



ID de Artículo: SLJ-Vol.3.N.1.002.2026

Original

Crímenes de odio digitales desde el derecho penal

Digital hate crimes from a criminal law perspective

Autores:

Olivia Elizabeth Álvarez Montalván¹, Perla Elizabeth Ventura Ramos²

¹ Universidad Autónoma de Guerrero, Estado de Guerrero, México, consejera1969@hotmail.com, <https://orcid.org/0009-0009-1398-2027>

² Universidad Autónoma de Guerrero, Estado de Guerrero, México, ventura-eliza31@hotmail.com, <https://orcid.org/0000-0001-8680-1703>

Autor de Correspondencia: Olivia Elizabeth Álvarez Montalván,
consejera1969@hotmail.com

Recepción: 02-October-2025 Aceptación: 20-November-2025 Publicación: 10-January-2026

How to cite this article:

Álvarez Montalván, O. E., & Ventura Ramos, P. E. (2026). Crímenes de odio digitales desde el derecho penal. *Sapiens Law and Justice*, 3(1), 1-17. <https://doi.org/10.71068/8p0xkx28>



©2026 por los Autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones bajo una licencia internacional Creative Commons Atribución 4.0. (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>



Resumen

El artículo analizó los crímenes de odio cometidos en redes sociales, motivados por razones de género, raza e ideología, en el contexto mexicano. El objetivo fue examinar sus efectos en las víctimas y los desafíos que plantearon para el derecho penal, tomando como caso de estudio la ciudad de Chilpancingo, Guerrero. La investigación se desarrolló desde un enfoque cualitativo con alcance descriptivo-analítico, mediante entrevistas semiestructuradas a personas vinculadas al ámbito académico y jurídico local, así como a través del análisis documental de normas, criterios jurisprudenciales e informes especializados en violencia digital y derechos humanos. Los resultados mostraron que plataformas como Facebook y WhatsApp concentraron las principales prácticas de desacreditación, mientras que el odio por género y el político-ideológico se identificaron como las formas más recurrentes de agresión. Asimismo, se evidenció una brecha significativa entre la experiencia del daño y su reconocimiento jurídico, reflejada en la incertidumbre sobre las posibilidades de denuncia y sanción. Se concluyó que los crímenes de odio digitales constituyeron expresiones de violencia estructural que afectaron la dignidad y la participación pública, lo que planteó la necesidad de respuestas jurídicas más claras y contextualizadas.

Palabras clave: crímenes de odio digitales; violencia digital; derecho penal; discurso de odio; dignidad humana.

Abstract

The article analyzed hate crimes committed on social media, motivated by gender, race, and ideological factors, within the Mexican context. The objective was to examine their effects on victims and the challenges they posed for criminal law, using the city of Chilpancingo, Guerrero, as a case study. The research adopted a qualitative approach with a descriptive-analytical scope, based on semi-structured interviews with participants linked to the local academic and legal fields, as well as documentary analysis of legal norms, jurisprudential criteria, and specialized reports on digital violence and human rights. The results showed that platforms such as Facebook and WhatsApp concentrated the main forms of discrediting practices, while gender-based and political-ideological hate were identified as the most recurrent forms of aggression. Likewise, a significant gap was identified between the experience of harm and its legal recognition, reflected in uncertainty regarding reporting and sanction mechanisms. It was concluded that digital hate crimes constituted expressions of structural violence that affected dignity and public participation, highlighting the need for clearer and context-sensitive legal responses.

Keywords: digital hate crimes; digital violence; criminal law; hate speech; human dignity.



1. INTRODUCCIÓN.

En los últimos años, las redes sociales han dejado de ser solo espacios de interacción cotidiana para convertirse en ámbitos centrales de deliberación pública, visibilidad social y disputa simbólica. En México, esta expansión digital tiene un peso evidente: en 2023, 87.9 millones de personas de 12 años y más usaron internet, equivalentes al 82.4 % de esa población; además, 18.4 millones de personas usuarias reportaron haber vivido alguna situación de ciberacoso, con una incidencia mayor en mujeres que en hombres (Instituto Nacional de Estadística y Geografía [INEGI], 2024).

Dentro de ese escenario, los ataques motivados por razones de género, raza, ideología, origen étnico u otras condiciones identitarias adquieren una relevancia especial, porque no se agotan en el desacuerdo entre personas usuarias ni en el exceso verbal propio de la discusión en línea. Se trata de agresiones que buscan degradar, intimidar, desacreditar o expulsar simbólicamente a ciertas personas del espacio público digital por lo que son, por lo que representan o por la comunidad con la que se les asocia. En esa línea, Perry (2001) ha sostenido que el crimen de odio no daña únicamente a la víctima directa, sino que también envía un mensaje de intimidación a la colectividad a la que esa persona pertenece.

Esta comprensión resulta todavía más pertinente cuando el análisis se traslada al ámbito digital. La velocidad de circulación de los mensajes, su permanencia, la posibilidad de anonimato y la reproducción masiva de contenidos amplifican el daño y multiplican el alcance de la agresión (Citron, 2014; Perry & Olsson, 2009). El propio Módulo sobre Ciberacoso 2023 muestra que entre las formas más frecuentes de agresión se encuentran el contacto mediante identidades falsas, los mensajes ofensivos y el envío de contenido sexual no solicitado; también identifica a Facebook como uno de los medios más reportados y reconoce consecuencias como enojo, desconfianza e inseguridad en las víctimas (INEGI, 2024).

