

Discursos de odio en la red social X tras el segundo debate presidencial en Ecuador 2025

Hate Speech on the Social Network X After the Second Presidential Debate in Ecuador 2025

Discursos de ódio na rede social X após o segundo debate presidencial no Equador 2025

María Fernanda García Mero¹
ORCID: 0009-0000-4756-8718

Juan Pablo Trámpuz¹
ORCID: 0000-0001-9999-177X

Oscar Molina-Bailon¹
ORCID: 0000-0002-2137-0911

¹Universidad Laica Eloy Alfaro de Manabí, Ecuador.
Correspondencia: oscar.molina@uleam.edu.ec

Recepción: 04/12/2025
Revisión: 30/01/2026
Aceptación: 18/02/2026

RESUMEN. Este trabajo analiza la prevalencia, tipología y dinámica temporal de los discursos de odio generados por usuarios en la red social X, en el contexto de la segunda vuelta presidencial en Ecuador. Se utilizó Python para el procesamiento del corpus de publicaciones, a partir del cual se realizó una subclasificación en categorías de odio (amenazas, difamación, insultos y descortesía); se emplearon reglas priorizadas con umbrales estadísticos, y se analizó la evolución diaria de las interacciones mediante intervalos de confianza de Wilson y técnicas de bootstrap. Los hallazgos mostraron que la hostilidad en X contra Daniel Noboa y Luisa González fue respuesta a circunstancias específicas, hubo picos de agresividad a inicios de abril, coincidiendo con el segundo debate, y un descenso hacia el 13 de abril, día electoral. El discurso de odio en la esfera digital ecuatoriana se presenta de forma estratégica para deslegitimar moralmente la alteridad política y se consolida como un recurso sistémico de la cultura política.

Palabras clave: comunicación política; discurso de odio; elecciones; redes sociales; violencia.

ABSTRACT. This study examines the prevalence, typology, and temporal dynamics of user-generated hate speech on the social media platform X in the context of the second round of the presidential election in Ecuador. Python was used to process the corpus of posts, which were subsequently subclassified into categories of hate (threats, defamation, insults, and incivility). A rule-based approach with prioritized criteria and statistical thresholds was applied, and the daily evolution of interactions was analyzed using Wilson confidence intervals and bootstrap techniques. The findings indicate that hostility on X toward Daniel Noboa and Luisa González was responsive to specific events. Peaks in aggression were observed in early April, coinciding with the second debate, followed by a decline toward April 13, Election Day. Hate speech in the Ecuadorian digital sphere appears to be deployed strategically to morally delegitimize political alterity and has become consolidated as a systemic feature of the country's political culture.

Keywords: political communication; hate speech; elections; social networks; violence.

RESUMO. Este trabalho analisa a prevalência, tipologia e dinâmica temporal dos discursos de ódio gerados por usuários na rede social X, no contexto do segundo turno presidencial no Equador. Utilizou-se o Python para o processamento do corpus de publicações, a partir do qual foi realizada uma subclassificação em categorias de ódio (ameaças, difamação, insultos e descortesia); empregaram-se regras priorizadas com limiares estatísticos, e analisou-se a evolução diária das interações por meio de intervalos de confiança de Wilson e técnicas de bootstrap. Os achados mostraram que a hostilidade no X contra Daniel Noboa e Luisa González foi uma resposta a circunstâncias específicas, houve picos de agressividade no início de abril, coincidindo com o segundo debate, e um declínio por volta de 13 de abril, dia da eleição. O discurso de ódio na esfera digital equatoriana apresenta-se de forma estratégica para deslegitimar moralmente a alteridade política e consolida-se como um recurso sistêmico da cultura política.

Palavras-chave: comunicação política; discurso de ódio; eleições; redes sociais; violência.

Introducción

El presente estudio se orienta en analizar la forma en que las redes sociales, en especial X —antes conocida como Twitter—, han influido en la confrontación digital de sus usuarios, en especial durante eventos políticos particulares como son los debates presidenciales, contexto en el que los ciudadanos suelen adoptar posturas radicales que se trasladan a la virtualidad. Para ello se parte desde la premisa de que las redes sociales han transformado la libertad de expresión a tal nivel que no solo facilitan la difusión instantánea y global de opiniones, sino también de discursos de odio. Por ende, la capacidad que estas herramientas tienen para viralizar contenido de forma anónima permite la propagación de mensajes extremistas y la deslegitimación espacios democráticos. Esta situación lleva a repensar los límites entre expresión legítima y discursos peligrosos que generan polaridad en la ciudadanía.

La red social X es una plataforma que para el 2025 registra alrededor de 611 millones de usuarios activos al mes, de estos al menos 245 millones acceden a la plataforma diariamente (S. Singh, 2025). En los últimos años se ha reportado un incremento en el uso de esta red social con fines informativos de acontecer y política, principalmente entre jóvenes y usuarios conservadores, pues, de acuerdo con Reuters Institute (2025), alrededor del 23 % de adultos en EE.UU. la utilizan como fuente informativa. Se indica también que, desde que fue adquirida por Elon Musk, la plataforma adoptó una política de “libertad de expresión, no de alcance”, lo que limita la difusión de determinados contenidos sin eliminarlos. Además, estudios, como el de Ye et al. (2025), muestran un sesgo algorítmico que da mayor amplificación a contenidos alineados con la derecha política, lo que influye en la agenda pública.

Desde un enfoque a la situación mundial, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco, 2024) ha informado que los discursos de odio en redes sociales han mostrado un aumento significativo, prueba de ello es que en España, por ejemplo, entre 2019 y 2023 los crímenes de odio tuvieron un incremento del 24.71 %, asociados a discursos hostiles en redes y sociedad, lo que a su vez ha sido potenciado por la polarización política y el respaldo a agendas extremistas (Rodríguez-Peral et al., 2025). Durante las elecciones de 2024 de EE. UU., se observaron alrededor de 50 millones

de cuentas en comunidades de odio interactuando entre sí, aumentando el contenido hostil y hostigamiento digital (Verma et al., 2024). En India, en cambio, los efectos políticos y electorales se han reflejado en el aumento del 74 % de discursos de odio contra minorías (K. Singh, 2025).

En Iberoamérica se ha hecho uso de la política para polarizar, deslegitimar adversarios y movilizar bases, sobre todo en periodos electorales. Se ha observado que las corrientes políticas suelen utilizar X como instrumento para la difusión bien planificada de discursos de odio mediante la automatización y cuentas falsas, lo que ha llevado incluso a socavar y exasperar el ambiente democrático de las naciones (Díez-Gutiérrez et al., 2022). Esta situación se agrava por la existencia de entornos de desinformación institucionalizados, desde los que se distribuye contenido divisivo y falsificado con la finalidad de influir en la opinión pública y debilitar la confianza en instituciones. Para ejemplificar, en Argentina en los últimos años se han amplificado ataques y hostigamiento digitales contra críticos y opositores, lo que da paso a nuevas narrativas políticas con efecto en la deliberación pública (Freedom House, 2024a).

En el Ecuador, los discursos de odio ganan cada vez más espacio y notoriedad, y tienen un impacto significativo en el debate político y electoral. Un informe del Observatorio de Comunicación de la Pontificia Universidad Católica del Ecuador ha dado a conocer que entre el 10 y el 21 de febrero de 2025 se detectaron más de 83.000 publicaciones con contenido de odio, acumulando más de 3 millones de interacciones, y con un 80 % de sentimiento negativo, en las que se evidenciaron narrativas polarizadas que refuerzan acusaciones como “narco” o “títere” contra figuras públicas (Cruz, 2025). Dicho informe revela además que los algoritmos favorecieron la viralización de contenido extremista, se reduce así el espacio para el debate constructivo. En 2023, en cambio, se documentaron 265 agresiones bajo discurso de odio dirigidas en gran parte a periodistas e indígenas, motivadas por críticas políticas en medios digitales (Freedom House, 2024b).

