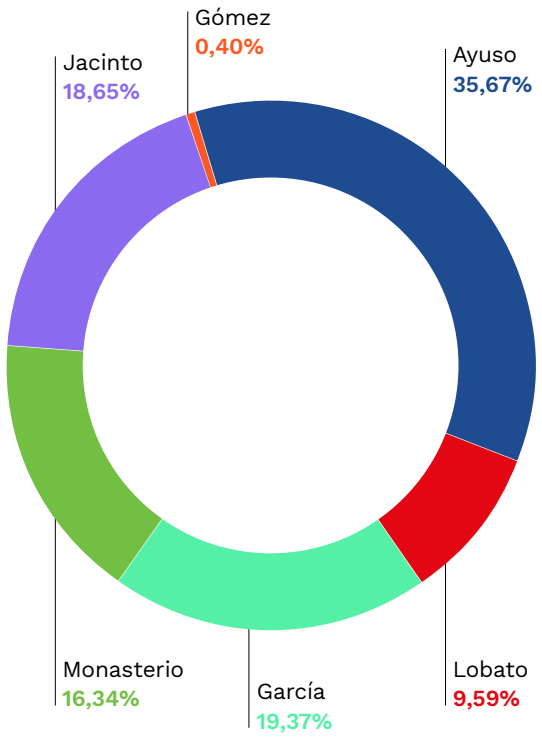
Para mitigar estos sesgos no deseados, las nuevas versiones de Perspective API han adoptado en su desarrollo iniciativas como aprender de las puntuaciones sugeridas, crear nuevos métodos de evaluación de los sesgos, y asociarse a colaboradores que proporcionan conjuntos de datos que ayudan a disminuir la sub-/sobrerrepresentación.

03
Resultados

Se parte de la premisa de que los seis candidatos que se presentaban a las elecciones autonómicas de la Comunidad de Madrid utilizaron sus perfiles en la red social hoy llamada X durante la totalidad de las jornadas de campaña electoral. Teniendo en cuenta este factor, se observa una clara diferencia en el nivel de *engagement*, medido en contestaciones, que sus publicaciones obtuvieron por parte de los usuarios.

Como se muestra en la figura 1, Isabel Díaz Ayuso (PP) se desmarca de sus rivales, consiguiendo más de un tercio de las respuestas recogidas en la muestra de este análisis. En sentido contrario, Juan Lobato (PSOE) no alcanza el umbral del 10 % de comentarios, y menor aún es el porcentaje de la candidata de Cs, Aruca Gómez, a quien iba dirigido menos del 1 % de los mensajes.

FIGURA 1. DISTRIBUCIÓN DE LA MUESTRA DE TUITS DE LOS CANDIDATOS.

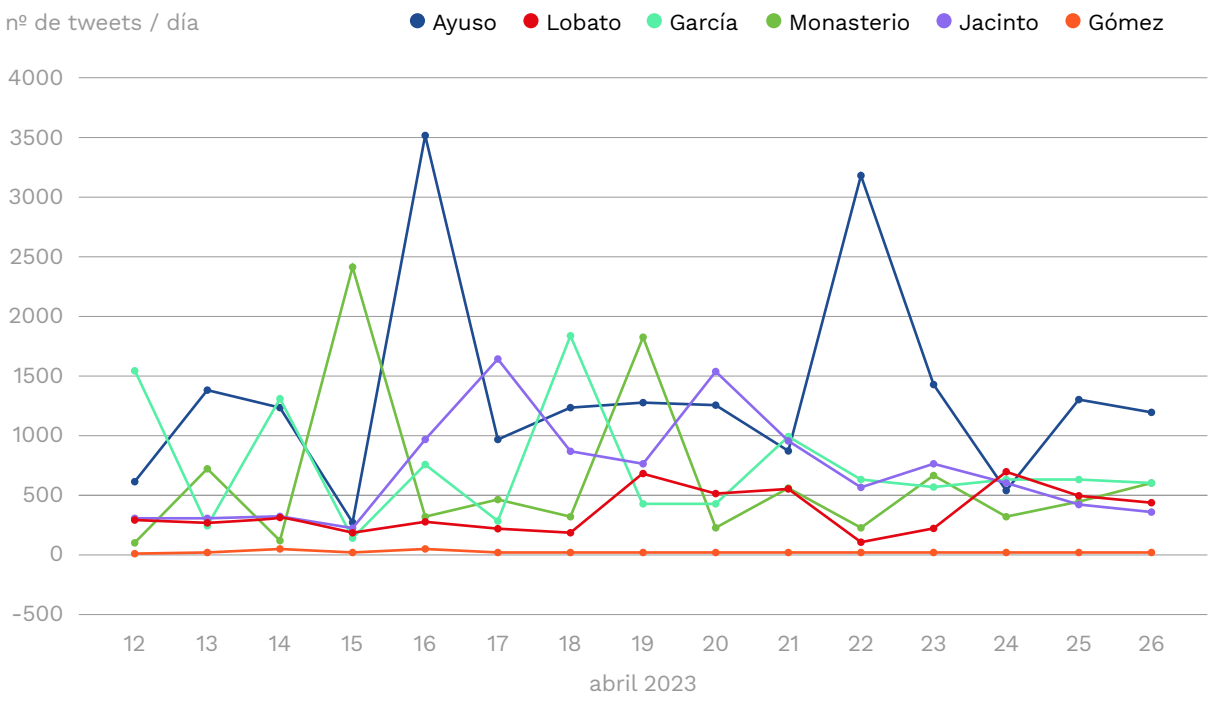


Fuente: Elaboración propia.

Estos datos globales proponen una rápida previsión de las métricas sobre el impacto diario que causan las publicaciones de cada político, que se confirma en la figura 2. De esta manera, los comentarios que mencionan a Ayuso, presidenta autonómica vigente, son los más populares hasta en seis ocasiones: 13, 23, 25, 26 de abril y, especialmente, 16 y 22, cuando las respuestas superaron la barrera de los tres mil mensajes en ambas jornadas. A esta notoriedad temporal de la líder popular, le siguen los mensajes de Mónica García (Más Madrid), que despuntaron en cuatro días, el doble de lo que precisaron los que apelaban a Rocío Monasterio (Vox) o Alejandra Jacinto (Podemos). Por su parte, la conversación del socialista Juan Lobato logró impactar en el mayor número de usuarios al menos el día 24 de abril, aunque con una escasa diferencia respecto a sus oponentes. Y en términos de visibilidad mínima se ha encontrado constantemente la candidata de Ciudadanos, Aruca Gómez, quien obtiene los peores resultados en cuanto a retroalimentación a lo largo de toda la campaña electoral.

Debido a la heterogeneidad de los resultados en cuanto al impacto temporal en relación con la variedad de las temáticas de conversación, no se puede dirimir que existan materias de discusión concretas que generen un mayor interés común entre los usuarios, y que, por tanto, se produzca un volumen de comentarios destacable. Sin embargo, se han obtenido nociones colectivas, en la fase de análisis de los mensajes recibidos por los candidatos, que ponen al descubierto nuevas reflexiones sobre la toxicidad que desprende el debate político *online*.

FIGURA 2. IMPACTO DIARIO DE LAS PUBLICACIONES DE LOS CANDIDATOS.



Fuente: Elaboración propia.

PI 1: CANDIDATOS Y SUS AUDIENCIAS

Existe, en un estado general de la cuestión, una tendencia a la conversación embarrada, pues en la totalidad de las jornadas de campaña electoral, los debates iniciados en los perfiles de todos los líderes políticos cuentan con comentarios considerados blasfemia; es decir, la ofensa no se dirige directamente hacia la persona política, pero los usuarios perciben cierta obscenidad en el contexto de una discusión cívica debido a las expresiones y ataques verbales contra elementos sagrados o religiosos. De hecho, todos los candidatos poseen en sus hilos de conversación picos diarios relativos a la profanidad que ascienden desde el 5,5 % de las respuestas hasta el máximo del 16,7 % en el caso de la dirigente de Ciudadanos.

Asimismo, otra noción común a todos los líderes autonómicos examinados señala a los insultos como el atributo tóxico con un valor más elevado entre los porcentajes de las seis categorías analizadas. En cifras, este hecho representa que, de media, el 5,85 % de los comentarios contiene insultos. Además, se trata de un suceso cotidiano: estos agravios contra los políticos se producen a diario independientemente de la temática de conversación o de los agentes interpelados en ella. Desde la perspectiva individual, cabe resaltar que este umbral genérico es superado en las cuentas de Monasterio (7,14 %), Jacinto (6,83 %) y Ayuso (5,95 %). Especialmente, los insultos contra las líderes de Vox y de Podemos se presentan como elementos protagonistas de sus hilos en varias jornadas, superando la barrera del 10 % de las contestaciones.

No obstante, el pico máximo de insultos (15,71 % de los comentarios) se registra en la discusión de Lobato con fecha del 13 de abril de 2023 (imagen 1). El mensaje raíz del que se inicia la conversación del socialista hace referencia a una elección entre el bienestar de las familias madrileñas o Ayuso, una dicotomía mostrada tanto en el texto como en el vídeo que acompaña al tuit. Se subraya esta estructura porque se adecúa al modelo de mensaje

electoral que también siguen las publicaciones más tóxicas del resto de los candidatos autonómicos.

Por una parte, el texto presenta a un adversario al que se culpa de un problema o se le opone contra una solución propuesta por el emisor. Tanto García, Jacinto como Gómez mencionan a Ayuso como lo hizo Lobato, mientras que la propia Ayuso se contrapone a Sánchez, presidente del Gobierno de España, y Monasterio alude de manera despectiva al antiguo líder de Podemos, Pablo Iglesias. Por otra parte, el mensaje siempre va acompañado de contenido audiovisual (vídeo, fotografía o enlace a una noticia) que funciona como elemento reforzador de las opciones dialécticamente desequilibradas entre las que debe elegir el votante (anexo 1).