En el caso mexicano, además, diversos informes han mostrado que esta violencia afecta con especial intensidad a mujeres con proyección pública, periodistas, defensoras de derechos humanos y otras personas expuestas a debates de interés social. ARTICLE 19 México y Centroamérica (2021) documentó amenazas, campañas de desprestigio, intimidaciones y hostigamientos en redes sociales contra mujeres periodistas, mientras que ONU Mujeres México (2020) advirtió que la violencia digital impacta de forma particular a mujeres y niñas, especialmente a aquellas que participan activamente en entornos digitales.



Desde el punto de vista jurídico, el problema no radica solamente en reconocer que existe daño, sino en determinar cómo debe responder el derecho penal frente a estas conductas sin desbordar sus propios límites. Aquí aparece una tensión decisiva para este trabajo: por una parte, existe la necesidad de proteger la dignidad humana, la igualdad y la seguridad de las personas frente a expresiones que promueven discriminación, hostilidad o violencia; por otra, el derecho penal está sujeto a exigencias estrictas de legalidad, taxatividad e intervención mínima. La Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (OACNUDH, s. f.) ha insistido en la necesidad de valorar el contexto y el umbral de daño, y la Suprema Corte de Justicia de la Nación (SCJN, 2020) ha reconocido que la libertad de expresión no protege manifestaciones de odio que lesionan la dignidad, la igualdad y la no discriminación.

En ese marco, el estado de la cuestión muestra avances importantes, pero también vacíos. Perry (2001) explicó el crimen de odio como una práctica de dominación que proyecta un mensaje de intimidación colectiva; Perry y Olsson (2009) mostraron que esa lógica se reconfigura en el entorno digital; Citron (2014) analizó la profundidad del daño producido en línea; y el Consejo de Europa (s. f.) distinguió entre hate crime y hate speech como categorías relacionadas, pero no equivalentes. Sin embargo, todavía persiste una articulación insuficiente entre el estudio del odio digital y su examen específico desde el derecho penal mexicano.

Por ello, la pertinencia de esta investigación se ubica en un doble plano. Por un lado, busca aportar una lectura jurídicamente fundada y socialmente situada del fenómeno; por otro, pretende examinar cómo los crímenes de odio cometidos en redes sociales, motivados por razones de género, raza o ideología, pueden ser comprendidos e investigados desde el derecho penal a partir de sus efectos reales sobre las víctimas y de los retos jurídicos que plantea el entorno digital.

2. DESARROLLO.

Los crímenes de odio pueden comprenderse como conductas delictivas motivadas, total o parcialmente, por prejuicios dirigidos contra una característica real o percibida de la víctima, como el género, la raza, la religión, la orientación sexual, la identidad cultural o la ideología. Lo que distingue a estos hechos de otras conductas punibles no es solo el resultado material que producen, sino la presencia de una motivación discriminatoria que amplía el alcance del daño (Perry, 2001). En una línea semejante, el Consejo de Europa (s. f.) define el hate crime como un acto criminal motivado por sesgo o prejuicio hacia una persona o grupo.

Esta comprensión resulta útil para el presente trabajo porque evita reducir el problema a conflictos individuales o a simples ofensas personales. Cuando la



agresión se dirige contra alguien por su pertenencia, real o atribuida, a un grupo históricamente vulnerable, el hecho adquiere una dimensión social más amplia: no afecta solo a la persona directamente atacada, sino también al colectivo al que simbólicamente se intimida o descalifica (Perry, 2001).

Los crímenes de odio digitales constituyen una manifestación contemporánea de esos mismos actos cuando se producen a través de internet, redes sociales, plataformas digitales o servicios de mensajería. La literatura especializada ha utilizado el término cyberhate para referirse a la circulación del odio por medios digitales, destacando que el entorno tecnológico facilita una difusión rápida, una audiencia amplia y una permanencia del contenido que intensifica el daño sobre las víctimas y sus comunidades (Perry & Olsson, 2009).

La especificidad de lo digital no elimina la lógica del crimen de odio tradicional, pero sí introduce elementos que el derecho debe considerar con especial cuidado. En redes sociales, un ataque puede multiplicarse en minutos, reaparecer en capturas o réplicas, conservarse durante largos periodos y extenderse a audiencias que exceden por mucho el momento inicial de la agresión. En ese sentido, lo digital no crea por sí solo el odio, pero sí le ofrece condiciones particularmente eficaces para circular y producir afectaciones duraderas (Citron, 2014).

El discurso de odio puede definirse como toda forma de expresión que promueve, incita, legitima o justifica la hostilidad, la discriminación o la violencia contra una persona o grupo por razones identitarias. La OACNUDH (s. f.) ha señalado que su valoración exige un umbral alto y un examen contextual cuidadoso, a fin de no confundir verdaderos actos de incitación al odio con expresiones críticas, polémicas o simplemente ofensivas.