A nivel local también se ha documentado el uso de discursos de odio como herramienta política para polarizar elecciones y deslegitimar candidatos, esto se evidencia más en campañas de elecciones seccionales, que son el foco para la proliferación de mensajes que fomentan noticias

falsas y crean noticias en forma de sátira para desacreditar la imagen de los personajes políticos. Esta realidad muestra coincidencia con patrones nacionales donde los discursos de odio y la violencia digital erosionan la deliberación democrática y se entrelazan con estrategias comunicativas electorales, que se amplifican por redes sociales, como X, que favorecen las emociones negativas y la polarización. Como efecto se ven reducidos los espacios que fomenten un debate informado y por el contrario se acentúan narrativas simplistas, que se alejan de las necesidades reales de la población.

Con base en lo mencionado anteriormente, se formuló la siguiente pregunta de investigación: ¿cómo el discurso de odio desde la red social X acentuó la parte negativa de los candidatos presidenciales de Ecuador en 2025 tras el segundo debate? Con miras a responder dicha interrogante se planteó como objetivo de estudio analizar los discursos de odio en la red social X tras el segundo debate presidencial en Ecuador en 2025. Con este planteamiento el presente trabajo busca, en primera instancia, identificar los discursos de odio en X tras el segundo debate, su distribución y por categorías, así como los marcos léxicos dominantes; en segundo término, analizar la evolución temporal de la tasa de odio y de sus categorías durante el periodo de estudio; y, finalmente, evaluar la repercusión de los posts de odio en la audiencia mediante métricas de interacción digital.

Para alcanzar estos objetivos, el presente trabajo se articula de la siguiente manera: en primer lugar, se desarrolla la revisión de literatura, la cual profundiza en las dinámicas de polarización en redes sociales, particularmente X, en Ecuador; en este apartado, se fundamentan y delimitan teóricamente las categorías de odio (insulto, difamación, amenaza y descortesía) que se analizan en este estudio. En segundo término, se detalla la metodología, describiendo el procesamiento del corpus mediante el modelo RoBERTuito y un diseño de gate binario que integra variables emocionales e ironía. Posteriormente, se exponen los resultados, analizando la prevalencia por categorías y la dinámica temporal, lo que permite identificar el cambio composicional y los picos de agresividad en el flujo de datos. Finalmente, el estudio culmina con una discusión y una conclusión que reflexiona sobre la institucionalización de la violencia digital como un recurso sistémico de la cultura política ecuatoriana.

Revisión literaria

Discurso de odio

El discurso de odio se ha establecido como una declaración pública, consciente y deliberada, destinada a denigrar a un grupo de personas, la recomendación de la Comisión Europea contra el Racismo y la Intolerancia n.º 15 (sobre la lucha contra el discurso de odio), del 8 de diciembre de 2015, también cita el odio, la humillación o el desprecio hacia una persona que pertenece a un grupo. Estos discursos generalmente se orientan a características identificativas, como la raza, el color, la religión, la etnia o la nacionalidad, así como al género, la identidad sexual o la orientación (Paz et al., 2020).

Desde un punto de vista similar, Perera et al. (2023) mencionan que el discurso de odio es aquel que ataca o menosprecia, que incita a la violencia o al odio contra grupos, basándose en características específicas como la apariencia física, la religión, el origen étnico o nacional, la orientación sexual, la identidad de género u otras, y puede manifestarse con diferentes estilos lingüísticos, incluso de forma sutil o cuando se utiliza el humor. En ese sentido, se destaca que este tipo de ataque no puede reflejarse de manera directa, sino que suele utilizarse también enmascarado en el humor o la sátira hacia las personas.

En un sentido más amplio, Castaño-Pulgarín et al. (2021) explican que el discurso de odio en la red son aquellas acciones que incluyen comportamientos que amenazan la democracia, niegan a las personas sus libertades personales o estereotipan a grupos sociales; y el discurso de odio fuera de línea se expresa como agresiones directas contra ideologías políticas, grupos religiosos o minorías étnicas. Por ejemplo, los rumores centrados en la raza y la etnia pueden dar lugar a violencia étnica y las personas ofendidas pueden verse amenazadas por su identidad grupal. Se debe indicar que, a veces, la libertad de expresión se considera una de las causas fundamentales para promover el discurso de odio en línea (Tareen et al., 2021).

Tipos de discursos de odio

El abordaje de los tipos de discurso es un elemento clave dentro de la investigación, para ello se toma el criterio de Wirz y Blassnig (2025), que ha categorizado los discursos de odio en cuatro grupos:

Insultos. Se subclasifica en 9 tipos:

- Insultos relacionados con la identidad grupal: acoso, insultos, humillación o ataques a personas por su pertenencia a un grupo.
- Ataques personales: ataques basados en características personales, rasgos, elecciones, entre otros.
- Deshumanización: agresión verbal deshumanizante, cosificación o nombres despectivos (por ejemplo, relacionados con animales).
- Intolerancia política: desacreditar opiniones o actores políticos opuestos como inválidos o ilegítimos, negar su derecho a existir o acusarlos de tener un impacto social extremadamente negativo.
- Racismo: discriminación, estereotipos o lenguaje de odio basados en la raza o el origen.
- Intolerancia socioeconómica: discriminación, estereotipos o lenguaje de odio basados en la educación, el estatus social o los ingresos.
- Sexismo o intolerancia hacia la libertad sexual: discriminación, estereotipos o lenguaje de odio basados en el género, la orientación o la identidad sexual.
- Intolerancia religiosa: discriminación, estereotipos o lenguaje de odio basados en las creencias religiosas.
- Amenazas a los derechos individuales: negación de la igualdad de derechos a individuos o grupos.

Difamación. Se subclasifican como:

- Mentiras o falsedades: afirmaciones falsas o distorsionadas sobre individuos o grupos basadas en su pertenencia a un grupo.
- Difamación institucional: afirmaciones falsas o distorsionadas, o lenguaje peyorativo hacia políticas, instituciones u organizaciones para menospreciar o cuestionar su credibilidad o legitimidad.
- Estereotipos ofensivos: clichés negativos simplificados sobre una persona o un grupo.

Las amenazas o incitaciones a la violencia. Engloba anuncios de sanciones negativas o agresivas o llamamientos a actos violentos o delitos contra un destinatario.

La descortesía. Se clasifica en seis subcategorías:

- Desprecio/condescendencia: comentarios negativos sobre personas, grupos o ideas que no atacan el honor de nadie.
- Sarcasmo/cinismo: comentarios cínicos, burlas o humor despectivo.
- Gritos: uso de mayúsculas o signos de exclamación.
- Lenguaje profano o vulgar: lenguaje obsceno, soez o grosero que es inapropiado para el discurso profesional.
- Nombres despectivos: etiquetas ofensivas y despectivas para individuos, grupos o instituciones que no se clasifican como discurso de odio.
- Ataques a los argumentos: declaraciones que atacan, descalifican o descartan un punto de vista o argumento.

Efectos del discurso de odio

El discurso de odio genera una variedad de efectos que pueden clasificarse en directos e indirectos. Los efectos directos recaen sobre quienes son destinatarios explícitos del mensaje y afectan su dignidad y su capacidad de participación en la conversación pública. Los efectos indirectos, en cambio, se producen a través de lenguaje codificado o eufemismos que, pese a su aparente sutileza, instalan estereotipos y generan entornos de exclusión y hostilidad. En cualquiera de sus formas, este fenómeno genera consecuencias sociales que vulneran la libertad de expresión de las personas, además de incentivar la agresión hacia grupos o individuos específicos (Tareen et al., 2021).