Además de los insultos, el rasgo más característico de las contestaciones a los políticos madrileños apunta hacia la toxicidad leve, entendida como un conjunto de potenciales descalificativos, de un nivel de agravio bajo, dirigidos hacia el respectivo candidato que ha iniciado la conversación. Pese al sentido baladí —incluso humorístico— de las ofensas, la existencia de estas respuestas implica que existe la posibilidad de que otros usuarios decidan abandonar la conversación, al encontrarse con comentarios que perciben como altamente tóxicos. En ese sentido, los resultados del análisis indican que los debates surgidos en las cuentas de Monasterio (5,73 %), Jacinto (4,95 %) y Lobato (4,35 %) rebasan la mediana que indica que colectivamente el 4,27 % de los comentarios contienen toxicidad leve.

IMAGEN 1. TUIT CON MAYOR TASA DE TOXICIDAD.

Fuente: @juanlobato_es.

Por el contrario, la tasa de toxicidad severa no alcanza en ningún líder político el umbral del 1 % de los comentarios, ya que apenas se han encontrado tuits en cuyo texto se agreda verbalmente de manera incisiva o se incite al odio. En ese mismo sentido, tampoco se ha detectado un impacto significativo en las respuestas de la audiencia que induzcan a posibles amenazas dirigidas a los candidatos de la Comunidad de Madrid. De nuevo, ningún líder llega siquiera a rozar la barrera del 1 % de los comentarios en relación con este atributo. Igualmente, este ínfimo porcentaje se repite cuando se examina la categoría referida a los ataques de identidad.

Según factores de género

Los resultados del análisis exhaustivo sobre las conversaciones de los candidatos autonómicos también muestran una conclusión genérica, que, además, respalda investigaciones previas. Estudios como Hua *et al.* (2020: 278) sugieren que «el género, por ejemplo, no es un factor significativo para los candidatos» en cuanto a predictor de interacciones adversas. De igual manera, se puede afirmar en este trabajo que en las seis categorías de toxicidad concernientes a todos los miembros políticos que conforman la muestra no se advierte ningún comportamiento particular por razones de género. Esta deducción apoyaría, además, la teoría defendida por autores como Hayes & Lawless (2015) y Murray *et al.* (2012), en las que se expresa que los votantes no están predispuestos contra la elección de mujeres.

Frente a esta corriente, sin embargo, la hipótesis según factores de género se sigue planteando como una gran preocupación y, por ello, es motivo de investigación, tanto entre diversas instituciones internacionales (IPU, 2016; OAS, 2015) como para los propios académicos (Ballington, 2018; Bardall, 2013; Bardall *et al.*, 2020; Bjarnegård, 2018). Y en ambos casos se ha llegado a una misma conclusión: la violencia política parece estar dirigida más intensamente hacia las mujeres que hacia los hombres, lo que plantea otro obstáculo grave para la consecución de la igualdad política de género. Instituciones como las Naciones Unidas (2021: 12) inciden en la necesidad de seguir estableciendo «iniciativas para responder a la violencia contra las mujeres que se dedican a la política». Una violencia que se ejerce psicológica y físicamente en persona, pero que también se ha traducido en acoso digital —mensajes, fotografías y grabaciones no solicitadas—, secuestro o suplantación de identidad de sus perfiles en redes sociales y todo tipo de difamación *online* (Håkansson, 2021).

Según factores ideológicos

Si nos enfocamos en una comparación de la conversación según bandos ideológicos, podemos advertir que la diferencia más reveladora se encuentra en las temáticas en las que centran el debate las diferentes organizaciones (figura 3).

Por una parte, los progresistas encuentran sus mayores ratios de retroalimentación en las publicaciones donde defienden sus propuestas. El líder del PSOE madrileño, Lobato, apuesta por la fiscalidad, aumentando los impuestos a las grandes fortunas. García, de Más Madrid, propone incorporar monitores en las instalaciones al aire libre para incentivar el ejercicio. Y Jacinto comenta la reforma del Código Penal impulsada por Unidas Podemos en relación con la ocupación de viviendas. Por el contrario, las líderes alineadas con el ala derecha de la política madrileña (Ayuso, Gómez y Monasterio) construyen sus relatos continuamente enfocadas en el enemigo a batir.

Sin embargo, a la hora de examinar las publicaciones con un índice mayor de toxicidad, encontramos un

FIGURA 3. NUBES DE PALABRAS CLAVE EN LA CONVERSACIÓN DE CADA CANDIDATO.

Fuente: Elaboración propia.

patrón común. Desde el mensaje inicial que lanza cada candidato, se utiliza un tono negativo para describir la situación actual y determinar directamente quién es el objetivo a culpar. El bando conservador mantiene su estrategia de señalar un enemigo: Sánchez, en el caso de la presidenta autonómica, mientras que la candidata de Vox alude de manera despectiva a Pablo Iglesias. La dirigente de Ciudadanos apunta hacia Ayuso, la oponente antagonista compartida también con los candidatos del PSOE, de Más Madrid y de Podemos a nivel autonómico. De esta manera, el lenguaje utilizado por las audiencias en este tipo de publicaciones destaca por palabras tóxicas como las mostradas en la figura 4.

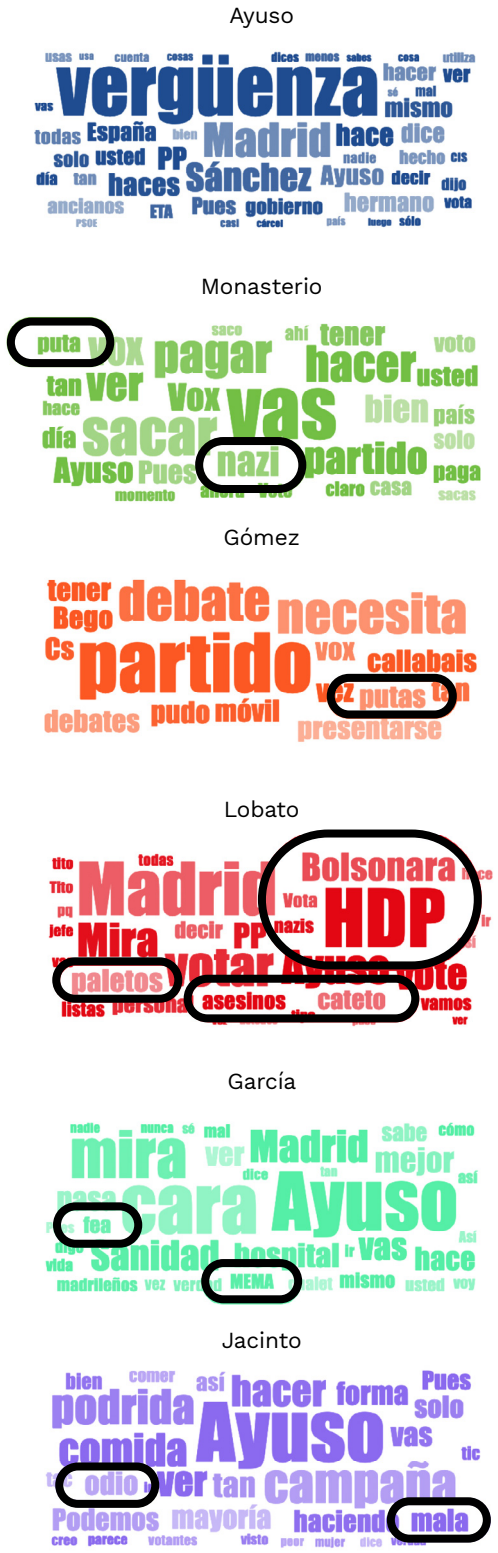
Los candidatos progresistas reciben mayor tasa de toxicidad leve que los líderes de formaciones conservadoras, pero superan incluso la media general, como también ocurre en el análisis de los atributos de blasfemia, toxicidad severa y amenazas. Pese a ser datos significativos estadísticamente, la diferencia porcentual con los políticos de derechas es tan ínfima que no se puede comprobar que exista una relación entre la tasa de toxicidad y el signo ideológico de la conversación. Por ende, independientemente del factor ideológico, cabe resaltar el señalamiento de un enemigo como monotemática de las publicaciones más tóxicas. Este patrón común destaca como uno de los elementos clave para desatar las polémicas, como se ha demostrado anteriormente, a través del tipo de estructura que fomenta la toxicidad.

PI 2.1: CANDIDATOS VS. PARTIDOS POLÍTICOS

El peso de unas siglas de partido siempre se ha considerado un *input* relevante para la reputación del candidato, puesto que la propia masa militante de la organización ejerce, durante la campaña electoral, un respaldo unívoco al líder, cuya autoridad se ve legitimada a los ojos de los votantes. Sin embargo, en la última década, este componente ha ido reduciendo su trascendencia en beneficio del propio liderazgo del candidato.

En términos comunicativos, se proporcionan dos motivos complementarios que explican esta transposición de poderes:

FIGURA 4. NUBES DE PALABRAS CLAVE EN LAS CONVERSACIONES TÓXICAS DE CADA CANDIDATO.



Fuente: Elaboración propia.

Esta investigación confirma la existencia de un arraigo comunicativo más estrecho entre la audiencia y los perfiles de los dirigentes, que entre los usuarios y las cuentas de los partidos

Por un lado, el foco de la cobertura de noticias se ha desplazado desde los partidos a los dirigentes (Pérez-Castaños *et al.*, 2023; Van Aelst *et al.*, 2012). Este movimiento viene precedido por el fenómeno de la «audiovisualización de la política», es decir, el creciente protagonismo de la imagen frente al argumento (Donofrio & Rubio, 2019).

Por otro, la aparición de las redes sociales ha dado respuesta a esa necesidad de identificación con el candidato, propiciando la personalización del discurso en la figura de los líderes (Del Rey-Morató, 2017). Y, en consecuencia, las plataformas sociales rompen la simbiosis que asociaba a los políticos con cualquier agente intermediario comunicativo. Anteriormente, durante la campaña electoral, los candidatos necesitaban llegar al público a través del mayor número de mecanismos mediáticos posibles, aun si estos no coincidían en los intereses ideológicos. En cambio, la conexión directa con sus simpatizantes que les permiten actualmente las redes sociales provoca cierta pérdida de influencia por parte de los medios de comunicación tradicionales.

