

Discursos de odio e inteligencia artificial en la era de la desinformación

El artículo examina cómo la desinformación en las redes sociales fomenta los discursos de odio en los jóvenes, incidiendo en la formación de prejuicios. Se subraya la relevancia de la alfabetización mediática temprana como elemento clave para prevenir la radicalización y defender los derechos humanos.

[ILUSTRACIÓN: SYLVERARTS / [ISTOCK](#)]

La creciente propagación de desinformación y contenidos maliciosos, reforzada por el uso de inteligencia artificial (IA), sitúa a los más jóvenes en el centro de la vulnerabilidad frente a discursos de odio y estereotipos discriminatorios. En este escenario, vemos como la cultura “figital” —esa convergencia de lo físico y lo digital— fusiona ambas realidades en una experiencia única e ininterrumpida. Desde edades cada vez más tempranas, las nuevas generaciones acceden a dispositivos móviles y a redes sociales, determinando su visión del mundo.

El concepto de cultura figital nos presenta un nuevo escenario en el que “las realidades sociales espaciales y temporalmente específicas ya no se limitan al mundo físico; los procesos de desplazamiento por el espacio y establecimiento de puntos en común pueden tener lugar ahora tanto online como offline” (Boellstorff, 2012: 52).

Esta integración progresiva de lo físico y lo digital, percibido como un solo entorno conectado, se hace especialmente patente entre la población joven, cuyas actividades cotidianas —estudio, ocio, socialización— se desarrollan de forma simultánea en ambos planos. El informe “Mi realidad conectada: Sí, también soy digital” de la Fundación Cibervoluntarios (2023), aporta datos relevantes sobre el acceso a la tecnología en España:

- El 57,5% de los y las jóvenes recibe su primer dispositivo entre los 10 y 12 años.
- El 32,6% lo obtiene entre los 14 y 16 años.
- Más del 90% de la juventud accede a su primer móvil antes de los 16 años.

Estos datos coinciden con las tendencias señaladas por el informe Common Sense Media (2019), que sitúa en Estados Unidos la edad media de inicio del uso del primer teléfono inteligente entre los 10 y 12 años. Paralelamente, el acceso a redes sociales suele situarse alrededor de los 13 años, impulsado en gran parte por la curiosidad y el deseo de integrarse en la vida social de sus semejantes. No obstante, la legislación de muchos países establece los 13 o 14 años como edad mínima para el uso de redes y, aun así, la realidad muestra que los menores a menudo comienzan a interactuar en plataformas digitales por debajo de ese umbral.

En el marco de la cultura figital, las fronteras entre lo digital y lo presencial se difuminan, lo que permite que información falsa, prejuicios y discursos de odio penetren la vida cotidiana de los jóvenes

La ausencia de una educación mediática temprana incrementa el riesgo de que los menores acepten narrativas manipuladas como verdades incuestionables. El intercambio de memes y contenido visual en redes puede banalizar un contenido discriminatorio sin la conciencia necesaria para procesar el impacto de este tipo de mensajes. El informe Common Sense Media, (2019) señala cómo, el acceso de los niños a medios digitales de forma independiente, desempeña un papel destacado que perfila “su sentido de identidad, su conciencia

social, su bienestar y su acceso a la información, desde la política hasta el último meme viral”.

Desinformación y discursos de odio amplificadas por la IA

La desinformación consiste en la divulgación, con o sin intencionalidad, de contenidos falsos o engañosos que alteran la comprensión de la realidad (Wardle & Derakhshan, 2017). En la era de la inmediatez digital, las fake news se comparten a enorme velocidad, alimentando percepciones sesgadas y comportamientos extremistas. Los discursos de odio —definidos por el Consejo de Europa (2019) como la incitación o justificación de la violencia y la discriminación contra colectivos— hallan en las redes sociales un vehículo para su amplificación, especialmente ante audiencias con escasas herramientas críticas, como niños y adolescentes.

La investigación “Romper cadenas de odio, tejer redes de apoyo: los y las jóvenes ante los discursos del odio” realizada en el marco del proyecto [No More Haters](#), señala cómo “los jóvenes afirman enfrentarse en redes a mensajes de odio con gran carga ideológica y política sin la preparación necesaria para asimilarlos”.