Cabe añadir que el discurso de odio puede llegar a causar daño a grupos sociales al crear un entorno de prejuicios e intolerancia, que fomenta la discriminación y la hostilidad, que en casos graves facilita actos violentos, como la falta de cortesía, los términos peyorativos, la vulgaridad o el sarcasmo, o la incivilidad (Castaño-Pulgarín et al., 2021). Las personas expuestas a discursos de odio suelen experimentar sentimientos negativos (por ejemplo, ira, tristeza o vergüenza), así como una disminución de los niveles de confianza. La exposición al discurso de odio también puede estar asociada con procesos de radicaliza-

ción política. Las víctimas pueden carecer de estrategias de afrontamiento adecuadas para mitigar el daño del discurso de odio e incluso buscar venganza (Kansok-Dusche et al., 2023).

Discursos de odio en la red social X

En los últimos años, las plataformas de redes sociales se han convertido en un elemento central en la forma en que las personas intercambian información, dan forma al discurso público e influyen en el comportamiento. Si bien plataformas como Twitter/X facilitan la rápida difusión de contenidos, también permiten la propagación de comportamientos nocivos, como el discurso de odio y la desinformación, que pueden tener importantes consecuencias sociales (Ghenai et al., 2025).

En concordancia con lo anterior, se ha dado a conocer que las redes sociales crean entornos ideales para la difusión de contenidos que incitan al odio, porque conectan a usuarios con ideas afines y fomentan las cámaras de eco, donde se refuerzan y polarizan las creencias políticas. Además, los usuarios suelen juzgar erróneamente la diversidad de su audiencia, lo que conduce a un colapso del contexto, que amplifica aún más el odio digital, el ciber-racismo y las perspectivas extremistas (Molina, 2024). Este discurso de odio suele estar impulsado por rasgos antisociales, como el sadismo y la psicopatía (Frischlich et al., 2021), y algunos usuarios están motivados por el deseo de obtener la aprobación social de sus pares (Walther, 2022).

En este marco de ideas se ha dado a conocer que, si bien la desinformación ha sido un elemento recurrente en X, esta se ha visto acentuada posterior a que fue adquirida por Elon Musk en 2022, lo cual puede deberse a cambios en su modelo de negocio y gestión. Pues, según se reporta, desde este suceso han incrementado en un 50 % los aspectos negativos (Gayo Avello, 2025), como un mayor surgimiento del discurso de odio, el aumento de *bots*, cuentas falsas controladas digitalmente que alteran el algoritmo y dan por válidos contenidos falsos que buscan redirigir la agenda política (Díaz et al., 2025). Como efecto se destacó que este contenido de odio en X promueve o fomenta la violencia, el acoso, la discriminación o la hostilidad hacia individuos o grupos por motivos de raza, etnia, origen, orientación sexual, género, identidad de género, afiliación religiosa, edad, discapacidad o enfermedad (Perera et al., 2023).

La política ecuatoriana y los discursos de odio

Conforme con esto, los discursos de odio en Ecuador se ven influenciados por un contexto de violencia pública que llega hasta el ámbito político, que afecta las libertades, el derecho de participación y la democracia. En torno a esto se ha dado a conocer que:

El proceso electoral general en Ecuador durante 2025 se llevó a cabo en un contexto de profunda crisis institucional, altos niveles de violencia política, presencia territorial de bandas criminales y un marcado debilitamiento de la confianza ciudadana en el sistema democrático. Ante esta realidad, el Observatorio Ciudadano de Violencia Política activó un proceso de monitoreo, análisis y documentación de hechos, tendencias y omisiones que marcaron la jornada electoral (Gómez & Mora, 2025, p. 3).

Estos autores añaden datos importantes, que fueron recabados con apoyo del Observatorio Ciudadano de Violencia Política, en los que se destaca que dentro del marco electoral del 2025 se registraron 80 agresiones políticas, de las cuales el 57.5 % fueron asesinatos, el 26.2 % amenazas, el 10 % atentados, el 5 % secuestros y el 1.3 % agresiones físicas (Gómez & Mora, 2025). Estos datos evidencian la magnitud de la violencia política que se traslada a las redes sociales, alimenta discursos de odio de distinta índole e influye en la percepción ciudadana.

Dentro de esta problemática se identifican grupos específicos a quienes se dirige en gran medida este tipo de ataques digitales. Por ejemplo, en los últimos procesos electorales del Ecuador se ha observado un notable incremento de discursos de odio hacia los migrantes, con base en lo cual se ha dado a conocer que:

los inmigrantes pobres son excluidos del orden comunitario por estar frecuentemente asociados a la violencia y al desorden, pero más aún porque su presencia es capaz de poner en evidencia la existencia de las desigualdades sociales y económicas que estos gobiernos pretenden naturalizar e invisibilizar (Pagliarone & Quiroga, 2021, p. 104).

Otro grupo afectado por los discursos de odio en el ámbito político corresponde a las mujeres, que enfrentan frecuentemente estigmas, señalamientos y violencia digital. En torno a esto Ordóñez et al. (2023) señala que:

En Ecuador, las mujeres que participan en política son afectadas y están expuestas a diferentes formas de violencia, ya sea al interior de los partidos políticos, sus familias, comunidades, medios de comunicación y redes sociales, donde, a través de discursos de odio, violentos o discriminatorios se ven expuestas a la difamación, la exclusión, invisibilización, ataques, agresiones, amenazas, entre otros actos vejatorios, que con el fin de menoscabar su imagen pública, atacan contra sus derechos de participación, limitan o anulan sus derechos políticos, y afectan al conglomerado de la sociedad y a la consecución de procesos plurales y democráticos efectivos (p. 282).

De acuerdo con esto, los discursos de odio en Ecuador son una constante dentro de lo político, exacerbado por el contexto de violencia social que se vive actualmente. Esto configura un detonante clave para que se trasladen al entorno digital argumentos cargados de odio hacia grupos determinados, como extranjeros y mujeres, lo que afecta la participación comunitaria y política, además de generar estigmas, que son aprovechados por la clase política para incentivar en mayor medida estos discursos.

La política ecuatoriana y los discursos de odio en X

En Ecuador se hace un uso frecuente de la red social X para difundir contenidos de odio, debido a contextos políticos que conllevan a la polarización. Esto se ha prestado en gran medida para distribuir contenido cargado de desinformación. En las elecciones anticipadas de 2023 “los actores políticos generaron campañas de desinformación dirigidas a los votantes, lo que plantean dudas sobre la responsabilidad de las redes sociales, aprovechándose de los bajos niveles de educación y conocimiento de la población” (Sacoto Cabrera, 2024, p. 149).

En este uso que se le da a la red social en cuestión se destaca una modalidad de ataque digital que corresponde al *trolling* o troleo, como instrumento para difundir

contenido cargado de odio y desinformación, ante esto se dio a conocer que:

el fenómeno del *trolling* durante el balotaje de la campaña electoral en Ecuador se ha hecho evidente a través de las diversas y sofisticadas prácticas de ataque masivo en la red social Twitter. Entre ellas estuvo la bufonería transgresora (a partir de memes); la difusión de contenido difamatorio, rumoroso, alarmante, manipulado (como *deepfakes*) y las llamadas *fake news*; y el uso de insultos, denostaciones y/o lenguaje fuerte o soez, dirigido tanto a los candidatos para desvalorizar su imagen y generar un rechazo masivo hacia ellos, así como a cualquier otro usuario que piense de manera distinta, con el fin de censurarlo, intimidarlo o minimizar ciertas expresiones (Martínez Orellana, 2025, p. 185).

De igual manera, en el país es frecuente el uso de la red social X para difundir discursos de odio hacia mujeres que participan en la política. En torno a ello se han observado patrones comunes en este tipo de ataques, en los que se observa con mayor recurrencia “menosprecio de capacidades intelectuales, roles de género, y alusiones al cuerpo y desprestigio sexual. Evidenciando los estereotipos que la sociedad tiene hacia las candidatas, haciendo énfasis en sus cuerpos, vidas privadas y labores otorgadas solamente a mujeres” (Peñarreta Feijoo & Mier Sanmartín, 2024, p. 73).