En México, la Suprema Corte de Justicia de la Nación ha reconocido que la libertad de expresión no protege manifestaciones que vulneran la dignidad humana y promueven discriminación o violencia. En el Amparo Directo en Revisión 4865/2018, la Corte sostuvo que las restricciones a la libertad de expresión pueden resultar válidas cuando buscan proteger derechos como la dignidad, la igualdad y la seguridad de las personas frente a manifestaciones de odio (SCJN, 2020).

La violencia digital comprende un conjunto amplio de acciones realizadas mediante medios tecnológicos con el propósito de causar daño emocional, psicológico, social, reputacional o incluso físico. En este universo se incluyen prácticas como el acoso reiterado, las amenazas, la difusión no consentida de contenido, la vigilancia abusiva y la suplantación de identidad. En México, el INEGI (2024) identificó entre las formas más frecuentes de ciberacoso el contacto mediante identidades falsas, los mensajes ofensivos y el envío de contenido sexual no solicitado; a su vez, ONU Mujeres México (2020) ha subrayado que estas agresiones afectan de manera particular a mujeres y niñas.



En esta investigación, la violencia digital se asume como el marco amplio dentro del cual pueden ubicarse los crímenes de odio digitales. Ello significa que no toda violencia en línea constituye por sí misma un crimen de odio ni toda agresión digital tiene automáticamente relevancia penal. Sin embargo, cuando la conducta incorpora un componente de prejuicio, se dirige contra características identitarias de la víctima y produce una afectación significativa, el análisis debe desplazarse hacia una lectura más específica.

Desde la perspectiva del derecho penal, el problema central consiste en determinar de qué manera puede identificarse, probarse y sancionarse la motivación discriminatoria cuando una agresión ocurre en entornos digitales. El Consejo de Europa (s. f.) distingue entre hate crime, donde existe un acto criminal motivado por prejuicio, y hate speech, referido a expresiones dirigidas contra personas o grupos por sus características o estatus; esa distinción ayuda a precisar que el análisis penal no puede perder de vista la conducta base ni sustituirla por valoraciones abstractas sobre el discurso.

Ahora bien, en el entorno digital esta cuestión plantea dificultades adicionales. La fragmentación del lenguaje, la ironía, la circulación masiva de mensajes, la reutilización de contenidos y la dificultad para individualizar autores obligan a un análisis especialmente riguroso. A ello se suma una tensión constitucional inevitable: el derecho penal no puede actuar desconociendo los principios de legalidad, taxatividad e intervención mínima. Por ello, el estudio de la motivación discriminatoria en redes sociales exige una lectura cuidadosa que permita distinguir entre expresiones ofensivas o polémicas y conductas que, por su contexto, contenido, finalidad y efectos, alcanzan un umbral de gravedad jurídicamente relevante (OACNUDH, s. f.; SCJN, 2020).

La dignidad humana constituye el principio rector que atraviesa todo el análisis del artículo. Su importancia no es meramente declarativa, sino interpretativa y práctica: permite comprender por qué ciertas agresiones, aunque se emitan en espacios virtuales, no pueden considerarse simples excesos del lenguaje cuando producen humillación, exclusión o amenazas dirigidas contra la condición misma de la persona (SCJN, 2020).

Asumir la dignidad como eje del análisis evita, además, una lectura puramente formal del derecho penal. El problema no consiste solo en definir si una conducta encaja o no en un tipo legal, sino en reconocer que detrás del odio digital existen experiencias concretas de silenciamiento, miedo y exclusión. En esa medida, la dignidad humana funciona aquí como un criterio que orienta la comprensión del daño y justifica la necesidad de respuestas jurídicas proporcionadas, razonables y respetuosas de los derechos fundamentales en juego.



3. METODOLOGÍA.

El diseño de la investigación adoptó un enfoque cualitativo con alcance descriptivo-analítico, adecuado para examinar un fenómeno jurídico-social a partir de percepciones, experiencias y significados construidos por las personas participantes, así como desde el análisis de fuentes normativas y documentales (Flick, 2015; Hernández Sampieri et al., 2014).

Campo de estudio. Se delimitó al estado de Guerrero, particularmente a la ciudad de Chilpancingo de los Bravo, con el fin de analizar el fenómeno en un contexto local específico y vincular la experiencia situada con la discusión jurídica nacional.

Participantes y muestreo. La población de interés estuvo integrada por personas adultas usuarias de redes sociales vinculadas al ámbito académico y jurídico local, entre ellas docentes universitarios, estudiantes de nivel superior y profesionistas del derecho. La selección se realizó mediante un muestreo intencional no probabilístico, buscando diversidad de género, edad y trayectoria profesional.

Criterios de selección. Se consideraron como razones de inclusión: ser mayor de edad, residir o desarrollar actividades académicas o profesionales en el contexto local estudiado, utilizar redes sociales y contar con experiencia de uso o reflexión informada sobre el fenómeno analizado. Se excluyeron los casos en los que no existió consentimiento para participar o no se contó con información suficiente para aportar al objeto del estudio.

Se realizaron 12 entrevistas semiestructuradas a personas vinculadas al ámbito académico y jurídico en Chilpancingo de los Bravo, distribuidas de la siguiente manera: 4 docentes universitarios, 4 estudiantes de nivel superior y 4 profesionistas del derecho. La selección se llevó a cabo mediante un muestreo intencional no probabilístico, procurando diversidad de género, edad y trayectoria profesional.