Este artículo parte de la constatación del espacio digital como un entorno que puede favorecer la difusión de estos mensajes, a los que la población más joven está más expuesta por el uso de las redes sociales e internet: Un 38,1% (de los jóvenes consultados) señaló haber visto, en el último año, “páginas donde la gente publica mensajes que atacan a ciertos individuos o colectivos”.

En este sentido, el estudio “Discurso de odio sexista en redes sociales y entornos digitales” (Sigma Dos, 2021) establece cuatro tipos de “facilitadores” (Sigma Dos, 2021) o mecanismos que fomentan los discursos de odio entre los que se encuentran: el anonimato y la distancia que favorecen la emisión de mensajes hostiles, la sensación de impunidad con esa idea de ausencia de consecuencias frente a determinados comportamientos en redes y las cámaras de eco y los algoritmos que hacen que los adolescentes se rodeen de opiniones afines, reforzando sesgos y creencias extremas.

Estos riesgos se amplifican con la llegada de sistemas de inteligencia artificial capaces de generar y difundir información falsa de manera masiva. Los algoritmos de recomendación crean “burbujas de filtro” (Pariser, 2011), en la que los ciudadanos quedan aislados de otras opiniones, reforzando una visión del mundo menos diversa.

La IA irrumpe en este escenario como un amplificador de la desinformación al elaborar vídeos e imágenes sintéticas (*deepfakes*) que pueden dañar la reputación de individuos o colectivos, extendiendo estereotipos racistas y bulos difíciles de refutar (Chesney & Citron, 2019). Woolley, S. C., & Howard, P. N. (2018) señalan cómo este tipo de tecnología, hace que la desinformación sea mucho “más difícil de rastrear”.

El impacto de la desinformación en la construcción del discurso de odio

La desinformación que presenta a ciertos grupos raciales como amenazas a la identidad nacional y a la seguridad ciudadana alimenta narrativas radicales estableciendo “una voluntad de supremacía ideológica, y un riesgo para las democracias liberales” (Del-Fresno García, 2019).

Los titulares sensacionalistas y emocionales se viralizan rápidamente sin la debida verificación.

La falta de criterios sólidos de los menores para evaluar la veracidad de los hechos incrementa el riesgo de interiorizar estereotipos

En estos mensajes también se emplea el humor y memes para legitimar el odio. Las dinámicas de meme o el humor sarcástico pueden enmascarar expresiones xenófobas, presentándolas como algo inofensivo. Los movimientos anti-derechos instrumentalizan el humor “hacia abajo” —disfrazado de rebeldía— para difundir

discursos de odio no punibles (Izquierdo-Montero et al., 2022). Shifman (2014) señala que los memes, como “unidades de contenido digital” replicables y fáciles de transformar, tienen un gran potencial para normalizar estereotipos. A su vez, autores como Sunstein (2017) advierten que esta manipulación se aprovecha de defectos cognitivos para instalar prejuicios sin que el receptor sea plenamente consciente.

La alfabetización mediática temprana como respuesta urgente

Los diversos estudios citados ponen de manifiesto la necesidad de reforzar la alfabetización mediática e informacional desde edades tempranas, dotando a los más jóvenes de competencias digitales que les permitan detectar bulos, cuestionar memes discriminatorios y reflexionar sobre cómo quieren que sea su participación en las redes sociales:

1. Formación en pensamiento crítico: para analizar informaciones virales y verificar su autenticidad.
2. Conciencia de los algoritmos: comprender los procesos de personalización que encierran al usuario en burbujas de filtro (Pariser, 2011).
3. Responsabilidad en el uso del humor: los memes y el sarcasmo pueden enmascarar discursos de odio que se legitiman como “bromas”.
4. Protección de la infancia y la adolescencia: las instituciones, las familias y las propias redes deben propiciar un clima de respeto a la diversidad.
5. Estudios como el de la Fundación Cibervoluntarios (2023), indican que el acceso al primer móvil se produce antes de los 16 años. Por ello, es esencial contar con herramientas educativas que contrarresten la desinformación migratoria difundida a través de memes, vídeos cortos y narrativas emocionales en las redes (Blanco-Herrero y Arcila-Calderón, 2023).