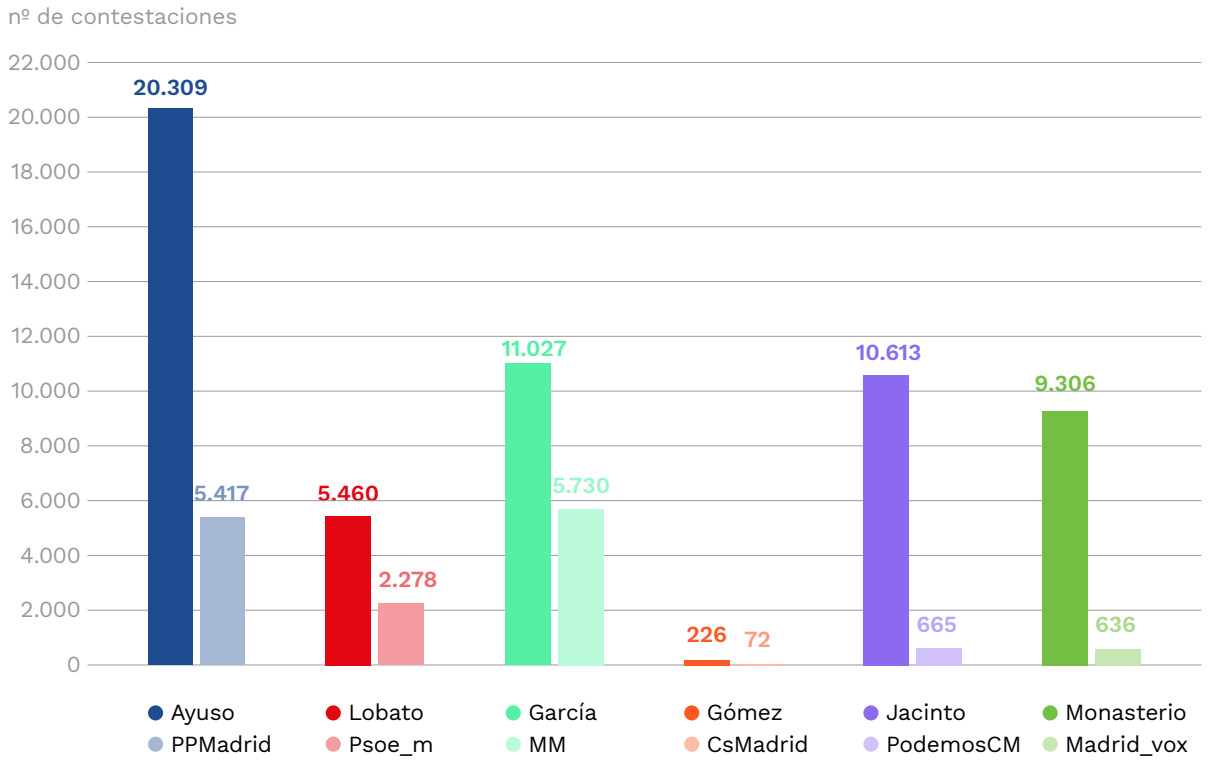
Otra de las ventajas capitales que han aportado las redes sociales para los líderes ha sido la democratización de una herramienta comunicativa que, en muchas ocasiones, obtiene un radar de alcance público igual o superior a los medios convencionales (Papacharissi, 2015). Este foco mediático, gratuito e ilimitado ha proporcionado a los candidatos un nuevo espacio de expresión que, ni en su objetivo de persuasión hacia el electorado, ni en la queja, ni siquiera en la confrontación con adversarios políticos y medios, en ningún caso se topa con un árbitro juzgador de sus verdades y mentiras, como sí suele existir en los filtros informativos que exigen los medios tradicionales (Marín-Pérez *et al.*, 2021).

Además, sus conversaciones no solo obedecen al control de las respuestas, sino también a sus preguntas retóricas. Por lo tanto, los mensajes que publican hacia sus audiencias pueden versar sobre su ideología, su programa, o, incluso, dar a conocer gustos, intereses y asuntos de su vida más privada. Convertir la propia intimidad en un nuevo elemento público es una vertiente más de las estrategias de comunicación política (Castelo, 2020) que han emergido a raíz del lanzamiento de las redes sociales como canales de campaña electoral.

Por ende, este nuevo canal ha permitido a los candidatos poder mantener un vínculo personal más próximo con los ciudadanos que, en cambio, los partidos, por su condición de entes no físicos, no alcanzan. Sin embargo, la posibilidad de una mayor relación con la audiencia, a veces establece conexiones fuertemente positivas, que permiten la creación de comunidades *online* que funcionan como redes de apoyo en torno al candidato; y en otras ocasiones provoca enfrentamientos que sugieren todo tipo de toxicidades (desinformación, discursos de odio o acoso, entre otros). En ese sentido, esta investigación confirma la existencia de un arraigo comunicativo más estrecho entre la audiencia y los perfiles de los dirigentes políticos, que entre los usuarios y las cuentas de los partidos políticos.

Como se observa en la figura 5, existe una distinción nítida en el nivel de *engagement*, medido en contestaciones, que las publicaciones de las organizaciones y las de los candidatos obtuvieron por parte de los usuarios. En todos los casos, los mensajes que lanza el líder político reciben más atención que los emitidos por sus partidos políticos correspondientes. A pesar de que, en volumen, Más Madrid y el Partido Popular madrileño se desmarcan de sus rivales, superando las cinco mil respuestas, cabe hacer hincapié especialmente en la influencia que ejercen los comentarios de Alejandra Jacinto y Rocío Monasterio frente a los perfiles autonómicos de Podemos y Vox respectivamente. Ambas con-

FIGURA 5. COMPARACIÓN DE LA RETROALIMENTACIÓN DE LÍDERES Y PARTIDOS.



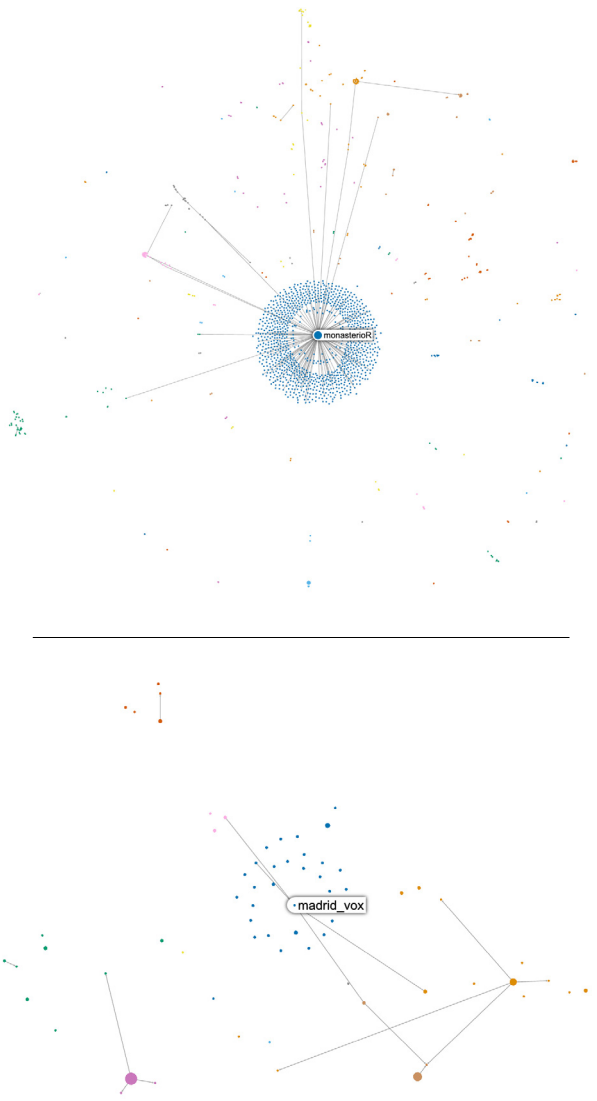
Fuente: Elaboración propia.

siguieron una retroalimentación aproximadamente quince veces mejor que el *feedback* recibido por sus partidos.

Como consecuencia del interés menor por la comunicación partidista, el nivel de las varias categorías de toxicidad también se reduce en las conversaciones que inician las organizaciones políticas, aunque las tasas se mantienen en el mismo orden. Los insultos siguen siendo el atributo tóxico más mencionado por los usuarios, pero su presencia disminuye alrededor de dos puntos porcentuales, rozando el 3,93 % de los comentarios. La toxicidad leve desciende más de un punto, rebajando la porcentual del 4,27 presente en los hilos de los dirigentes hasta un más discreto 2,94 %. También la blasfemia de las conversaciones aminora su presencia en medio punto, con un 2,78 % de las respuestas ofensivas respecto a su tasa en los candidatos (3,25 %); y las amenazas en los debates generados por los partidos políticos se vuelven inexistentes, cayendo el porcentaje al 0 %.

Los usuarios dependen cada vez más unos de otros en sus redes sociales para obtener recomendaciones de noticias e información política

FIGURA 6. COMPARACIÓN +70% DE TOXICIDAD ENTRE CANDIDATA (MONASTERIO) Y PARTIDO (VOX).



Fuente: Elaboración propia.

En términos individuales se produce una conducta similar, como se comprueba a modo ilustrativo en la figura 6. En una comparación entre el partido (Madrid_VOX) y su candidata (Rocío Monasterio), la cuenta de esta última recibe mayor número de mensajes tóxicos que el perfil de la formación política. Una tendencia que se mantiene en los comportamientos de los cinco líderes y organizaciones autonómicas restantes (anexo 2).

