

COLECCIÓN TEMAS EN DERECHOS HUMANOS

TOMO V

**INTELIGENCIA ARTIFICIAL
Y DERECHOS HUMANOS**

JORNADAS



Centro
Bajo los auspicios
de UNESCO



Centro Internacional
para la Promoción de
los Derechos Humanos

Directora Ejecutiva: Mg. Fernanda Gil Lozano
Director de Desarrollo de las Capacidades Institucionales: Dr. Rodrigo Karasik (a cargo)
Directora de Estudios e Investigación: Lic. Carolina Tellería (a cargo)
Director Técnico Administrativo: CPN Guido Napolitano
Coordinadora de Asuntos Jurídicos: Dra. Gabriela Pantuso
Responsable General de Gestión: María Tatiana Souza Korolkov
Responsable de Comunicación Institucional: Lic. Diana Martí
Responsable de Comunicación de Proyectos: Lucero Vega
Coordinación Proyecto Inteligencia Artificial: Lic. Elena Hanono
Compiladoras: Lic. Elena Hanono, María Tatiana Souza Korolkov, Dra. Gabriela Pantuso
Diseño Editorial: Lic. Marcela Braccelarghe
Corrección y Edición Gráfica: Anna Souza
Asistente de Diseño: Mercedes Araujo
Coordinadora Colección: Mg. Marina Guimpel

Publicado por:

CENTRO INTERNACIONAL PARA LA PROMOCIÓN DE LOS DERECHOS HUMANOS bajo los auspicios de UNESCO

Av. España 2591, Buenos Aires, Argentina

Tel: +5411 5300 4000 extensión 78300

Las ideas y opiniones expresadas en esta obra son las de las y los autores; no reflejan necesariamente el punto de vista del CIPDH ni comprometen a la organización. Los gráficos expuestos han sido elaborados en base a la información proporcionada por cada autor

Cómo citar esta publicación: Centro Internacional para la Promoción de los Derechos Humanos bajo los auspicios de UNESCO (CIPDH). (2024). "Colección Temas en Derechos Humanos. Tomo V – Inteligencia artificial y derechos humanos", Buenos Aires, Argentina.

ISBN 978-631-90305-9-4

ISBN 978-631-90305-9-4



9 786319 030594

ÍNDICE

Prólogo	5
Introducción.....	7
Introducción a la inteligencia artificial y derechos humanos María Dolores Souza (Chile)	13
Cómo entiendo y cómo me entiende la inteligencia artificial Sebastián Waisbrot (Argentina)	29
Desórdenes informativos y narrativas hostiles María José Labrador (México)	41
Transformación digital y derechos humanos. Un diálogo a construir Gabriel Levy (Colombia)	55
Agencia humana a prueba de futuro. Cómo proteger los derechos humanos frente a las máquinas Cordel Green (Jamaica)	69
Inteligencia artificial para proteger a las personas en estado de vulnerabilidad Samridhi Arora (India)	79
Democratización del conocimiento a través de la tecnología Matías Karlsson (Argentina)	91
Educación e inteligencia artificial: un caso argentino Alicia Penayo (Argentina)	101
Ética de la inteligencia artificial: equidad de género Natalia González Alarcón (Colombia)	111

DESÓRDENES INFORMATIVOS Y NARRATIVAS HOSTILES

MARÍA JOSÉ LABRADOR (MÉXICO)

Doctora en Ciencias de la Información por la Universidad Complutense de Madrid. Licenciada en Ciencias y Técnicas de la Comunicación por la Universidad Anáhuac México Norte. Directora de Investigación en la Universidad del Desarrollo de Chile. Miembro del Grupo AI4AI¹ del proyecto IFAP² UNESCO. Conferencista de trayectoria internacional.

RESUMEN

Esta presentación se estructura en tres grandes ejes. El primero aborda el fenómeno de la desinformación, para lo cual es prioritario hacer énfasis sobre su concepto y significación, así como su clasificación y caracterización. En el segundo eje se analizan las vulnerabilidades y las perspectivas éticas frente a la inteligencia artificial en este mismo marco de la desinformación. Para concluir, se efectúa una breve reseña sobre las investigaciones y estudios que se están realizando en el Observatorio Desinformación Chile, del cual la autora es fundadora.

Palabras clave

inteligencia artificial, desinformación, clickbait, vulnerabilidades, desiertos informativos, troll, perspectivas éticas, bot

ABSTRACT

This presentation is structured around three main axes. The first addresses the phenomenon of disinformation, emphasizing its concept and significance, as well as its classification and characterization. The second axis analyzes vulnerabilities and ethical perspectives regarding artificial intelligence within the same framework of disinformation. To conclude, a brief overview is provided of the research and studies being conducted at the Disinformation Observatory in Chile, of which the author is a founder.