El aprendizaje mediático y digital temprano es clave para repensar la cultura digital y detener la propagación de mensajes discriminatorios potenciados por la IA

La colaboración entre educación, investigación y ética tecnológica, junto con la alfabetización sobre bulos migratorios impulsada por las instituciones públicas (Molina Cañabate y Magallón Rosa, 2019), se revela como el mejor de los antídotos contra la desinformación y los discursos de odio, garantizando así sociedades más inclusivas y democráticas.

Eva Herrero destaca la vulnerabilidad de los niños ante la desinformación y los discursos de odio en redes. Por ello, es crucial incluir la alfabetización mediática en el currículo escolar para que aprendan a identificar información falsa (Herrero-Curiel, 2023). Este esfuerzo debe ser colaborativo, involucrando a docentes, familias, empresas tecnológicas, gobiernos y el tercer sector para fomentar un pensamiento crítico, esencial en la actualidad.

Blanco-Herrero, D., & Arcila-Calderón, C. “Deontología y noticias falsas: Estudio de las percepciones de periodistas digitales y tradicionales en España”. En R. Magallón (Ed.), *Comprometidos con la verdad. Propuestas para combatir la desinformación*. Madrid. Editorial CLABE (pp. 131-139).

Boellstorff, T. (2012). *Coming of age in second life: An anthropologist explores the virtually human*. Estados Unidos. Princeton University Press.

Cibervoluntarios. *Mi realidad conectada: Sí, también soy digital*. 2023. Disponible en: <https://www.cibervoluntarios.org/es/actualidad/post/cibervoluntarios-presenta-el-estudio-de-investigacion-mi-realidad-conectada-si-tambien-soy-digital>

Common Sense Media. *The Common Sense Census: Media Use by Tweens and Teens*, 2019. Disponible en: <https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2019>

Council of Europe. *Recommendation of the Committee of Ministers to member States on preventing and combating sexism*, 2019. Disponible en: <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>

- Chesney, R., & Citron, D. K.** (2019): "Deep fakes: A looming challenge for privacy, democracy, and national security" en *California Law Review*. Disponible en: <https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security>
- Del-Fresno-García, M.** "Desórdenes informativos: Sobreexpuestos e infrainformados en la era de la posverdad". *El Profesional de la Información*, 2019. 28(3), e280302. Disponible en: <https://doi.org/10.3145/epi.2019.may.02>
- Herrero-Curiel, E., & Arcila-Calderón, C.** (2023). Luchar contra la desinformación desde la infancia. El papel de la AMI en la escuela. En R. Magallón (Ed.), *Comprometidos con la verdad. Propuestas para combatir la desinformación*. Madrid. Editorial CLABE. (pp. 211-220).
- Izquierdo-Montero, A., Laforgue-Bullido, N., Quirós-Guindal, A. y Lorón-Díaz, I.** Adolescentes frente a los discursos de odio. Una investigación participativa para identificar escenarios, agentes y estrategias para afrontarlos. Madrid, *Centro Reina Sofía Sobre Adolescencia y Juventud, Fad*, 2022. Disponible en: <https://zenodo.org/records/6581238>
- Pariser, E.** (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Estados Unidos. Penguin Books.
- Molina Cañabate, J. P., & Magallón Rosa, R.** "Procedimientos para verificar y desmontar informaciones falsas basadas en el discurso del odio: El caso de Maldita migración" en *Revista de la Asociación Española de Investigación de la Comunicación*, (2019, 6(12), 89-104).
- Sigma Dos.** (2021). *Discursos de odio sexistas en redes sociales y entornos digitales*. Madrid: Centro Reina Sofía Sobre Adolescencia y Juventud, Fad.
- Shifman, L.** (2013). *Memes in digital culture*. The MIT Press. Disponible en: <https://doi.org/10.7551/mitpress/9429.001.0001>
- Sunstein, C.** (2017). *#Republic: Divided Democracy in the Age of Social Media*. Estados Unidos. Princeton University Press.
- Wardle, C., & Derakhshan, H.** (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Council of Europe.
- Woolley, S. C., & Howard, P. N.** *Computational Propaganda Worldwide: Executive Summary*. Oxford Internet Institute, 2018. Disponible en: <https://demtech.oii.ox.ac.uk/research/posts/computational-propaganda-worldwide-executive-summary>