PI 2.2: CANDIDATOS VS. LÍDERES DE OPINIÓN

A pesar del liderazgo que ejercen los propios candidatos políticos, la investigación académica ha demostrado que los usuarios dependen cada vez más unos de otros en sus redes sociales para obtener recomendaciones de noticias e información política, y que sus conocimientos, opiniones y comportamientos se ven afectados por la dinámica social del entorno (Duffy *et al.*, 2020; Strömback *et al.*, 2020). Esta influencia en el tipo de contenidos que recibe la audiencia provoca que los usuarios se sometan a un proceso de «exposición selectiva» (Messing & Westwood, 2014), un fenómeno que se ha visto potenciado por el aumento de emisores, canales y contenidos en la actualidad, pero cuyo origen se remonta a la década de los cuarenta.

En la primera ola de este fenómeno, denominada «exposición selectiva *de facto*» (Sears & Freedman, 1967), se entendía que la predilección por un producto comunicativo recaía sobre la influencia del entorno social. En cambio, en la segunda ola de redes egocéntricas (Fischer, 1982), la influencia de Internet y la televisión por satélite abrieron la perspectiva en detrimento de la información local, y la audiencia comenzó a seleccionar el contenido según sus intereses personales. Al cambiar los factores de exposición, desde una perspectiva social a una individual, también aumentó el riesgo de una polarización política e identitaria. El panorama descrito podría encajar en las circunstancias actuales en las que se enmarca esta investigación; no obstante, Barnidge & Peacook (2019) abogan por que las redes sociales han dado lugar a una tercera ola, estructurada por la facilitación para crear nuevos espacios informativos hiperpartidistas y por la diversificación de las conexiones sociales.

Esta primera condición —los nuevos espacios informativos hiperpartidistas— puede fomentar la exposición de los usuarios a ciertos tipos de toxicidad, como la desinformación y las teorías conspiranoicas (Mourão & Robertson, 2019), que incrementan aún más el riesgo de polarización. Ante un escenario con tal cantidad de información, fuentes y riesgos, la audiencia rechaza su exposición directa ante los medios de comunicación, y cubre sus necesidades

a través del contenido acotado que les proporcionan otros usuarios intermediarios, hacia los que la audiencia opta por exponerse porque los considera «líderes de opinión». Este concepto, sin embargo, nace ya en la década de los años cincuenta a través del modelo Two-Step Flows of Communication (Katz, 1957). Los medios de comunicación emiten sus mensajes hacia una audiencia masiva, pero en el camino son interceptados por un tercer actor, los líderes de opinión. Estos intermediarios son vistos por el público como individuos con un acceso mayor y más cercano a los medios de comunicación y, por ello, con mejor comprensión del contenido.

De este modo, los líderes de opinión funcionan como receptores del mensaje de los medios y como emisores encargados de transmitir el contenido a la sociedad (Katz & Lazarsfeld, 1964). Se pone de manifiesto, entonces, una evolución en el modo de transmisión de la información, potenciando las relaciones interpersonales en detrimento del poder directo que solían ejercer los medios de comunicación. En aquella época, todos estos modelos teóricos se basaban en la unidireccionalidad de la comunicación, mientras que las plataformas digitales surgieron a inicios del siglo XXI, concebidas como espacios donde interaccionar aspirando a una comunicación bidireccional.

En la actualidad, sin embargo, todavía no se ha conseguido ese objetivo en su plenitud, y por ello en el entorno de las redes sociales conviven perfiles que no interactúan, funcionando solamente como creadores de contenido, y cuentas que actúan como prosumidores, es decir, consumidores y productores simultáneos de contenido. No obstante, este último apunte sobre la influencia personal, que recogían Katz y Lazarsfeld, aproxima las teorías sobre flujos hacia las dinámicas con las que se entiende hoy en día la comunicación a través de redes sociales. Los usuarios de redes sociales no son individuos pasivos, pues la posibilidad de una comunicación bidireccional ha otorgado a la audiencia el poder de una participación más activa.

Esta nueva cualidad origina el carácter prosumidor de los usuarios en redes sociales, el cual les ha proporcionado un poder mayor frente a los medios de comunicación. Los receptores tienen la capacidad de seleccionar qué mensajes ver, pero también de designar de manera popular a los intermediarios, reduciendo así la autoridad legitimadora de los medios convencionales para nombrar a los líderes de opinión. Al ser seleccionados por los propios usuarios individualmente, estos nuevos transmisores adquieren el potencial de ser influyentes en la configuración de las actitudes y comportamientos políticos de sus pares *online* (Weeks *et al.*, 2017). Además, su alto interés y compromiso con las noticias, y su tendencia a compartir contenidos, los convierten en fuentes confiables (Barberà *et al.*, 2015). Esto significa que el respaldo de los líderes de opinión en las redes sociales también da credibilidad ante terceros (Dubois *et al.*, 2020).

Conforme al conjunto de teorías expuestas, se cuestiona en qué dirección se ubica la toxicidad creada en las interacciones entre usuarios de las redes sociales. Por un lado, se denomina «dirección vertical» a los flujos de comunicación que ocurren en el hilo principal nacido del tuit inicial de la persona candidata. Los perfiles que participan en la conversación, pese a presumir que no van a obtener respuesta, dirigen el contenido de sus mensajes en este hilo principal al emisor del primer mensaje. Por el otro, se considera «dirección horizontal» a la retroalimentación de las interacciones entre individuos pares. Aunque todos estos usuarios hayan llegado por el interés surgido de la publicación inicial del político, el aliciente último que los lleva a participar en estas discusiones de segundo nivel puede venir desencadenado por algún comentario realizado por líderes de opinión.

Para responder a esta tercera pregunta, se han tomado dos momentos como referencia y se han comparado. El primer estadio equivale al estado natural (estado 0) de la conversación, midiendo todas las reacciones hacia el político (eje vertical) y las interacciones entre los usuarios (eje horizontal). El segundo estadio viene dado por el grado de toxicidad que se ha seleccionado como umbral para considerar un mensaje como tóxico. En este caso, se ha seleccionado mostrar todas las interacciones que la herramienta Perspective API ha catalogado con un porcentaje tóxico de mínimo el 70 %.

En síntesis, las conversaciones de los seis actores que forman el objeto de estudio han coincidido en las dinámicas (anexo 3). Como se observa en el mapa de interacciones de Ayuso (figura 7), en un primer momento (estado 0) recoge todas las intercomunicaciones, con independencia del grado de toxicidad de las mismas. Se pueden ver los comentarios dirigidos directamente a la candidata, que se concentran formando un círculo en cuyo centro convergen. Y también se monitorizan las conversaciones que nacen de la interacción entre usuarios, a niveles secundarios, representadas en clústers exteriores al círculo azul. Estas concentraciones de usuarios interconectados en la conversación de segundo nivel y siguientes, no en el hilo principal, son creadas por líderes de opinión (círculos en color verde, naranja y rojo), que alteran el objetivo de las audiencias, cuyo interés acaba residiendo en mayor medida en el debate surgido en esos clústers.

FIGURA 7. GRAFO DE INTERACCIONES EN EL ESTADO 0 DE LA CONVERSACIÓN.

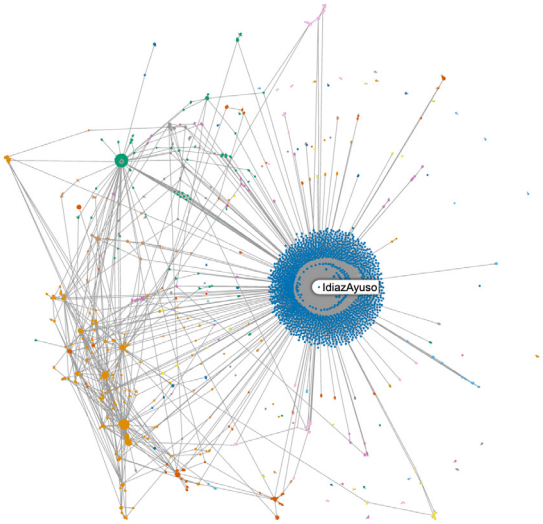
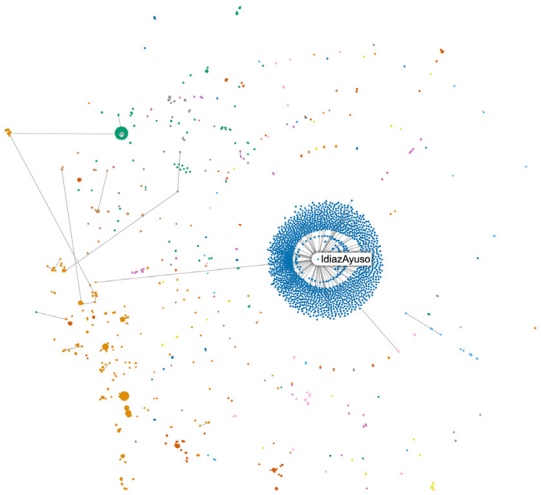


FIGURA 8. GRAFO DE INTERACCIONES CON AL MENOS UN 70 % DE TOXICIDAD.



Fuente: Elaboración propia.

Una vez se conoce el grado natural de la conversación, se procede a buscar concretamente las interacciones más tóxicas. En caso de encontrar una mayor tasa de toxicidad posicionada en el eje comunicativo vertical, se puede intuir que se trata de comportamientos intencionados de la audiencia hacia el emisor. En contraste, de topar un mayor grado de toxicidad en el eje horizontal, se potencia la idea de que ese usuario ha sido incentivado por otros comentarios de parte de líderes de opinión y, además, se ha sentido reforzado socialmente en público por los comentarios de otros pares que componen ese debate de segundo nivel.

En ese sentido, cuando se acota la participación de los usuarios a relaciones con al menos un 70 % de toxicidad en los comentarios, se distingue claramente que el contenido más tóxico se focaliza en mensajes directos a la candidata que publicó el tuit (figura 8). Siguiendo el ejemplo de la conversación de Ayuso, se observa que los comentarios más tóxicos se mantienen en la conversación vertical entre usuario y candidata, mientras que la mayoría de las interacciones entre usuarios desaparecen, puesto que su retroalimentación no contiene niveles de toxicidad tan altos.