Técnicas de recolección. Se emplearon dos métodos principales: entrevistas semiestructuradas y análisis documental. Se emplearon entrevistas semiestructuradas y análisis documental. Las entrevistas se orientaron a identificar formas de odio digital, afectaciones percibidas, conocimiento de mecanismos de denuncia y opiniones sobre la intervención penal. El análisis documental incluyó la revisión de normas jurídicas, criterios jurisprudenciales e informes especializados en materia de violencia digital y discurso de odio.

Instrumentos. Para el trabajo empírico se utilizó un guion de entrevista con preguntas abiertas, organizado en torno a cuatro ejes: formas de agresión en línea, afectaciones percibidas, conocimiento jurídico del problema y valoración de las respuestas institucionales. Para la revisión documental se empleó una



ficha de análisis que permitió sistematizar definiciones, criterios jurídicos, coincidencias, tensiones y vacíos identificados en las fuentes revisadas.

Procedimiento y análisis. La información recolectada se organizó en categorías temáticas previamente definidas y luego ajustadas a partir de la lectura del material: formas de odio digital, afectaciones percibidas, respuestas institucionales, percepción de impunidad, tensiones entre libertad de expresión y discurso de odio, y vacíos legales advertidos en torno al tratamiento penal de estas conductas. La información fue organizada mediante codificación temática, a partir de categorías predefinidas y emergentes: formas de odio digital, afectaciones percibidas, respuestas institucionales, percepción de impunidad, tensiones entre libertad de expresión y discurso de odio, y vacíos legales. El análisis se desarrolló bajo un enfoque interpretativo, articulando los testimonios con el marco teórico y jurídico del estudio.

Rigor metodológico. Para fortalecer la consistencia del análisis se recurrió a la triangulación de fuentes, contrastando la información obtenida en las entrevistas con la revisión normativa, jurisprudencial y documental. Esta estrategia permitió evitar una lectura reducida a un solo plano y aportó mayor solidez a la interpretación final (Flick, 2015; Hernández Sampieri et al., 2014).

Consideraciones éticas. Antes de aplicar los instrumentos se explicó a cada participante el propósito académico del estudio, el uso de la información y las medidas previstas para resguardar su identidad. La participación fue voluntaria, se garantizó la confidencialidad de los datos y se obtuvo el consentimiento informado en todos los casos.

4. RESULTADOS.

Los resultados que se presentan a continuación derivan del análisis cualitativo de las entrevistas y de la revisión documental realizada en el contexto local de Chilpancingo de los Bravo, Guerrero. Aunque el estudio no tuvo un propósito estadístico, sí permitió identificar regularidades en la manera en que las personas participantes reconocen, describen y valoran los crímenes de odio digitales.

Un primer hallazgo fue la percepción de que el odio digital forma parte de prácticas frecuentes y fácilmente reconocibles en redes sociales de uso cotidiano. Las personas participantes no lo describieron como un hecho excepcional, sino como una dinámica reiterada que aparece en conversaciones públicas, publicaciones, comentarios y cadenas de difusión.

Entre las manifestaciones más mencionadas se ubicaron los comentarios degradantes, las burlas reiteradas, las amenazas, las campañas de desprestigio, los mensajes misóginos, las expresiones racistas y los ataques dirigidos contra posturas ideológicas. En varios testimonios, estas prácticas fueron percibidas



como formas de hostilidad que terminan normalizándose dentro del intercambio digital cotidiano.

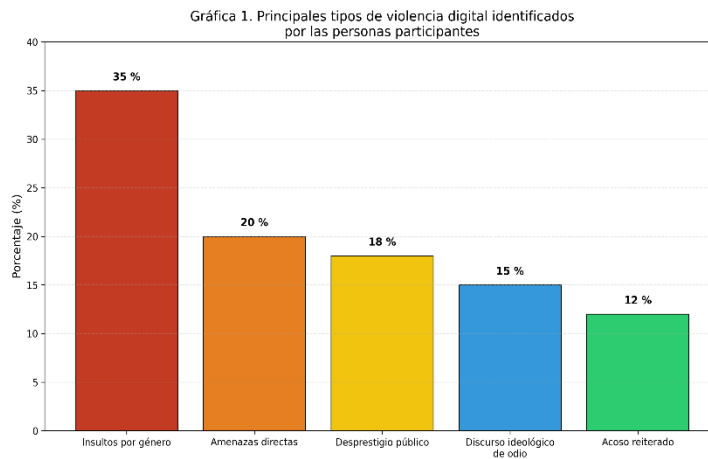
Otro hallazgo relevante fue el reconocimiento del daño real que estas conductas producen. Las personas participantes asociaron las agresiones digitales con miedo, malestar emocional, inseguridad, desgaste psicológico, retraimiento en la participación y, en algunos casos, autocensura. Desde su experiencia, el problema no se agota en la ofensa verbal, sino que repercute en la vida cotidiana y en la manera de habitar los espacios digitales.