Como ejemplo de ello se toma el caso de las elecciones generales del 2019, donde solo en la red social X se registraron 189 expresiones de odio hacia las mujeres (Observatorio Ciudadano de Violencia Política, 2024). También se ha indicado que estos discursos se acompañan de hashtags que suelen tener un impacto significativo en la opinión pública, lo que lleva a que se genere tanto debate como polarización entre los usuarios (Chica-Pincay, 2025).

Materiales y métodos

Este estudio combina análisis de contenido y técnicas de lingüística de corpus para describir y explicar los discursos hostiles en X durante la segunda vuelta presidencial de 2025. Se trabajó sobre un corpus depurado de trinos sin

retuits textuales, vacíos ni duplicados, y se delimitó el análisis al subconjunto previamente identificado como “odio”. La lectura del lenguaje se enfocó en actos de agresión verbal, delimitados, según Wirz y Blassnig (2025), en insulto, descortesía, difamación y amenaza; y en el tono emocional en ira, asco, miedo, desprecio; se combinaron los indicios léxicos, de estilo y de contexto para decidir casos y matices sin depender de una sola señal. Para extraer sentido colectivo, se resumió la superficie del discurso con frases clave y asociaciones de términos, y se esbozaron temas recurrentes con ejemplos representativos. Como referencia técnica, se consideró RoBERTuito (modelo para español en X/Twitter) para calibrar y contrastar señales de odio y emoción, usándolo como apoyo metodológico y no como fuente única de decisión. La implementación se realizó en Python dentro de VS Code. El método técnico de análisis se especifica en el Apéndice.

Disponibilidad de los datos

El corpus de trinos utilizado en este análisis se encuentra alojado en el repositorio de Harvard Dataverse, bajo el nombre de *Trinos sobre la segunda vuelta presidencial en Ecuador 2025: del segundo debate a las elecciones* (Molina, 2025), el mismo cuenta con 13.173 publicaciones. Se seleccionó ese conjunto de datos por la ecuación booleana empleada en su captura y por su pertinencia analítica respecto a los objetivos del estudio. Por otra parte, el script con el flujo de análisis en Python se encuentra disponible en Figshare para su verificación y reproducibilidad (García, 2025).

Consideraciones éticas

El estudio analiza publicaciones de acceso público en X y aplica criterios orientados a evitar la exposición innecesaria de usuarios y la difusión de contenido sensible. Para ello, se trabajó con datos depurados y, salvo cuando fue indispensable para contextualizar el análisis, se mencionó únicamente a figuras políticas bajo escrutinio público. A lo largo del texto, se evitó identificar cuentas individuales, y los resultados se presentan de manera agregada. Se puso especial atención en no amplificar contenido perjudicial, para ello se limitó la reproducción de insultos al mínimo necesario para el análisis, así como en describir el fenó-

meno sin estigmatizar a colectivos y en cumplir con las condiciones de uso de la plataforma.

Limitaciones

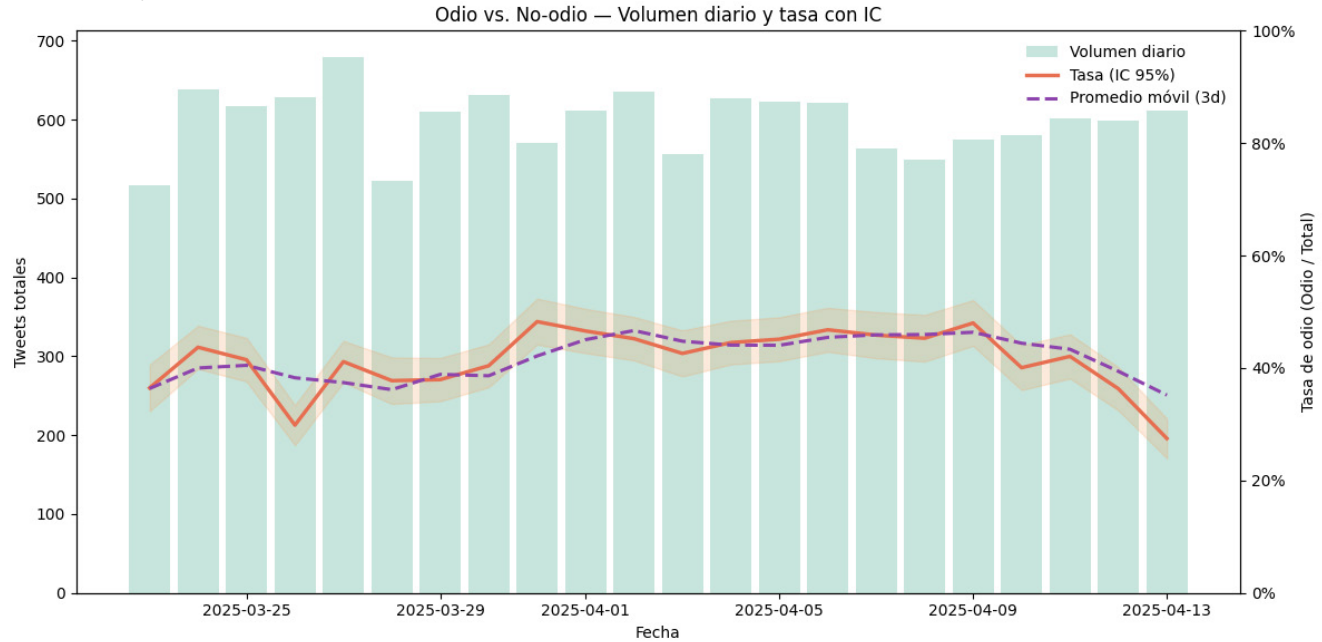
Los hallazgos dependen del recorte temporal, de la ecuación de captura y de la propia naturaleza de la plataforma, por lo que no son exhaustivos ni necesariamente representativos del debate público general. La clasificación automática puede incurrir en errores asociados al sarcasmo, la ambigüedad y los localismos, lo que pudo afectar la subclasificación de los tipos de odio. Además, el análisis se limita a publicaciones en X, por lo que los resultados pueden estar afectados por sesgos de cobertura y no ser plenamente representativos del conjunto de la población.

Resultados

En el conjunto de datos se identificaron 5459 tuits (41.44 %) clasificados como odio y 7714 (58.56 %) como no-odio. Esta clasificación se obtuvo mediante un *gate* automático, que combinó la probabilidad de odio estimada por RoBERTuito con señales emocionales de ira, asco y miedo; un tuit se consideró no-odio cuando tanto $p(hate)$ como dichas emociones estuvieron por debajo de umbrales predefinidos y calibrables, en caso contrario se clasificó como odio. La evolución temporal del fenómeno se presenta en la Figura 1, donde se detalla, día a día, el volumen total y la proporción de mensajes de odio con sus respectivos intervalos de confianza de Wilson (95 %).

En el periodo analizado, la Figura 1 muestra un repunte acotado de la tasa de odio en la primera semana de abril, del 1 al 6 de abril la tasa media fue 0.452, frente al 23 al 27 de marzo con una media de 0.386. Esta etapa de debate y post-debate mantuvo una baja proporción de odio en la conversación, pero los marcos agresivos se consolidaron y difundieron en los días posteriores, especialmente justo antes del cierre de campaña. La diferencia absoluta entre ambos tramos es de + 6.6 puntos porcentuales, lo que indica un cambio de tono con mayor proporción de odio, no un aumento de volumen, sino de agresividad. Luego, al acercarse y pasar la fecha de la votación, el 13 de abril, el tono desciende; se observa una meseta del 7 al 9 de abril y una baja visible hacia el 10 al 13 de abril. Con el volumen diario estable, el cambio es

Figura 1
Volumen diario y tasa con IC de odio vs. no-odio



composicional, es decir, no hubo más tuits en total, sino una mayor proporción de mensajes con carga de odio en esos días de inicio de abril. Se puede sugerir que al entrar en la etapa final y el día de la elección el 13 de abril no se concentra el máximo de odio: la tasa retrocede respecto del pico de inicios de abril, que se ve por debajo de la línea punteada de esos días. Esto sugiere que el pico no está directamente relacionado con la jornada electoral en sí, sino con un aumento de tensión o polarización en los días previos a la votación.