Keywords

artificial intelligence, disinformation, clickbait, vulnerabilities, information deserts, troll, ethical perspectives, bot

El fenómeno de la desinformación, clasificación y caracterización

En un contexto de sobreabundancia informativa derivada de la democratización y el amplio acceso a internet (Vizoso y Vázquez-Herrero, 2019) surge la desinformación como uno de los fenómenos globales de mayor trascendencia y todavía impredecible (Wardle y Derakhshan, 2017), expresado mediante desórdenes informativos que inciden en todo tipo de procesos sociales y democráticos.

Este cambio de paradigma ha traído un aumento muy grande en la producción de la circulación de información y, junto con ello, una desinformación sin precedentes cuyas características están ligadas a sesgos cognitivos acoplados y a la llamada economía de la atención.

La desinformación es una de las principales amenazas a las sociedades libres y a las democracias latinoamericanas y del mundo. Así lo han hecho ver instituciones globales como el Fondo Económico Mundial desde 2013, la UNESCO desde 2020 a la fecha y la Unión Europea con su plan de acción 2018-2024 que finaliza con las leyes de Servicios Digitales³ y la de Mercados Digitales⁴ (Unión Europea, 2018). Estas tienen por objeto crear un espacio digital más seguro en el que se protejan los derechos fundamentales de los usuarios y establecer condiciones de competencia equitativas para las empresas.

El plan de acción de lucha contra la desinformación, de la Unión Europea (2018), y a la vez la Comisión Europea, a partir del informe de resultados de un enfoque multidisciplinario a la desinformación, elaboran una distinción basada en desinformación (*disinformation*), información errónea (*misinformation*) e información maliciosa (*malinformation*) (Wardle & Derakhshan, 2017).

Los desórdenes informativos incluirían formatos, productos informativos y presentación de la información caracterizada como sátira y parodia, conexión falsa, contenido engañoso, contenido impostor, contenido fabricado, contexto falso y contenido manipulado.



En el marco de los procesos electorales estos desórdenes informativos se han convertido en una amplia preocupación tanto en la región como en Chile.

En el análisis del impacto de lo que algunos han llamado la “crisis de la información” se plantea una idea de una “crisis epistémica” de la democracia (Bimber y Gil de Zúñiga, 2020), ya que la desinformación en contextos electorales supone una gran amenaza al sistema democrático. En este entorno, la cuestión central es cómo los ciudadanos están filtrando, verificando, discerniendo y circulando la información y, en definitiva, quién define lo verdadero y prioriza lo importante en el espacio público (Dahlgren, 2018) y de qué forma se generan capacidades en la ciudadanía para detectar la desinformación (UNESCO, 2023).

Ante esta necesidad surge el chequeo de datos como un movimiento transnacional (Graves, 2016), conformado por profesionales de diversos sistemas de medios y culturas periodísticas, quienes bajo procedimientos específicos (Moreno Gil y Salgado de Dios, 2023) y conscientes de lo perjudicial de los desórdenes informativos (Mena, 2018), se dan la tarea de verificar datos. Como explican López y Rodríguez (2019), el actual movimiento de plataformas de verificación no se concreta en departamentos asignados a una redacción que corrigen los errores antes de publicar, sino que su objetivo es recopilar afir-

maciones sobre hechos, encontrar pruebas que los confirmen o cuestionen e informar sobre las conclusiones, desacreditando así contenidos viralizados en redes sociales, que circulan con rapidez y en cascadas informativas.

¿Cuál es la categorización que tiene la desinformación?

Existen diferentes categorías de la desinformación. Una de ellas es el *clickbait*, que es un cebo, una técnica de redacción que consiste en crear encabezados y descripciones sensacionalistas para llevar a la opinión pública a determinados tipos de enlace. En los años pasados, junto con ellos se han generado y se han abierto camino estas granjas de *clickbait*. Es decir, una cantidad enorme de *clickbait* con una funcionalidad que en Facebook era Instant Articles⁵, y que permite consumir mucho contenido periodístico.

También Ad Breaks, a partir de sus pausas publicitarias, que permiten además monetizar el video, ha fomentado una cantidad enorme de estas granjas. En un momento dado, Allen y otros investigadores del MIT⁶ registraron que hasta un 60 % de estos dominios o estas tácticas se estaban utilizando en Facebook. De hecho, frente al caso Cambridge Analítica⁷, Facebook tiene que comparecer ante el Congreso de Estados Unidos y, a partir de ello, cambia su nombre a Meta. Cambia también entonces la lógica de su dueño, Mark Zuckerberg, quien en el lanzamiento de la red declara que la privacidad había terminado y, a partir de estos hechos, Meta explica que la privacidad es un tema fundamental para la empresa, lo que admite un análisis en materia de investigación.

Otra categorización son los bots, que son cuentas automatizadas de redes sociales administradas por un algoritmo y no por una persona real. Esto se ve habitualmente sobre todo en el anonimato, que es uno de los problemas más graves de la desinformación, por el tema de la opacidad. Es decir, no se puede saber con claridad el origen de la desinformación.