No obstante, las dinámicas que se han formado en estos clústers, en sus interacciones más tóxicas, van a seguir señalando a los adalides, ya que son los centralizadores de las conversaciones secundarias. Aun cuando los líderes de opinión no sean quienes emiten o reciben comentarios tóxicos, su persona en sí, o sus comportamientos, han sido el motivo por el cual el usuario que ha intervenido de manera tóxica ha decidido introducirse en la conversación.

04
Conclusiones

PI 1. ¿Existe alguna tendencia que pueda prever una retroalimentación tóxica de la audiencia a los mensajes que emiten los candidatos políticos?

De acuerdo con los hallazgos en el análisis de las diferentes variables de toxicidad, se puede extraer como conclusión la **imposibilidad de mantener una conversación limpia y sana** cuando se trata de temática política. Los niveles de toxicidad no sugieren un clima severo de crispación, pero sí un gran ambiente de polarización, del cual emergen discusiones que contienen menosprecios y burlas, principalmente dirigidos hacia el/la candidato/a. En ese sentido, las descalificaciones más comunes suelen evocar mayoritariamente insultos y ofensas leves.

La toxicidad se genera en la conversación que crean los usuarios; no obstante, también puede venir incitada desde el mensaje inicial que envían los candidatos. Tal y como se observa en la imagen 1, el **modelo de mensaje** de las publicaciones más tóxicas durante esta campaña electoral #28M en la Comunidad de Madrid subraya una estructura clara: **el texto presenta a un enemigo al que se culpa del problema, y se acompaña de un contenido audiovisual (vídeo, fotografía o enlace a una noticia) que funciona como elemento reforzador de la elección que se transmite al votante**.

Otro componente a tener en consideración es la temática de las publicaciones, donde se observa una clara distinción entre bandos ideológicos. El ala progresista intercala propuestas públicas con mensajes acusatorios hacia las políticas de Isabel Díaz-Ayuso, actual presidenta de la Comunidad de Madrid. En cambio, las candidatas de derechas mantienen la misma estrategia discursiva a lo largo de toda la campaña electoral, poniendo el foco en el enemigo. Por ello, se percibe una mayor diferencia en el tipo de lenguaje utilizado en las audiencias del ala progresista dependiendo de las publicaciones que envían. Asimismo, las ratios de toxicidades entre las formaciones de izquierdas son escasamente superiores en la mayoría de categorías.

Desde una perspectiva de género, no se ha demostrado ninguna diferencia sustancial por parte de la audiencia que indique un trato peyorativo hacia las líderes femeninas respecto al género opuesto, o viceversa. Existe, sin embargo, una corriente científica que evidencia «una preferencia [de los votantes] por las características “masculinas” en los niveles más altos de cargos» (Huddy & Terkildsen, 1993: 503). En respuesta a los estereotipos de género de los votantes, algunas dirigentes femeninas optan por el empleo de **comportamientos masculinos** de cara a evitar las implicaciones potencialmente negativas asociadas al género femenino (Okimoto & Brescoll, 2010). No obstante, a pesar de que existe literatura que confirma ciertos casos, en esta investigación no se pueden apreciar cambios significativos en ese sentido.

Junto con el *statu quo* ya adquirido, ciertas características en sus tratamientos y discursos pueden ser percibidos por la audiencia como los prototípicos de un líder masculino capacitado para ejercer la política y, por tanto, ser visto como un rival al que temer (Okimoto & Brescoll, 2010). Complementariamente, dichas candidatas usan su condición de mujer como punto favorable en su reputación, es decir, *purplewashing*, una estrategia de lavado de imagen en materia feminista (Mascia, 2022). Por ende, de cara a aseverar la existencia o no de una brecha en materia de género, se recomienda una profundización en el estudio de sus conductas que, posteriormente, se pueda poner en relación con los niveles de toxicidad que reciben en sus conversaciones.

PI 2. ¿Quiénes generan mayores niveles de toxicidad al ejercer su liderazgo político?, ¿los candidatos o los partidos?, ¿los candidatos o los líderes de opinión?

En cuanto al liderazgo, una de las características esenciales en el campo de la persuasión y en el de la política, se ha demostrado un cambio en la última década. **La importancia de las siglas de los partidos ha disminuido en comparación con la influencia personal de los candidatos**. El dirigente ha pasado de ser la cara de representación de una organización política a presentarse bajo la figura de un líder que, además, cuenta con el respaldo de una formación. Este cambio se ha visto influenciado por la creciente atención mediática a los líderes en lugar de a los partidos, un fenómeno amplificado por la audiovisualización de la política y su presencia en las redes sociales, desempeñando un papel central en la formación de la opinión pública.

En efecto, las redes sociales han jugado un papel crucial al permitir la personalización del discurso, proporcionando a los candidatos un espacio de expresión directo y sin restricciones, lo que ha llevado a una conexión más estrecha con los votantes, aunque también puede dar lugar a comunidades *online* de apoyo y de animadversión. En ese sentido, se destaca que los líderes políticos reciben considerablemente más atención y *engagement* que los partidos, lo que subraya la importancia de la conexión directa con la audiencia. Además, se observa una disminución general de toxicidad en las conversaciones iniciadas por los partidos en comparación con las de los líderes, aunque persiste un nivel mínimo de ofensa.

Sin embargo, a pesar del creciente liderazgo de los candidatos políticos en redes sociales, la investigación académica sugiere que los usuarios dependen cada vez más de recomendaciones e información compartidas por otros usuarios en sus redes para formar opiniones políticas. Este fenómeno, conocido como «exposición selectiva», ha evolucionado hacia una tercera ola, caracterizada por la formación de espacios informativos hiperpartidistas y la diversificación de conexiones sociales. Este contexto permite a los usuarios seleccionar mensajes y designar líderes de opinión, a quienes se exponen, otorgándoles mayor poder de influencia sobre los contenidos en detrimento de la autoridad que solían ejercer los medios convencionales.

En el análisis de la toxicidad en las interacciones en redes sociales, se ha podido verificar que la participación activa de usuarios genera clústers, donde se señala **la influencia de los líderes de opinión en debates secundarios, incluso si no son directamente responsables de la toxicidad** observada. No obstante, la toxicidad se concentra en mayor cantidad en las dinámicas relativas al eje de comunicación vertical, lo cual permite intuir que se trata de comportamientos intencionados de la audiencia hacia el emisor.

Estos usuarios han entrado directamente en la conversación, sin el aliciente de ningún intermediario ni el respaldo de un público social afín, y han ejercido una conducta con un nivel considerablemente tóxico por su propia voluntad. En consecuencia, este comportamiento, o un cúmulo de comentarios similares, pueden provocar que otros usuarios decidan no participar en la discusión o incluso abandonar la conversación.

Bibliografía

Ballington, J. (2018). Turning the Tide on Violence against Women in Politics: How Are We Measuring Up? *Politics & Gender, 14(4),* 695-701. https://doi.org/10.1017/S1743923X18000636

Barberà, P., Jost, J. T., Nagler, J., Tucker, J. A., Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science, 26(10)*, 1531-1542. https://doi.org/10.1177/0956797615594620

Bardall, G. (2013). Gender-specific election violence: The role of information and communication technologies. *Stability: International Journal of Security and Development, 2(3*), 1-11. http://dx.doi.org/10.5334/sta.cs

Bardall, G., Bjarnegård, E., & Piscopo, J. M. (2020). How is political violence gendered? Disentangling motives, forms, and impacts. *Political Studies, 68(4*), 916-935. https://doi.org/10.1177/0032321719881812

Barnidge, M., & Peacock, C. (2019). A third wave of selective exposure research? The challenges posed by hyperpartisan news on social media. *Media and communication*, *7(3)*, 4-7. https://doi.org/10.17645/mac.v7i3.2257

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

Berrocal-Gonzalo, S., Waisbord, S., & Gómez-García, S. (2023). Polarización política y medios de comunicación, su impacto en la democracia y en la sociedad. *Profesional de la información / Information Professional*, *32(6)*. https://doi.org/10.3145/epi.2023.nov.22

Bjarnegård, E. (2018). Making gender visible in election violence: Strategies for data collection. *Politics & Gender, 14(4)*, 690-695. https://doi.org/10.1017/S1743923X18000624

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

Carral, U., & Elías, C. (2024). Aplicación de herramientas de IA como metodología para el análisis de la toxicidad en la conversación en redes sociales: Estudio de caso de la política española en Twitter. *Revista Latina de Comunicación Social*. https://doi.org/10.4185/rlcs-2024-2205

Casero-Ripollés, A., Feenstra, R. A., & Tormey, S. (2016). Old and new media logics in an electoral campaign: The case of Podemos and the two-way Street mediatization of politics. *The International Journal of Press/Politics, 21(3)*, 378-397. https://doi.org/10.1177/1940161216645340

Castelo, S. (2020). *Cuando la intimidad se hace pública. Un análisis de la publicitación de lo privado como estrategia de comunicación política*. En Pont-Sorribes, C., & Gutiérrez-Rubí, A. (2020). *Instagram en la construcción del liderazgo político*. Editorial Gedisa.

Chadwick, A., & Vaccari, C. (2019). *News sharing on UK social media: Misinformation, disinformation, and correction.* Loughborough University. Report.

Del Rey-Morató, J. (2017). *Comunicación política, Internet y campañas electorales. De la democracia a la ciberdemocracia*. Madrid: Ed. Tecnos.