Con la distribución de discursos de odio identificada, se desagregaron los trinos clasificados como odio según la tipología de Wirz y Blassnig (2025), que incluye insultos, difamación, amenazas y descortesía. Esta desagregación presenta la granularidad que se detalla en la Tabla 1.

Como muestra la Tabla 1, los discursos de odio durante el balotaje de Ecuador se expresan con mayor tendencia como insultos; la categoría amenazas —definida como el llamamiento a actos violentos o delitos que representen daños físicos o materiales hacia alguien— es relativamente marginal en el corpus analizado.

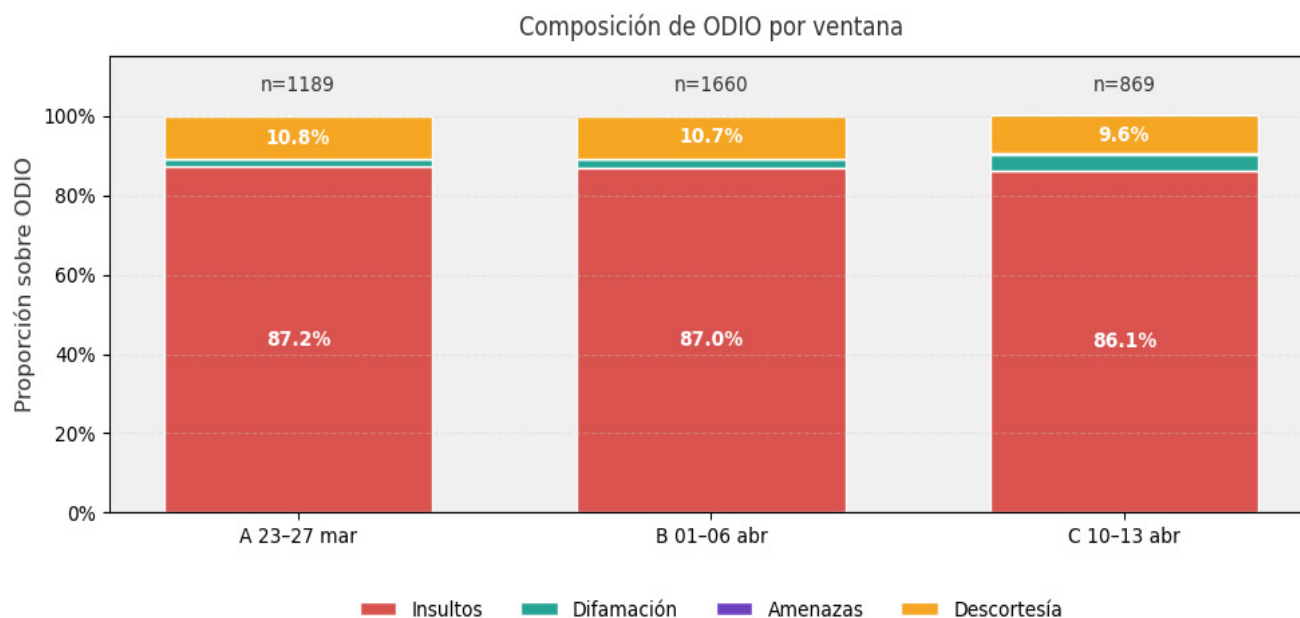
Para examinar en detalle las tendencias identificadas en la Figura 1, se analizó una serie temporal de tres

Tabla 1
Distribución global por categoría de trinos con carga de odio

Categoría	Cantidad	Porcentaje
Insultos	4747	87.0 %
Descortesía	567	10.4 %
Difamación	143	2.6 %
Amenazas	2	≈ 0 %

ventanas clave representadas en la Figura 2. Este análisis permite verificar un patrón donde los insultos dominan la conversación, oscilando entre el 85 y el 90 % a lo largo del periodo del 23 de marzo al 13 de abril, con picos alrededor del 1 al 6 de abril, que reflejan el repunte general de odio observados previamente en el volumen diario y en la tasa con IC de odio vs. no-odio. Descortesía, por su parte, fluctúa entre el 8 y el 14 %, con variaciones menores y sin una tendencia definida. En cuanto a la difamación, esta se mantiene entre el 2 y el 6 %, registra un leve aumento hacia el final del periodo del 10 al 13 de abril, lo que podría indicar un incremento en acusaciones

Figura 2
Ventanas clave según proporción de cada categoría dentro de odio



o rumores cercanos a la votación. La categoría amenazas, con una representación prácticamente nula (cercana a 0 %), resulta casi inocua en el análisis. Estos hallazgos muestran que el aumento de trinos con cargas de odio previo a la elección, destacado en la Figura 1, se debe a los insultos, con un aporte secundario de difamación durante el pico, seguido de una moderación del tono de la conversación en la fase de días antes de las elecciones marcada por la disminución de insultos.

La evolución diaria observable en la Figura 3 ratifica que los insultos se mantienen como la categoría predominante, oscilando entre 85 y 90 % a lo largo del periodo del 23 marzo al 13 de abril, con picos leves alrededor del 1 al 6 de abril, que reflejan el repunte general de odio. Descortesía varía entre 8 y 14 %, con fluctuaciones menores, pero sin una tendencia clara. Difamación se mantiene entre el 2 y 6 %, con un leve aumento hacia el final del periodo del 10 al 13 de abril. Amenazas se mantiene con una representación cercana a 0 %, permaneciendo prácticamente ausente, lo que corrobora su irrelevancia en los trinos analizados.

En cuanto a la repercusión de los mensajes, los datos de la interacción por categoría se presentan en la Tabla 2, aunque la difamación no es la categoría con mayor representatividad, en esta tabla se destaca que posee un 44.8 % de tuits con al menos un favorito, y un percentil 95 de 75 de interacción con “Me gusta”, esto indica una alta capacidad de difusión y alcance cuando aparecen trinos con esta carga de odio. Seguidamente, insultos, pese a su predominio, tienen un *engagement* más moderado con un 38.1 % y P95 = 21; mientras que descortesía muestra un comportamiento intermedio, con 40.4 % de *engagement* y un P95 = 22. Amenazas con cero interacciones refuerza su marginalidad. Esta mayor resonancia de la difamación puede explicarse por su naturaleza, basada en afirmaciones falsas o distorsionadas, como noticias o coyunturas (Wirz & Blassnig 2025), que, en una red social como X, con cámaras de eco y sesgos de confirmación, tiende a mantener una interacción elevada, lo cual es coherente con el alto contagio observado.