Las narrativas recogidas también mostraron que el impacto no se distribuye de manera homogénea. De acuerdo con lo expresado por las personas participantes, la exposición al odio digital se intensifica cuando la presencia en redes está vinculada con actividad académica, participación política, proyección pública o posicionamientos visibles sobre temas sensibles.

En el plano jurídico, uno de los resultados más notorios fue la persistente incertidumbre sobre la posibilidad real de denunciar, investigar o sancionar estas conductas. Las entrevistas mostraron dudas acerca de cuándo una agresión en redes sociales puede considerarse un delito, qué elementos permiten reconocer una motivación discriminatoria y qué tan efectivas resultan las vías institucionales existentes.

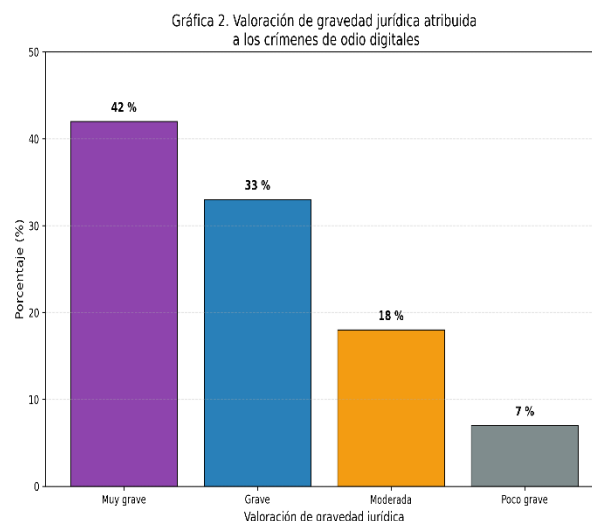
Asimismo, se identificaron obstáculos prácticos vinculados con la dificultad para individualizar a las personas agresoras, la circulación masiva de mensajes, el uso de perfiles falsos y la percepción de que la respuesta institucional suele ser tardía o insuficiente. En términos generales, la distancia entre el daño vivido y su reconocimiento jurídico apareció como una constante en el material analizado.

La gráfica 1 presenta los principales tipos de violencia digital identificados por las personas participantes en el contexto local analizado. La categoría con mayor incidencia percibida corresponde a los insultos por género (35%), seguida de las amenazas directas (20%), el desprestigio público (18%), el discurso ideológico de odio (15%) y el acoso reiterado (12%). En conjunto, estos resultados muestran que la violencia digital no se manifiesta de una sola manera, sino a través de expresiones diversas que combinan agresión verbal, hostigamiento, intimidación y afectación de la imagen pública de las personas.



Nota. Fuente: elaboración propia con base en la sistematización de entrevistas semiestructuradas y revisión documental del estudio.

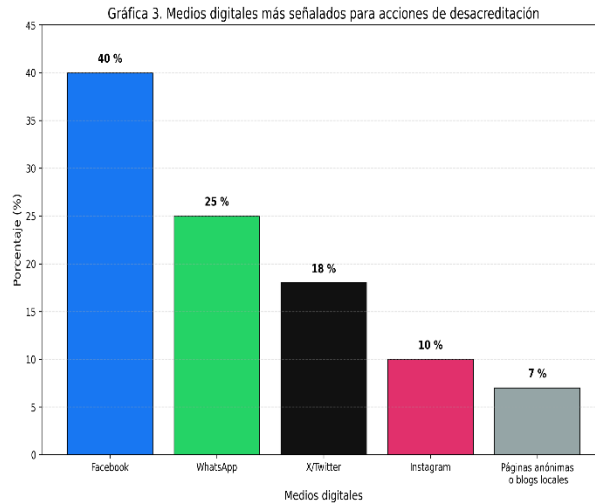
La gráfica 2 muestra la valoración de gravedad jurídica atribuida a los crímenes de odio digitales por las personas participantes. El 42 % los consideró una conducta muy grave, el 33 % los calificó como grave, el 18 % como moderada y el 7 % como poco grave. En conjunto, estos resultados evidencian una percepción predominantemente alta sobre la relevancia jurídica del fenómeno, ya que la mayoría de las personas participantes lo ubica en los niveles de mayor gravedad. Esto permite advertir que, en el contexto analizado, los crímenes de odio digitales no son vistos como expresiones menores o irrelevantes, sino como conductas que ameritan atención jurídica.



Nota. Fuente: elaboración propia con base en la sistematización de entrevistas semiestructuradas y revisión documental del estudio.

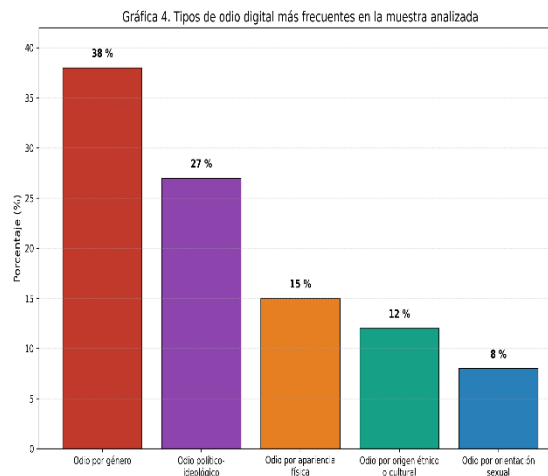
La gráfica 3 identifica los medios digitales más señalados para la realización de acciones de desacreditación en el contexto local analizado. Facebook concentra el mayor porcentaje (40 %), seguido de WhatsApp (25 %), X/Twitter (18 %), Instagram (10 %) y las páginas anónimas o blogs locales (7 %). Estos resultados muestran que las prácticas de desacreditación se ubican principalmente en

plataformas de uso cotidiano y de alta circulación, lo que favorece la rápida difusión de contenidos ofensivos, rumores o mensajes de hostigamiento. En conjunto, la información permite advertir que tanto las redes sociales abiertas como los servicios de mensajería y los espacios anónimos funcionan como canales relevantes para este tipo de agresiones.