Digital News Report 2023. (2023). Reuters Institute for the Study of Journalism. Disponible en: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023

Donofrio, A., & Rubio-Moraga, A. L. (2019). De Berlusconi a Trump: la comunicación convertida en espectáculo. *Estudios sobre el Mensaje Periodístico, 25(1)*, 113-127. https://doi.org/10.5209/ESMP.63719

Dubois, E., Minaeian, S., Paquet-Labelle, A., & Beaudry, S. (2020). Who to trust on social media: How opinion leaders and seekers avoid disinformation and echo chambers. *Social media + Society, 6(2*). https://doi.org/10.1177/2056305120913993

Duffy, A., Tandoc, E., & Ling, R. (2020). Too good to be true, too good not to share: the social utility of fake news. *Information, Communication & Society, 23(13)*, 1965-1979. https://doi.org/10.1080/1369118X.2019.1623904

Fischer, C. S. (1982). *To dwell among friends: Personal networks in town and city*. Chicago: University of Chicago Press.

Graves, L., & Anderson, C. W. (2020). Discipline and promote: Building infrastructure and managing algorithms in a "structured journalism" project by professional *fact-checking* groups. *New media & society*, *22*(2), 342-360. https://doi.org/10.1177/1461444819856916

Gruzd, A., Mai, P., & Vahedi, Z. (2022). Studying Anti-Social Behaviour on Reddit with Communalytic. En Quan-Haase, A., & Sloan, L. (eds.). *The SAGE Handbook of Social Media Research Methods,* 503-520. Disponible en: https://www.semanticscholar.org/reader/45a79dee30fe649583889573d16be1b316cf72d1

Håkansson, S. (2021). Do women pay a higher price for power? Gender bias in political violence in Sweden. *The Journal of Politics, 83(2)*, 515-531. https://doi.org/10.1086/709838

Hayes, D., & Lawless, J. (2015). A Non-Gendered Lens? Media, Voters, and Female Candidates in Contemporary Congressional Elections. *Perspectives on Politics, 13(1)*, 95-118. https://doi.org/10.1017/S1537592714003156

Hua, Y., Ristenpart, T., & Naaman, M. (2020). Towards Measuring Adversarial Twitter Interactions against Candidates in the US Midterm Elections. *Proceedings of the International AAAI Conference on Web and Social Media, 14,* 272-282. https://doi.org/10.1609/icwsm.v14i1.7298

Huddy, L., & Terkildsen, N. (1993). The consequences of gender stereotypes for women candidates at different levels and types of office. *Political Research Quarterly, 46(3),* 503-525. https://doi.org/10.1177/106591299304600304

Inter-Parliamentary Union (IPU). (2016). *Sexism, harassment and violence against women parliamentarians: issues brief*. Disponible en: https://www.ipu.org/resources/publications/issue-briefs/2016-10/sexism-harassment-and-violence-against-women-parliamentarians

Katz, E. (1957). The Two-Step Flow of Communication. *Political Opinion Quarterly, 21(1)*, 61-78. https://doi.org/10.1086/266687

Katz, E., & Lazarsfeld, P. F. (1964) (*e-book*, 2017). *Personal Influence, The part played by people in the flow of mass communications*. New York: Routledge. https://doi.org/10.4324/9781315126234

Lampe, C., & Johnston, E. (2005). Follow the (slash) dot: effects of feedback on new members in an online community. *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, 11-20. https://doi.org/10.1145/1099203.1099206

Lozano-Blasco, R., Mira-Aladrén, M., & Gil-Lamata, M. (2023). Redes sociales y su influencia en los jóvenes y niños: Análisis en Instagram, Twitter y YouTube. *Comunicar, 31(74)*, 125-137. https://doi.org/10.3916/C74-2023-10

Marín-Pérez, B. M., Pérez-Sánchez, J., & Pop, A. I. (2021). La focalización mediática en los principales informativos de televisión en España en las campañas electorales generales de 2019: partidos políticos *versus* líderes. *Dígitos. Revista de comunicación digital*, *1(7)*, 15. https://doi.org/10.7203/rd.v1i7.209

Mascia, C. (2022, 30 abril). Marine Le Pen: el triunfo del *'purplewashing'*. *El País*. Disponible en: https://elpais.com/opinion/2022-04-30/marine-le-pen-el-triunfo-del-purplewashing.html

Messing, S., & Westwood, S. J. (2014). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research, 41(8)*, 1042-1063. https://doi.org/10.1177/0093650212466406

Miller, S., Menard, P., Bourrie, D., & Sittig, S. (2022). Integrating truth bias and elaboration likelihood to understand how political polarization impacts disinformation engagement on social media. *Information Systems Journal*, 1-38. https://doi.org/10.1111/isj.12418

Mourão, R. R., & Robertson, C. T. (2019). Fake news as discursive integration: An analysis of sites that publish false, misleading, hyperpartisan and sensational information. *Journalism Studies, 20(14),* 2077-2095. https://doi.org/10.1080/1461670X.2019.1566871

Munger, K. (2019). All the news that's fit to click: The economics of clickbait media. *Political Communication, 37(3*), 376-397. https://doi.org/10.1080/10584609.2019.1687626

Murray, R., Krook, M. L., & Opello, K. A. (2012). Why are gender quotas adopted? Party pragmatism and parity in France. *Political Research Quarterly, 65(3*), 529-543. https://doi.org/10.1177/1065912911414590

Naciones Unidas (2021, 16 abril). *Informe sobre los progresos realizados en la aplicación del Plan Estratégico para la Igualdad de Género y el Empoderamiento de las Mujeres 2018-2021*. Junta Ejecutiva para la Igualdad de Género y el Empoderamiento de las Mujeres. Disponible en: https://www.upm.es/gsfs/SFS08922

Okimoto, T., & Brescoll, V. (2010). The Price of Power: Power Seeking and Backlash Against Female Politicians. *Personality and Social Psychology Bulletin, 36,* 923-936. https://doi.org/10.1177/0146167210371949

Organization of American States (OAS). (2015). *Declaration on Political Harassment and Violence against Women*. Sixth Conference of the States Parties to the Belém do Pará Convention. Disponible en: https://www.oas.org/en/cim/docs/ViolenciaPolitica-FactSheet-EN.pdf

Papacharissi, Z. (2015). *Affective Publics. Sentiment, Technology and Politics*. Oxford: Oxford University Press.

Pascual-Ferrá, P., Alperstein, N., Barnett, D. J., & Rimal, R. N. (2021). Toxicity and verbal aggression on social media: Polarized discourse on wearing face masks during the COVID-19 pandemic. *Big Data & Society*, *8(1)*. https://doi.org/10.1177/20539517211023533

Pérez-Castaños, S., Antón-Merino, J., & García-Santamaría, S. (2023). Emociones, liderazgo y redes sociales. Propuesta para su medición en materiales de campaña electoral. *Revista Española de Investigaciones Sociológicas*, 184, 125-136. https://doi.org/10.5477/cis/reis.184.125

Pérez-López, C., & Santín-González, D. (2007). Minería de datos: Técnicas y herramientas. *Ediciones Paraninfo, SA*.

Renee, D. (2018). Computational propaganda: If you make it trend, you make it true. *The Yale Review, 106(4)*, 12-29. https://doi.org/10.1353/tyr.2018.0030

Rogers, R. (2020). *Deplatforming*: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication*, *35(3)*, 213-229. https://doi.org/10.1177/0267323120922066

Sánchez-Duarte, J. M., & Magallón-Rosa, R. (2023). Desinformación. *Eunomía. Revista en Cultura de la Legalidad, 24*, 236-249. https://doi.org/10.20318/eunomia.2023.7663

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

Sears, D. O., & Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, *31(2)*, 194-213. https://doi.org/10.1086/267513

Strömback, J., Tsfati, Y., Boomgaarden, H., Damstra, A., Lindgren, E., Vliegenthart, R., & Lindholm, T. (2020). News media trust and its impact on media use: Toward a framework for future research. *Annals of the International Communication Association, 44(2)*, 139-156. https://doi.org/10.1080/23808985.2020.1755338

Teves, L. S., & Cueto, J. J. (2020). El Análisis de Redes Sociales (ARS) aplicado a problemáticas actuales. *AWARI: Revista de la Asociación Latinoamericana de Análisis de Redes Sociales*, 1. Disponible en: http://sedici.unlp.edu.ar/handle/10915/110952

Van Aelst, P., Sheafer, T., & Stanyer, J. (2012). The personalization of mediated political communication: A review of concepts, operationalizations and key findings. *Journalism, 13(2)*, 203-220. https://doi.org/10.1177/1464884911427802

Weeks, B. E., Ardèvol-Abreu, A., & Gil de Zúñiga, H. (2017). Online influence? Social media use, opinion leadership, and political persuasion. *International Journal of Public Opinion Research, 29(2*), 214-239. https://doi.org/10.1093/ijpor/edv050

Yang, K. C., Varol, O., Davis, C. A., Ferrara, E., Flammini, A., & Menczer, F. (2019). Arming the public with artificial intelligence to counter social bots. *Human Behavior and Emerging Technologies, 1(1)*, 48-61. https://doi.org/10.1002/hbe2.115

Anexos

1. LOS TUITS CON MÁS TOXICIDAD DE CADA CANDIDATO

a) @IdiazAyuso: 25/05/2023.

b) @juanlobato_es: 13/05/2023.

c) @Monica_Garcia_G: 16/05/2023.

d) @monasterioR: 13/05/2023.

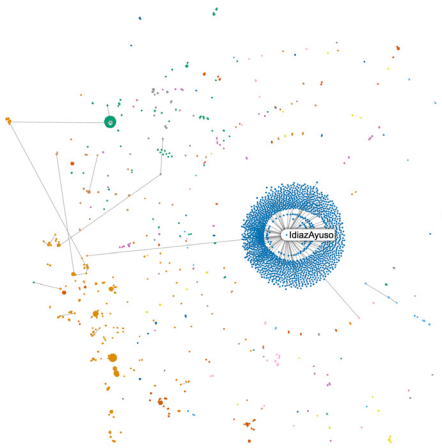
e) @AleJacintoUrang: 21/05/2023.