Figura 3
Evolución diaria por categoría de odio

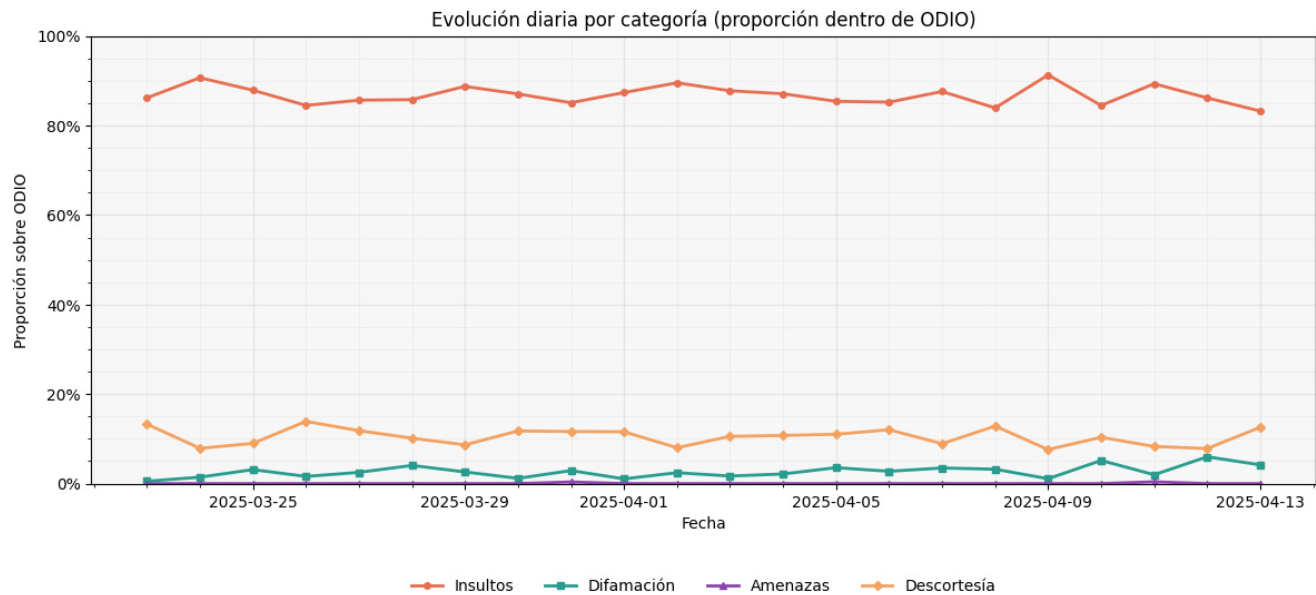


Tabla 2
Engagement (*me encanta*) por categoría

Categoría	N	% tuits con ≥ 1 favorito	P95 favoritos
Insultos	4747	38.1	21
Difamación	143	44.8	75
Amenazas	2	0.0	0
Descortesía	567	40.4	22

Nota. N: número de interacciones con el botón Me gusta. % tuits con ≥ 1 favorito: indica la proporción de tuits de la categoría que reciben al menos un favorito (probabilidad mínima de interacción). P95 favoritos: percentil 95 de favoritos: el 95 % de los tuits queda en o por debajo de esa cifra; resume la cola alta de *engagement* sin sesgo por extremos.

El léxico del odio se visualiza en la Figura 4, a través de una nube de palabras que se elaboró con trinos etiquetados como odio por un *gate* y, tras depurar el texto y contar términos individuales y coocurrencias frecuentes se muestra un léxico dominado por marcos de deslegitimación e insulto como “títere”, “vago”, “incendiario”, junto con criminalización con términos como “narco delictiva”,

“más corrupto”, y racialización “indio salvaje”, consignas de cierre “nunca más”, y parte de propuestas de campaña como “gestores paz” y apodos “rana rené”. La densidad de etiquetas y frases que remiten a la candidatura de González y a actores identificados con la izquierda sugiere una asimetría, donde una parte relevante del odio circuló dirigida hacia ese campo político. No implica que no existan ataques en sentido contrario, pero sí que, en el periodo y el corpus analizados, los marcos agresivos más visibles se concentran alrededor de esa candidatura.

Además, para continuar con el análisis del contenido de odio, se construyeron grafos egocentrados alrededor de cada candidato con trinos etiquetados como odio por el *gate*, de los cuales se extrajeron *tokens* y bigramas, lo que incluyó un filtrado de palabras espurias. Para cada candidatura se tomó el subconjunto de tuits que los mencionan y se contaron coocurrencias documentales entre el nombre del candidato y términos como insulto, criminalización, racialización, consignas y apodos.

En Daniel Noboa, el grafo de la Figura 5 se estructura principalmente en torno a un marco de criminalización y corrupción, en el que “narco” concentra una alta representatividad y se conecta con aristas de mayor peso a

Figura 5
Grafo de coocurrencias de odio alrededor de Daniel Noboa

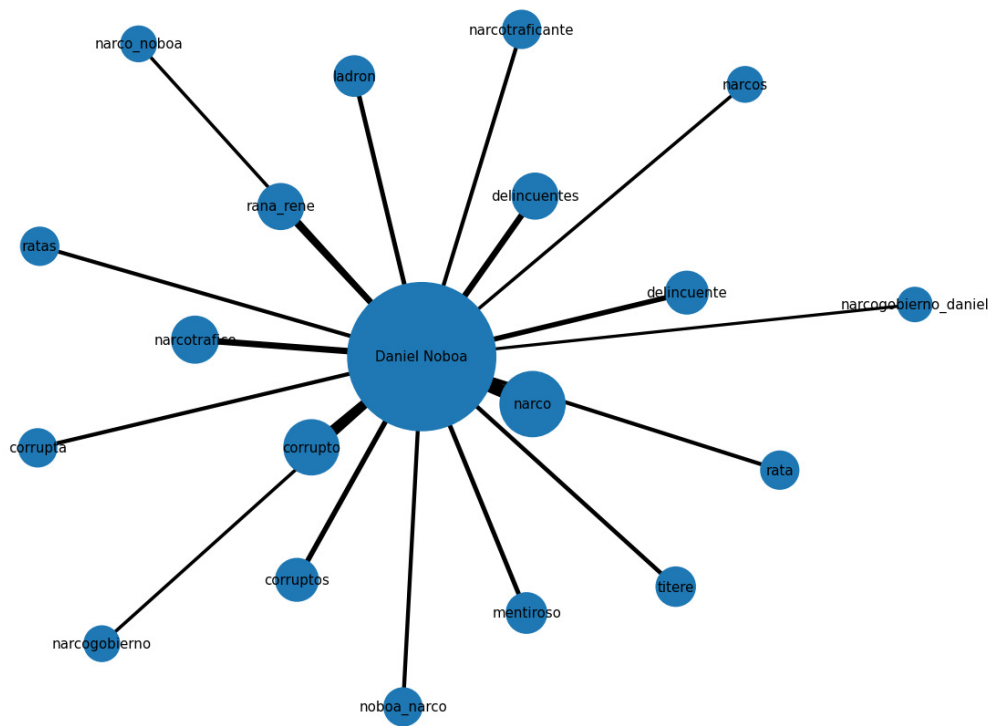
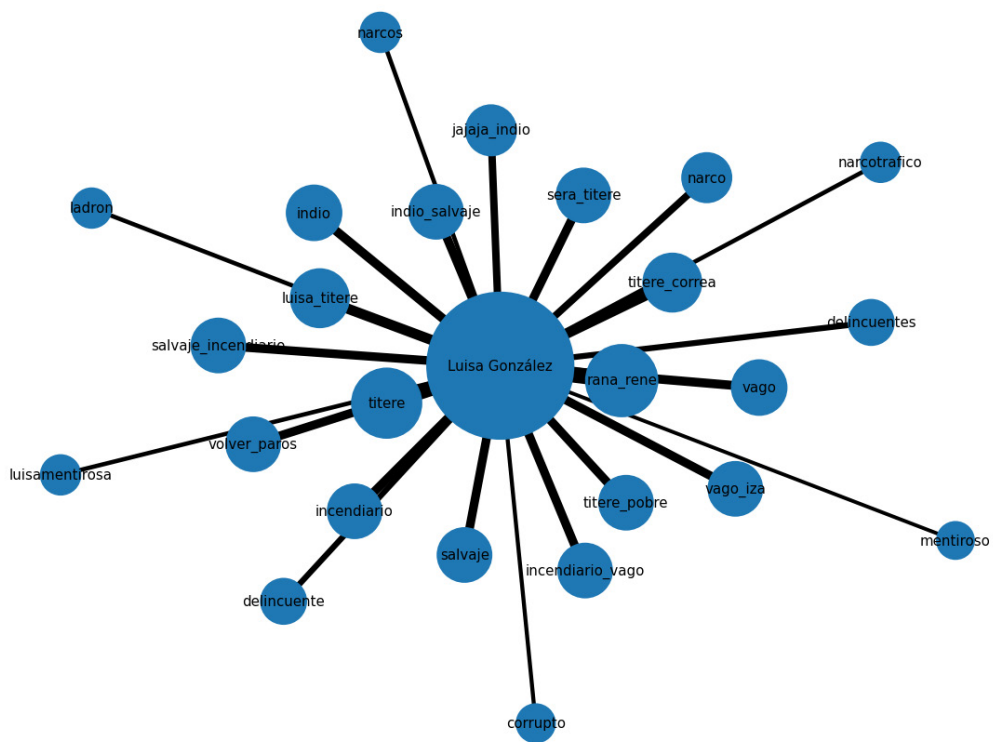