Nota. Fuente: elaboración propia con base en la sistematización de entrevistas semiestructuradas y revisión documental del estudio.

La gráfica 4 muestra los tipos de odio digital más frecuentes en la muestra analizada. El odio por género ocupa el primer lugar (38 %), seguido del odio político-ideológico (27 %), el odio por apariencia física (15 %), el odio por origen étnico o cultural (12 %) y el odio por orientación sexual (8 %). En conjunto, estos resultados evidencian que las manifestaciones de odio digital identificadas por las personas participantes se concentran principalmente en categorías vinculadas con el género y la ideología, aunque también permanecen presentes expresiones asociadas a prejuicios sobre la apariencia, el origen y la orientación sexual. Esto permite advertir que el odio digital reproduce distintas formas de discriminación ya existentes en el ámbito social y las proyecta en los entornos digitales.





Nota. Fuente: elaboración propia con base en la sistematización de entrevistas semiestructuradas y revisión documental del estudio.

En conjunto, las gráficas muestran una concentración del fenómeno en plataformas de uso cotidiano y una presencia marcada de agresiones vinculadas con el género y la ideología. También permiten observar que la violencia digital adopta formas diversas, desde ataques directos hasta prácticas de desprestigio y hostigamiento reiterado.

5. DISCUSIÓN.

Los resultados obtenidos permiten sostener que el odio digital, en el contexto estudiado, no aparece como una anomalía comunicativa, sino como una forma reiterada de agresión que combina desprestigio, intimidación y exclusión. A partir de esa constatación, la discusión debe desplazarse del simple registro empírico hacia la pregunta por su sentido jurídico y social: qué tipo de violencia expresa, a quiénes afecta con mayor intensidad y hasta dónde puede llegar la respuesta penal sin sacrificar garantías fundamentales.

En primer lugar, los hallazgos dialogan con la noción clásica de crimen de odio desarrollada por Perry (2001), en la medida en que muestran que la agresión no se dirige solo contra una persona individual, sino también contra el grupo con el que se la asocia. Esta lectura coincide con las definiciones institucionales que distinguen entre delito motivado por prejuicio y discurso de odio, y subrayan que el sesgo discriminatorio es el elemento que dota al hecho de una lesividad específica (Consejo de Europa, s. f.). Además, las revisiones de Paz et al. (2020) e Izquierdo Montero et al. (2022) muestran que el hate speech sigue siendo un campo conceptual amplio y en disputa, lo que ayuda a explicar por qué su traducción jurídica no siempre resulta sencilla.

En segundo término, el estudio confirma que el daño digital no puede ser minimizado por el solo hecho de producirse en internet. Citron (2014) ha mostrado que la violencia en línea produce afectaciones acumulativas en la vida emocional, social y profesional de las víctimas, y Perry y Olsson (2009) advirtieron tempranamente que el entorno digital potencia la circulación del odio. En la misma dirección, las tendencias de investigación recientes muestran que el control del discurso de odio en redes sociales se ha convertido en un problema persistente en el debate académico y regulatorio (Sánchez Sánchez et al., 2023). Lo hallado en Chilpancingo se inscribe precisamente en esa discusión: el daño existe, se percibe y modifica la forma en que las personas participan en el espacio público digital.

Un tercer punto de discusión se relaciona con la desigual distribución de la exposición al daño. El predominio del odio por género y del odio político-ideológico en los hallazgos empíricos coincide con reportes que han documentado un impacto diferenciado sobre mujeres, periodistas y personas



con visibilidad pública (ARTICLE 19 México y Centroamérica, 2021; ONU Mujeres México, 2020). A ello se suma que la literatura reciente ha insistido en que la violencia digital y el sesgo de género representan un desafío central para construir entornos digitales inclusivos (Römer-Pieretti et al., 2025). Por ello, una lectura jurídicamente seria del problema no puede prescindir de la perspectiva de género ni del reconocimiento de desigualdades previas que se reactivan en línea.

En el plano jurídico, los resultados ponen de manifiesto una tensión central: la experiencia social del daño es más nítida que su inteligibilidad penal. Las personas participantes reconocen con relativa claridad cuándo una agresión en redes sociales rebasa el desacuerdo tolerable, pero al mismo tiempo expresan dudas sobre su encuadre jurídico, su denunciabilidad y la posibilidad real de sanción. Este punto se articula con lo señalado por la SCJN (2020), que admite restricciones a la libertad de expresión cuando están en juego la dignidad, la igualdad y la no discriminación, y con el estándar de la OACNUDH (s. f.), que exige valorar el contexto, la intención y el umbral de daño. El desafío, entonces, no es meramente nominal, sino interpretativo y probatorio.