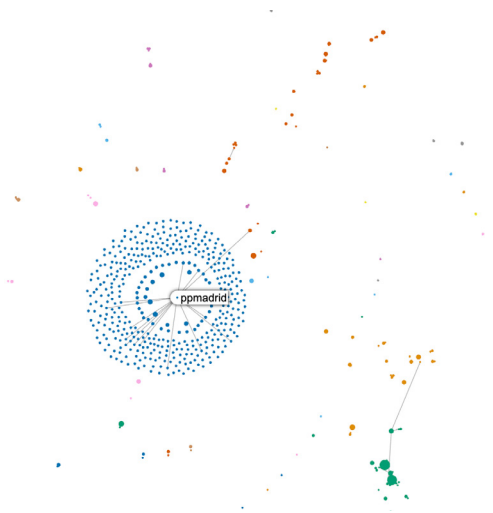
f) @ArucaGomez: 19/05/2023.

2. LA COMPARATIVA DE TOXICIDAD DE CANDIDATOS VS. PARTIDOS

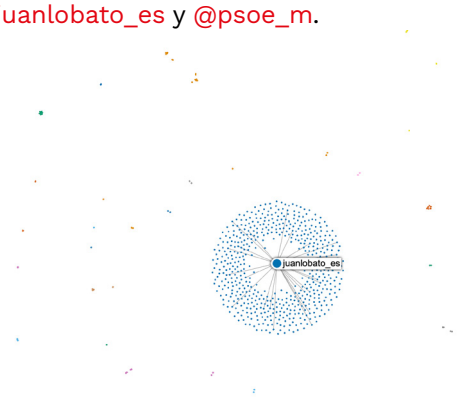
(Grafo de interacciones con al menos un 70 % de toxicidad en candidatos y partidos).

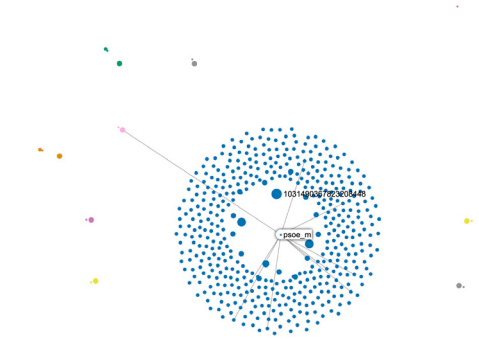
a) @IdiazAyuso y @ppmadrid.



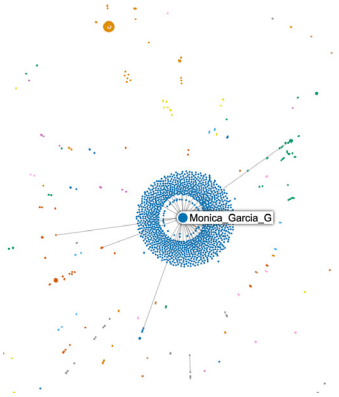


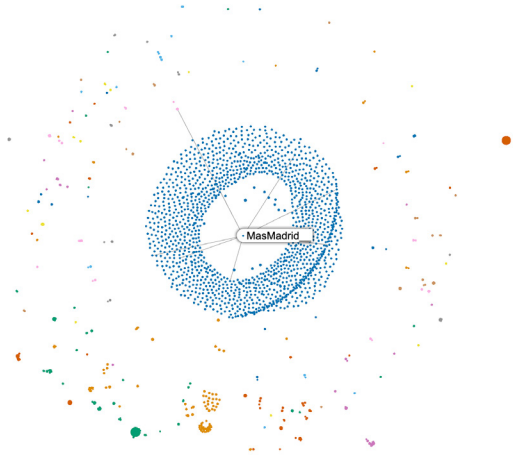
b) @juanlobato_es y @psoe_m.



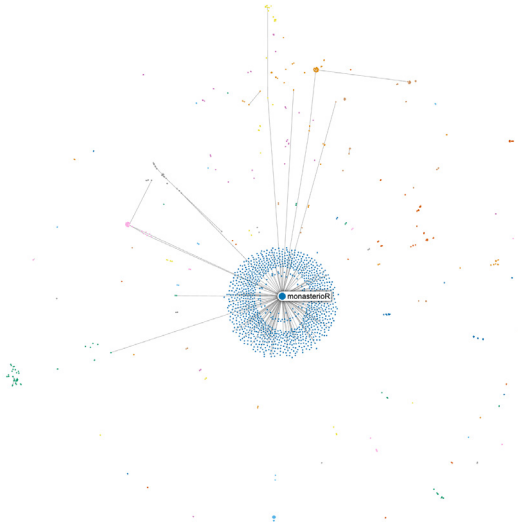


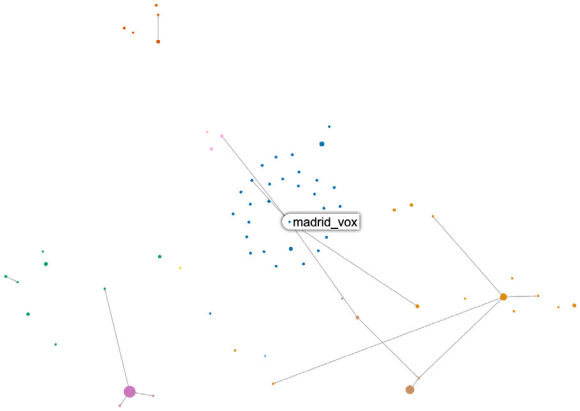
c) @Monica_Garcia_G y @MasMadrid__.



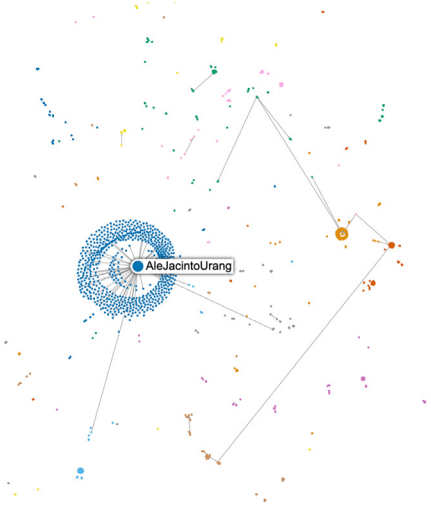


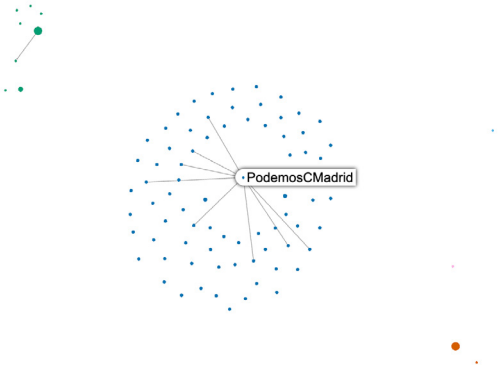
d) @monasterioR y @madrid_vox.



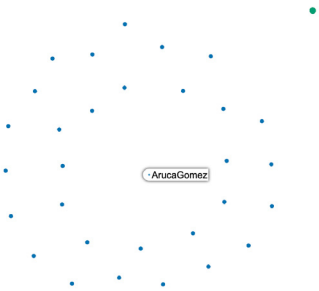


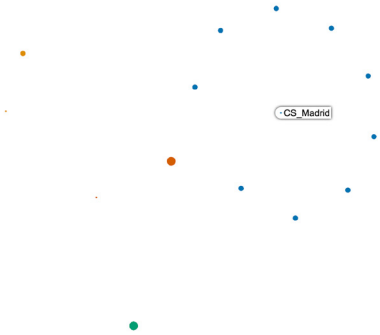
e) @AleJacintoUrang y @PodemosCMadrid.





f) @ArucaGomez y @CS_Madrid.

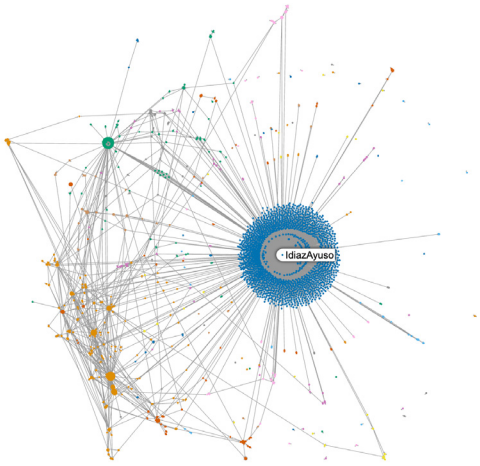


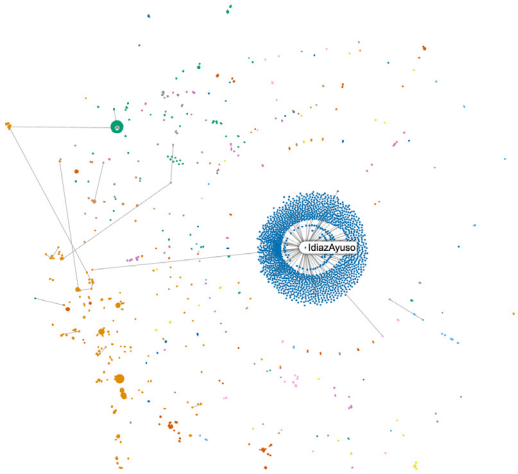


3. LAS INTERACCIONES DE CADA UNO DE LOS CANDIDATOS

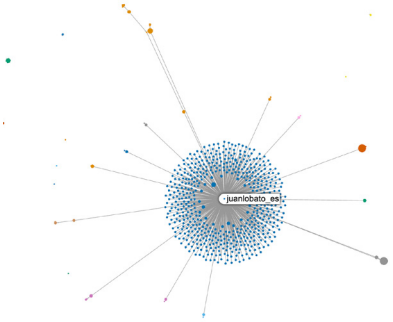
(Grafo de interacciones en estado inicial y grafo con al menos un 70 % de toxicidad).

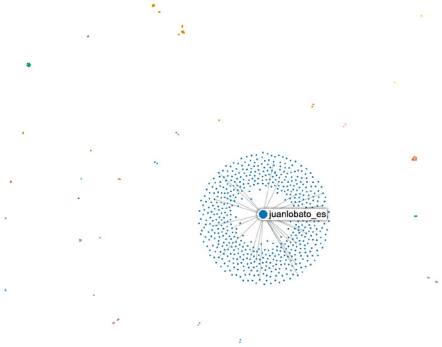
a) @IdiazAyuso.