Figura 6
Grafo de coocurrencias de odio alrededor de Luisa González



Conclusiones

Tras el segundo debate presidencial, la conversación en X estuvo marcada por manifestaciones de hostilidad que se expresaron sobre todo como insultos y descortesías como difamaciones personales. En el corpus analizado no se hallaron discursos extremistas, como amenazas hacia el daño físico de alguna persona. La taxonomía aplicada permitió distinguir cuatro formas de agresión y mostró que el tono de conversación dominante se asentó en marcos de deslegitimación e incriminación, con apodos y rótulos que simplifican la discusión pública y desplazan el intercambio argumentado.

El análisis léxico apuntó, además, a una asimetría en los blancos del discurso, con mayor densidad de etiquetas negativas alrededor de una de las candidaturas y de su entorno político. En términos temporales, la tasa de mensajes con carga de odio aumentó en los primeros días de abril y luego descendió al acercarse la jornada electoral del 13 de abril. Este comportamiento sugiere que los picos de agresividad no obedecen al día de la votación en sí, sino a la intensificación de la contienda en los días previos. El cambio observado fue principalmente composicional: no se trató de más volumen total de publicaciones, sino de una mayor proporción de contenidos hostiles dentro de un caudal diario estable.

Respecto de la repercusión en la audiencia, se halló que, aunque los insultos fueron la forma más frecuente, la difamación –minoritaria en términos de presencia– concentró una interacción mayor. Dicho de otro modo, las piezas que introducen acusaciones o rumores plausibles tienen mayor capacidad de atraer atención y propagarse, aun cuando no sean la forma predominante del agravio.

Como propuesta de línea de investigación futura, es pertinente incorporar un componente de aprendizaje supervisado que, a partir de un protocolo de etiquetado humano, entrene modelos capaces de combinar señales de contenido de marcos y postura, red con patrones de interacción, homofilia y centralidad, y metadatos para tres fines complementarios: aproximar la pertenencia a redes ideológicas, identificar cámaras de eco a partir de la baja diversidad de exposición y la alta modularidad comunitaria, y detectar desórdenes informativos.

Referencias

- Castaño-Pulgarín, S., Suárez-Betancur, N., Vega, M., & López, M. (2021). Internet, social media and online hate speech. Systematic review. *Aggression and Violent Behavior*, 58, 101608. <https://doi.org/10.1016/j.avb.2021.101608>
- Chica-Pincay, J. (2025). Impacto de hashtags en las elecciones del Ecuador 2021: Influencia en la opinión pública. *ComHumanitas: Revista Científica de Comunicación*, 16(1), artículo 1. <https://doi.org/10.31207/rch.v16i1.527>
- Cruz, J. (2025, 24 de febrero). *Discursos de odio en redes sociales tras las elecciones 2025*. Observatorio de Comunicación (Odecom). <https://odecom.puce.edu.ec/discursos-de-odio-en-redes-sociales-tras-las-elecciones-2025/>
- Díaz, J. B., Vicario, L. M., & Sobrepere, M. P. (2025). De Twitter a X: Orígenes, desinformación y el éxodo de medios tradicionales. Estudio de caso de los medios The Guardian y La Vanguardia. *Hipertext.net*, 30, artículo 30. <https://doi.org/10.31009/hipertext.net.2025.i30.07>
- Díez-Gutiérrez, E.-J., Verdeja, M., Sarrión-Andaluz, J., Buendía, L., & Macías-Tovar, J. (2022). Discurso político de odio de la ultraderecha desde Twitter en Iberoamérica. *Comunicar: Revista Científica de Comunicación y Educación*, 30(72), 101-113. <https://doi.org/10.3916/C72-2022-08>
- Freedom House. (2024a). *Argentina: Freedom on the Net 2024 country report*. <https://goo.su/fe2UR>
- Freedom House. (2024b). *Ecuador: Freedom on the Net 2024 country report*. <https://goo.su/iKv7mY>
- Frischlich, L., Schatto-Eckrodt, T., Boberg, S., & Wintterlin, F. (2021). Roots of Incivility: How Personality, Media Use, and Online Experiences Shape Uncivil Participation. *Media and Communication*, 9(1), 195-208. <https://doi.org/10.17645/mac.v9i1.3360>
- García, M. F. (2025). *Conjunto de datos - Discursos de odio en la red social X tras el segundo debate presidencial en Ecuador 2025* [Conjunto de datos]. Figshare. <https://doi.org/10.6084/m9.figshare.30005530.v1>
- Gayo Avello, D. (2025, 12 de febrero). *Hate speech has increased by 50 % on the social network X after its purchase by Elon Musk*. Science Media Centre España. <https://sciencemediacentre.es/en/hate-speech-has-increased-50-social-network-x-after-its-purchase-elon-musk>
- Ghenai, A., Noorian, Z., Moradisani, H., Abadeh, P., Erentzen, C., & Zarrinkalam, F. (2025). Exploring hate speech dynamics: The emo-