Ahora bien, esa constatación no autoriza una expansión irreflexiva del derecho penal. El entorno digital plantea dificultades particulares para identificar autorías, reconstruir contextos comunicativos y distinguir entre insulto, hostigamiento, incitación y agresión motivada por prejuicio. En este punto conviene ser cautelosos. Heinze (2025) advierte que las estrategias automáticas de remoción del discurso de odio pueden tensionar seriamente la libertad de expresión, mientras que Anttila y Domínguez-Armas (2025) recuerdan que el problema también involucra formas de exclusión argumentativa del espacio deliberativo. Por eso, la respuesta penal solo puede ser legítima si se mantiene excepcional, taxativa y rigurosamente contextualizada.

El aporte principal del estudio radica en mostrar, desde una escala local, cómo se entrecruzan tres dimensiones que con frecuencia se analizan por separado: la experiencia concreta del daño, la estructura discriminatoria que lo sostiene y la dificultad jurídica para traducirlo en categorías operativas. Además, la evidencia local sugiere que las plataformas no son soportes neutros, pues sus dinámicas de interacción favorecen patrones específicos de propagación, reiteración y visibilización del odio, tal como lo muestran investigaciones recientes sobre comportamiento de propagadores en redes sociales (Perera et al., 2023). Incluso estudios centrados en población escolar han mostrado que las expresiones de odio en línea y fuera de línea no operan de forma aislada, sino conectada (Cáceres et al., 2025), lo que refuerza la necesidad de mirar el fenómeno más allá de la pantalla.

En términos de implicaciones, los hallazgos sugieren la necesidad de fortalecer criterios jurídicos para reconocer la motivación discriminatoria en entornos



digitales, mejorar las capacidades institucionales para documentar y valorar el contexto de las agresiones, y ofrecer rutas de denuncia más claras para las víctimas. Ello no implica adoptar una lógica puramente punitiva, sino construir respuestas proporcionales y coherentes con un enfoque de derechos humanos. En suma, el estudio muestra que el odio digital exige una respuesta jurídica más precisa y más situada, capaz de proteger eficazmente la dignidad humana sin diluir las garantías propias de un Estado constitucional.

6. CONCLUSIÓN.

El estudio permite concluir que los crímenes de odio digitales constituyen una manifestación contemporánea de violencias estructurales que encuentran en las redes sociales un medio especialmente eficaz para circular, amplificarse y permanecer. Su relevancia jurídica no proviene únicamente del lenguaje agresivo que emplean, sino del componente discriminatorio que los atraviesa y de la capacidad que tienen para producir silenciamiento, exclusión e intimidación pública.

Desde la perspectiva del derecho penal, el hallazgo más significativo no es solo la existencia de estas agresiones, sino la brecha entre la experiencia social del daño y su reconocimiento jurídico. El trabajo muestra que las personas participantes identifican con relativa claridad la gravedad del fenómeno, pero siguen enfrentando incertidumbre sobre los criterios para denunciarlo, investigarlo y sancionarlo. Esa distancia revela que el problema no se agota en la ausencia de figuras específicas, sino también en las dificultades interpretativas, probatorias e institucionales que plantea el entorno digital.

En ese sentido, la conclusión central del artículo es que una eventual respuesta penal frente al odio digital debe construirse desde un estándar estricto de legalidad, taxatividad, proporcionalidad y respeto a la libertad de expresión. No se trata de ampliar sin medida el castigo, sino de contar con criterios jurídicos suficientemente precisos para distinguir cuándo una expresión ofensiva permanece en el ámbito del discurso tolerable y cuándo, por su contexto, contenido, finalidad y efectos, se convierte en una agresión motivada por prejuicio jurídicamente relevante.

El estudio también confirma que el análisis del fenómeno exige una mirada situada. En un contexto local como Chilpancingo, donde las relaciones sociales, la reputación y la circulación de mensajes se entrelazan de manera particularmente intensa, el impacto de una agresión digital puede extenderse con rapidez al ámbito emocional, académico, profesional y comunitario de la víctima. Por ello, las respuestas institucionales no deberían construirse únicamente desde abstracciones normativas, sino a partir de la manera concreta en que estas violencias operan en cada territorio.



Finalmente, la investigación aporta una base para seguir profundizando en la relación entre violencia digital, discurso de odio y derecho penal en México. Su alcance es necesariamente local y cualitativo, por lo que no busca generalizaciones estadísticas; sin embargo, sí ofrece elementos relevantes para futuras investigaciones, para la discusión doctrinal y para el diseño de criterios institucionales más claros. Proteger la dignidad humana en los espacios digitales exige hoy una comprensión jurídica más fina del odio en línea y una respuesta estatal que sea, al mismo tiempo, efectiva, prudente y constitucionalmente fundada.

REFERENCIAS BIBLIOGRÁFICAS.