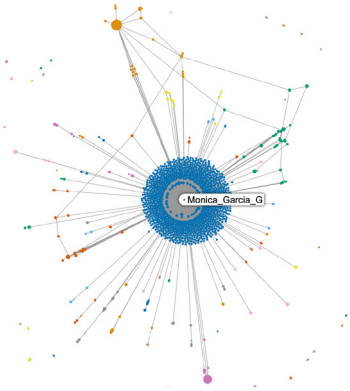


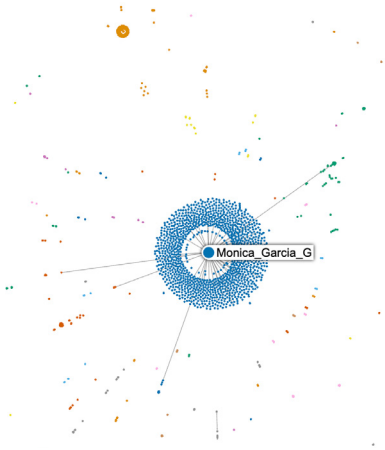
b) @juanlobato_es.



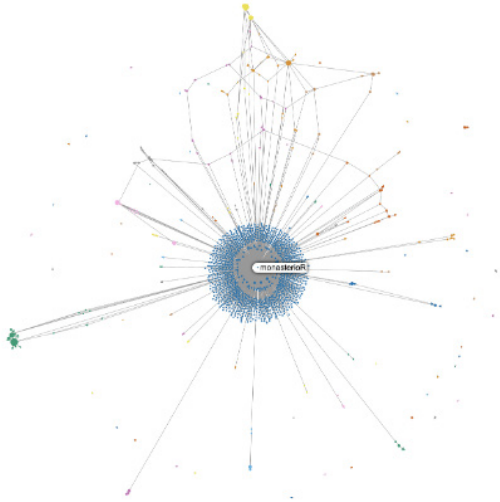


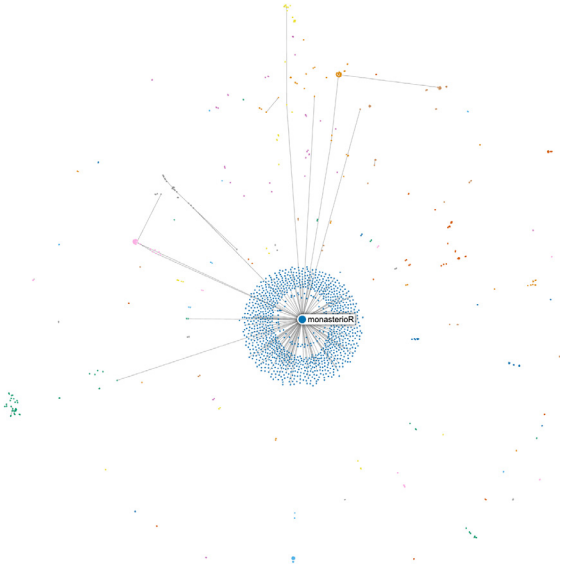
c) @Monica_Garcia_G.



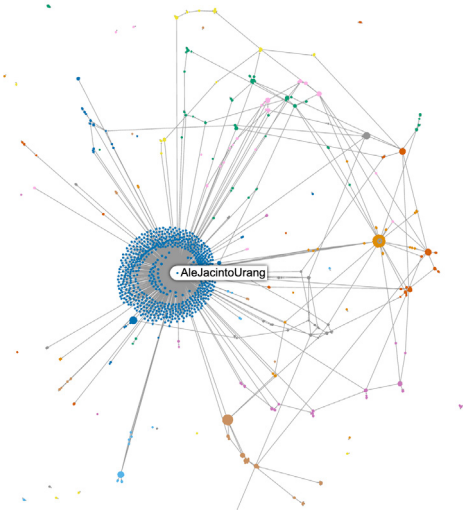


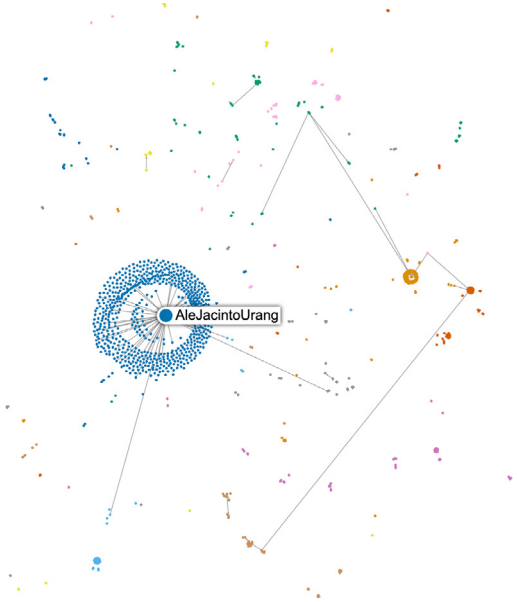
d) @monasterioR.



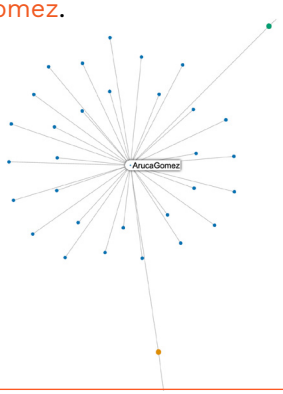


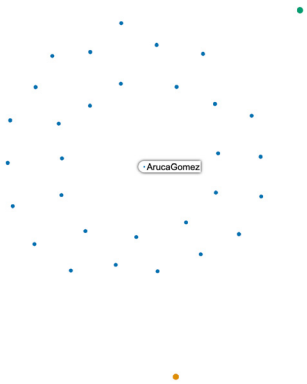
e) @AleJacintoUrang.





f) @ArucaGomez.





Ganadoras de las ediciones anteriores del Premio Cátedra Ideograma – UPF

1ª edición del Premio: **Sandra Pallarés Navarro** con el trabajo **«El discurso populista en Twitter. Análisis del caso de Santiago Abascal en la campaña electoral del 10N»** de la UNED.

Resumen: La irrupción de partidos populistas en Europa y Estados Unidos se ha ido consolidando durante la última década. La victoria de Trump en EEUU y el Brexit en el Reino Unido son claros exponentes de una tendencia que, aprovechando una crisis de representación, pretende romper con el establishment. Si bien la aparición de estos partidos populistas de extrema derecha se ha ido afianzando en Europa Occidental, parecía que España se resistía a su irrupción. Sin embargo, la entrada de VOX en las instituciones públicas supuso la llegada de este fenómeno a España. En las elecciones andaluzas de 2018, el partido liderado por Abascal logró 12 escaños. Cuatro meses después, la formación logró 24 diputados en la convocatoria del 28A. Y tras la repetición electoral del 10N, ascendió hasta los 52 escaños. Cabe subrayar el importante papel que han jugado las redes sociales para la expansión de este partido, constituidas como medios alternativos para difundir su discurso dada la tensa relación de VOX con los medios de comunicación tradicionales.

En este contexto, se plantea el estudio del uso de Twitter por Santiago Abascal durante la campaña electoral del 10N (10 de noviembre de 2019) con el objetivo de analizar su estrategia discursiva y determinar si aparecen rasgos característicos de la retórica populista en su discurso. Se analiza la cuenta del candidato dado el auge del personalismo en política y se estudia la campaña electoral del 10N puesto que fue un periodo decisivo que supuso la consolidación de VOX como tercera fuerza política.

2ª edición del Premio: **Carmen Luhua Martín Mira** con el trabajo **«La comunicación institucional de la Casa Real británica en la gestión de la crisis del Megxit»** de la Universidad de Sevilla.

Resumen: «La comunicación institucional de la Casa Real británica en la gestión de la crisis del *Megxit*» es un trabajo de investigación en el cual se analizan las estrategias, técnicas y herramientas de comunicación para la gestión de crisis por parte de la Casa Real británica aplicado al caso del *Megxit*.

Para ello, se evalúan las principales herramientas empleadas por la Familia Real, examinando cómo han gestionado esta crisis comunicativamente y los medios de los que se han servido. A través de un análisis de los mensajes emitidos en Twitter por la cuenta oficial de la Casa Real (@RoyalFamily) se ha podido dilucidar el carácter de las declaraciones de Buckingham Palace y de Su Majestad la Reina, la imagen del conflicto y la imagen mediática construida sobre los duques de Sussex. Previamente se había realizado un estudio del tratamiento de la comunicación institucional por parte de la Corona en Twitter, la plataforma escogida para el estudio.

Como conclusión de la investigación se observa que la Casa Real británica se encuentra arraigada en su manera de comunicar y gestionar las crisis, lo que no le impide aprender de sus errores y aplicar mejoras con las cuales consigue mantener a la opinión pública a su favor, demostrando la supervivencia inherente a la Casa de Windsor

---

3ª edición del Premio: **Uxía Carral** con el trabajo **«¿Cuán tóxica es la conversación sobre política española en Twitter? Estudio de este fenómeno a través de la IA»** de la CEU San Pablo.

Resumen: Este trabajo pretende conocer cómo diversas herramientas de IA pueden ayudar a mejorar la calidad de la conversación y combatir la polarización en redes sociales. Para ello, se han analizado 43.165 tuits publicados del 18 al 24 de octubre de 2021 y correspondientes a siete políticos españoles y a la cascada de respuestas de los usuarios. Los resultados apuntan a los insultos como la categoría tóxica predominante en los comentarios, independientemente de la ideología de los usuarios. Además, se observa una tendencia a la propagación de comentarios negativos directamente hacia el político, pero una conversación menos tóxica entre pares.

Cátedra Ideograma-UPF
de Comunicación Política y Democracia