- tional, linguistic, and thematic impact on social media users. *Information Processing & Management*, 62(3), 104079. <https://doi.org/10.1016/j.ipm.2025.104079>
- Gómez, A., & Mora, D. (2025). *Violencia política en las Elecciones Generales 2025 - Ecuador*. Observatorio Ciudadano de Violencia Política (OCVP); Asociación de Mujeres por la Equidad de Género y la Autonomía (MEGA). <https://library.fes.de/pdf-files/bueros/quito/22084.pdf>
- Kansok-Dusche, J., Ballaschk, C., Krause, N., ZeiBig, A., Seemann-Herz, L., Wachs, S., & Bilz, L. (2023). A Systematic Review on Hate Speech among Children and Adolescents: Definitions, Prevalence, and Overlap with Related Phenomena. *Trauma, Violence & Abuse*, 24(4), 2598-2615. <https://doi.org/10.1177/15248380221108070>
- Martínez Orellana, R. C. (2025). *Análisis del troleo político a partir de la identificación de cuentas troll en la circulación de hashtags controversiales posicionados como tendencia en Twitter durante el balotaje de la campaña electoral en Ecuador 2021* [Tesis de maestría, Flacso Ecuador]. Repositorio Digital Flacso Andes. <http://repositorio.flacsoandes.edu.ec/handle/10469/23597>
- Molina, O. (2024). Reseña del libro El entorno digital: Breve manual para entender cómo vivimos, aprendemos, trabajamos y pasamos el tiempo libre hoy, de P. Boczkowski y E. Mitchelstein. *Estudios sobre el Mensaje Periodístico*, 30(3), 683-684. <https://doi.org/10.5209/emp.96552>
- Molina, O. (2025). *Trinos sobre la segunda vuelta presidencial en Ecuador 2025: del segundo debate a las elecciones* [Conjunto de datos]. Harvard Dataverse. <https://doi.org/10.7910/DVN/SLSBBE>
- Observatorio Ciudadano de Violencia Política. (2024). *Violencia política en las elecciones anticipadas de 2023*. Friedrich-Ebert-Stiftung. <https://library.fes.de/pdf-files/bueros/quito/21063.pdf>
- Ordóñez, A., Espinoza, F., & Vizuete, C. (2023). Libertad de expresión, discurso de odio y violencia política de género en procesos electorales. *Revista Enfoques de la Comunicación*, 10, 249-296. <https://doi.org/10.1000/j34nh087>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2024). *Global report on hate speech online*. <https://www.unesco.org/>
- Pagliarone, M., & Quiroga, M. (2021). Discursos políticos de odio en Argentina y Ecuador. El inmigrante pobre como otredad. *Revista IUS*, 15(47), 103-132. <https://doi.org/10.35487/rius.v15i47.2021.663>
- Paz, M., Montero-Díaz, J., & Moreno-Delgado, A. (2020). Hate Speech: A Systematized Review. *Sage Open*, 10(4), 2158244020973022. <https://doi.org/10.1177/2158244020973022>
- Peñarreta Feijoo, P. S., & Mier Sanmartín, A. C. (2024). *Discursos de odio hacia candidatas a presidencia y asamblea nacional en las elecciones nacionales del Ecuador 2023 en la red social X* [Tesis de Grado, Universidad Técnica Particular de Loja]. Repositorio Institucional. <https://dspace.utpl.edu.ec/handle/123456789/68075>
- Perera, S., Meedin, N., Caldera, M., Perera, I., & Ahangama, S. (2023). A comparative study of the characteristics of hate speech propagators and their behaviours over Twitter social media platform. *Heliyon*, 9(8), e19097. <https://doi.org/10.1016/j.heliyon.2023.e19097>
- Reuters Institute. (2025). *Digital News Report 2025: X as news source in the U.S*. <https://goo.su/yiFr>
- Rodríguez-Peral, E. M., Gómez Franco, T., & Rodríguez-Peral Bustos, D. (2025). Propagation of Hate Speech on Social Network X: Trends and Approaches. *Social Inclusion*, 13, 9317. <https://doi.org/10.17645/si.9317>
- Sacoto Cabrera, E. J. (2024). Redes sociales, violencia política, enajenación social y su normalización en el Ecuador. En *Violencia. Dimensiones e implicaciones* (1ª ed., pp. 139-165). SPUE. <https://doi.org/10.17163/abyaups.58.461>
- Singh, K. (2025, 10 de febrero). *Anti-minority hate speech in India rose by 74% in 2024, research group says*. Reuters. <https://www.reuters.com/world/india/anti-minority-hate-speech-india-rose-by-74-2024-research-group-says-2025-02-10/>
- Singh, S. (2025, 19 de noviembre). *How Many People Use Twitter (X) - Daily & Monthly Users*. DemandSage. <https://www.demandsage.com/twitter-statistics/>
- Tareen, M. K., Tareen, H. K., Noreen, S., & Tariq, M. (2021). Hate Speech and social media: A Systematic Review. *Turkish Online Journal of Qualitative Inquiry*, 12(8), 1086-1094. <https://www.google.com/search?q=https://www.researchgate.net/profile/Muhammad-Tariq-41/publication/354545789>
- Verma, A., Sear, R., & Johnson, N. (2024). How U.S. Presidential elections strengthen global hate networks. *Npj Complexity*, 1(1), 18. <https://doi.org/10.1038/s44260-024-00018-8>
- Walther, J. B. (2022). Social media and online hate. *Current Opinion in Psychology*, 45, 101298. <https://doi.org/10.1016/j.copsyc.2021.12.010>
- Wirz, D. S., & Blassnig, S. (2025). A matter of perception? Investigating subjective and objective exposure to hate speech with a survey and mobile longitudinal linkage study. *Information, Communication & Society*, 28(4), 723-743. <https://doi.org/10.1080/1369118X.2025.2461646>
- Ye, J., Luceri, L., & Ferrara, E. (2025). Auditing Political Exposure Bias: Algorithmic Amplification on Twitter/X During the 2024 U.S. Presidential Election. *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, 2349-2362. <https://doi.org/10.1145/3715275.3732159>

Cómo citar: García Mero, M. F., Trámpuz, J. P., & Molina-Bailon, O. (2026). Discursos de odio en la red social X tras el segundo debate presidencial en Ecuador 2025. *Dixit*, 40, e4996. <https://doi.org/10.22235/d.v40.4996>

Financiamiento: Este estudio no recibió ninguna financiación externa ni apoyo financiero.

Disponibilidad de datos: El conjunto de datos que apoya los resultados de este estudio se encuentra disponible en el repositorio de Figshare con el nombre: *Conjunto de datos - Discursos de odio en la red social X tras el segundo debate presidencial en Ecuador 2025*. Figshare, Dataset. <<https://doi.org/10.6084/m9.figshare.30005530.v1>>.

Conflicto de interés: Los autores declaran no tener ningún conflicto de interés.

Contribución de los autores (Taxonomía CRediT): 1. Conceptualización; 2. Curación de datos; 3. Análisis formal; 4. Adquisición de fondos; 5. Investigación; 6. Metodología; 7. Administración de proyecto; 8. Recursos; 9. Software; 10. Supervisión; 11. Validación; 12. Visualización; 13. Redacción: borrador original; 14. Redacción: revisión y edición.

M. F. G. M. ha contribuido en 1, 5, 6, 13, 14; J. P. T. en 1, 3, 5, 6, 10, 11; O. M.-B. en 2, 3, 6, 9, 12.

Editora responsable: A. L.

Apéndice

Se trabajó con Python. El manejo del corpus (estructuración tabular, agregaciones, percentiles y transformaciones) se realizó con Pandas y NumPy; las visualizaciones (barras, líneas, escalas porcentuales e intervalos) con matplotlib. La inferencia automática en español se obtuvo con pysentimiento (modelo RoBERTuito), del que se derivó por tuit la probabilidad de discurso de odio (p_hate), las probabilidades emocionales de ira, asco y miedo, y, cuando estuvo disponible, ironía (p_irony). Para la tendencia temporal se empleó statsmodels (GLM binomial), y para las reglas lingüísticas auxiliares, expresiones regulares y utilidades numéricas básicas. El flujo fue directo: tras cargar el CSV, se normalizaron las fechas al huso local y se limpiaron las métricas. A cada trino se le calculó el p_hate, emo_anger, emo_disgust, emo_fear y, si existía, p_irony. Con esas señales se aplicó un gate binario conservador: se consideró no-odio cuando p_hate < 0.50 y las emociones estaban por debajo de 0.35 (si había ironía, además se exigía p_irony < 0.60); en caso contrario, el tuit se etiquetó como odio. Dentro de ese subconjunto, la subclasificación en amenaza, difamación, insulto y descortesía se resolvió con reglas priorizadas y umbrales, amenazas por patrón explícito o por ira ≥ 0.60 junto con p_hate ≥ 0.60 y segunda persona; difamación por acusación explícita o por asco ≥ 0.55 o miedo ≥ 0.60 con p_hate ≥ 0.50 ; insulto si ira ≥ 0.50 o asco ≥ 0.50 con p_hate ≥ 0.50 ; y descortesía como residuo cuando hay agresión verbal sin cumplir los criterios anteriores (umbrales de referencia: TH_ANGER_INS = 0.50; TH_DISG_INS = 0.50; TH_DIF_DISG = 0.55; TH_DIF_FEAR = 0.60; TH_AMEN_ANG = 0.60; TH_PHATE_STR = 0.60). Para la dinámica temporal se agregó por día el volumen total y el de odio, se calculó la tasa diaria con intervalo de Wilson (95 %) y se trazó la serie con promedio móvil de 3 días. Las comparaciones por ventanas se estimaron con bootstrap del intervalo de la diferencia de proporciones (y una prueba z como verificación paramétrica). La tendencia se modeló con GLM binomial (tiempo en días como predictor, ponderado por volumen), lo que reportó coeficiente, p-valor y cambio estimado a 7 días. Finalmente, se midió el engagement por categoría (porcentaje con ≥ 1 favorito y percentil 95 de favoritos) y se presentó de forma comparativa.