- Anttila, S., & Domínguez-Armas, Á. (2025). Argumentative exclusion and the case of online hate speech. *Topoi*, 44, 695-705.
<https://doi.org/10.1007/s11245-025-10165-9>
- ARTICLE 19 México y Centroamérica. (2021, 11 de junio). Machismo y censura: mujeres periodistas enfrentan violencia diferenciada en redes sociales. ARTICLE 19. Disponible en: <https://articulo19.org/machismo-y-censura-mujeres-periodistas-enfrentan-violencia-diferenciada-en-redes-sociales/>
- Cáceres, J. M. B., Martínez, P. R., Espinosa, J. J. S., & Joya, L. M. (2025). The perpetration of on- and offline hate speech among secondary school students. *Frontiers in Human Dynamics*, 7, 1632091.
<https://doi.org/10.3389/fhumd.2025.1632091>
- Citron, D. K. (2014). *Hate crimes in cyberspace*. Harvard University Press.
Disponible en: <https://www.hup.harvard.edu/books/9780674659902>
- Consejo de Europa. (s. f.). Hate crime and hate speech [thematic factsheet]. Council of Europe. Disponible en: <https://www.coe.int/en/web/portal/-/new-thematic-factsheet-on-hate-crime-and-hate-speech>
- Flick, U. (2015). *Introducción a la investigación cualitativa* (5.ª ed.). Morata.
https://edmorata.es/wp-content/uploads/2021/02/Flick.-Introduccion-a-la-investigacion-cualitativa_prw.pdf
- Heinze, E. (2025). Commentary to: AI-based removal of hate speech from digital social networks: chances and risks for freedom of expression. *AI and Ethics*, 5, 2957-2959. <https://doi.org/10.1007/s43681-024-00646-9>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación* (6.ª ed.). McGraw-Hill.
<https://apiperiodico.jalisco.gob.mx/api/sites/periodicooficial.jalisco.go>



[b.mex/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf](https://www.b.mex/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf)

Instituto Nacional de Estadística y Geografía [INEGI]. (2024). Módulo sobre Ciberacoso (MOCIBA) 2023. INEGI. Disponible en: <https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2024/MOCIBA/MOCIBA2023.pdf>

Izquierdo Montero, A., Laforgue-Bullido, N., & Abril-Hervás, D. (2022). Hate speech: a systematic review of scientific production and educational considerations. *Revista Fuentes*, 24(2), 222-233. <https://doi.org/10.12795/revistafuentes.2022.20240>

Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos [OACNUDH]. (s. f.). Freedom of expression and incitement to hatred. Naciones Unidas. Disponible en: <https://www.ohchr.org/en/freedom-of-expression>

ONU Mujeres México. (2020). Violencia contra mujeres y niñas en el espacio digital. ONU Mujeres. Disponible en: <https://mexico.unwomen.org/es/digiteca/publicaciones/2020-nuevo/diciembre-2020/violencia-digital>

Paz, M. A., Montero-Díaz, J., & Moreno-Delgado, A. (2020). Hate speech: A systematized review. *SAGE Open*, 10(4). <https://doi.org/10.1177/2158244020973022>

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>

Perry, B. (2001). In the name of hate: Understanding hate crimes. Routledge. Disponible en: <https://www.routledge.com/In-the-Name-of-Hate-Understanding-Hate-Crimes/Perry/p/book/9780415927734>

Perry, B., & Olsson, P. (2009). Cyberhate: The globalization of hate. *Information & Communications Technology Law*, 18(2), 185-199. <https://doi.org/10.1080/13600830902814984>

Reuters. (2024, 25 de enero). Foreign-born Miss Japan sparks debate on what it means to be Japanese. Reuters. Disponible en: <https://www.reuters.com/world/asia-pacific/foreign-born-miss-japan-sparks-debate-what-it-means-be-japanese-2024-01-25/>

Römer-Pieretti, M., Esteban-Ramiro, B., & Canelón Silva, A. (2025). Violence, hate speech, and gender bias: Challenges to an inclusive digital environment. *Social Inclusion*, 13. <https://doi.org/10.17645/si.9941>



Sánchez Sánchez, A. M., Ruiz-Muñoz, D., & Sánchez Sánchez, F. J. (2023). Research trends in the control of hate speech on social media (2016-2022). Cuadernos.info, 56, 89-116.
<https://doi.org/10.7764/cdi.56.60093>

Suprema Corte de Justicia de la Nación [SCJN]. (2020). Amparo Directo en Revisión 4865/2018 [resumen]. SCJN. Disponible en:
<https://www.scjn.gob.mx/derechos-humanos/sites/default/files/sentencias-emblematicas/resumen/2020-12/Resumen%20ADR%204865-2018%20DGDH.pdf>

Conflicto de Intereses: Los autores afirman que no existen conflictos de intereses en este estudio y que se han seguido éticamente los procesos establecidos por esta revista. Además, aseguran que este trabajo no ha sido publicado parcial ni totalmente en ninguna otra revista.

Financiación: Los autores declaran que este estudio no recibió ningún tipo de financiación externa por parte de agencias públicas, privadas, ni de organizaciones sin ánimo de lucro. Todas las actividades de investigación, análisis y desarrollo fueron realizadas con recursos propios.

CONTRIBUCIÓN DE AUTORÍA:

Autor: Olivia Elizabeth Álvarez Montalván(OEAM), Perla Elizabeth Ventura Ramos(PEVR)

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