Los trolls son usuarios reales que amplifican algunos mensajes con bots, y han irrumpido con fuerza los llamados *deepfake*, videos que en la actualidad tienen un análisis y una investigación muy profundos porque son un tipo particular

de medios sintéticos. Es decir, una persona, una imagen, un video existente es reemplazado por la imagen de otra persona. Se utilizan técnicas muy poderosas de aprendizaje automático y de inteligencia artificial y se manipula y se genera el contenido visual y el contenido de audio. Esto tiene un alto potencial de engaño y está siendo material de análisis para los investigadores, precisamente por los graves riesgos a los que conlleva.

Finalmente, el Brain Hacking o las teorías conspirativas, que tienen un efecto exponencial en los tipos de desinformación y en las categorías y los discursos de odio o las narrativas hostiles, las encontramos vinculadas, generalmente, a grupos con problemáticas específicas, por ejemplo, migrantes o grupos de enfermos o con algún tipo de discapacidad.

Vulnerabilidades y perspectivas éticas frente a la IA en el marco de la desinformación y las narrativas hostiles

El segundo eje de análisis de esta presentación tiene que ver con las vulnerabilidades y las perspectivas éticas frente a la inteligencia artificial en el marco de la desinformación.

¿CUÁLES SON ESTAS VULNERABILIDADES?

Por ejemplo, en Chile existen grupos de mujeres, grupos de adultos mayores y personas con necesidades, que son vulnerables, a los que se presta especial atención.

Un aspecto que incide en tales vulnerabilidades tiene que ver con los llamados desiertos informativos, con las brechas de género y los paradigmas de liderazgo y, junto con ello, se exponen algunas perspectivas éticas al respecto. Un reciente estudio que hizo la Asociación de Mujeres Geógrafas en Chile da cuenta de que cuatro de cada diez mujeres de Chile viven en zonas de nivel alto o muy alto de vulnerabilidad.

En regiones como Santiago, Antofagasta, Valparaíso y Concepción hay un alto nivel de vulnerabilidad de las mujeres. En 2021 la Subsecretaría de Telecomunicaciones de Chile (SUBTEL) refería que el 77 % de las mujeres presentan mayor dificultad y vulnerabilidad en el momento de acceder a servicios digitales.

Es decir, existe un problema de raíz que tiene que ver primero con la brecha digital y, por tanto, aumenta esta vulnerabilidad respecto de los avances relacionados con inteligencia artificial o la sociedad tecnológica.

Los desiertos informativos representan un problema que se ha expuesto a nivel regional y que por tanto plantean múltiples desafíos. El primero sería que millones de latinoamericanos viven en zonas rurales, en zonas urbanas pobres, con lo cual habitan auténticos desiertos de noticias. Es decir, no existen ecosistemas mediáticos robustos y, por tanto, enfrentamos un problema anterior: no tener información, y que la información que llega está distorsionada o manipulada. Además, hay problemas posteriores que tienen que ver con el pluralismo informativo. Son comunidades que carecen de estos medios locales y que proporcionan, además, contenidos creíbles y completos. Es decir, muchas veces tienen algo de las noticias, pero con precarización.

Existe una precarización del periodismo y de otras fuentes locales, y estas comunidades no tienen acceso a información esencial, ni siquiera sobre eventos locales, y mucho menos de gobierno o de elecciones y otros temas importantes que, particularmente en este 2024, son de mucha relevancia.

La Inter Press Service⁸ y otras organizaciones, así como el *Time Journalist Review*, han dado cuenta del caso interesantísimo de Brasil, que tiene prácticamente un 48,7 % de municipios o de comunidades que carecen de medios de comunicación. Argentina bordea también el 47 %. El caso Colombia, todos ellos dignos de estudio y a pesar de que surgen cada año centenares de medios en línea, muchos no son independientes o están sesgados porque dependen, muchas veces, de autoridades locales, de alcaldías u otro tipo de organizaciones. En Venezuela siete millones de personas viven en municipios donde no hay medios de comunicación, y la cifra sube prácticamente al doble si se consideran como semidesiertos los municipios que tienen uno o dos medios de comunicación. Por lo tanto, los desafíos están en la sostenibilidad de estos medios de comunicación y en la viabilidad como empresas digitales.

Este año se acaban de ejecutar iniciativas del Consejo de la Unión Europea como “Medios locales para la democracia”⁹. Tam-

bién un estudio da cuenta de la importancia de la movilización de los agentes públicos y privados para contribuir a reactivar el panorama de los medios de comunicación locales, y la importancia de reforzar la independencia, la sostenibilidad y la capacidad de innovación de los medios de comunicación.

A la vez, desde las perspectivas de la educación es fundamental concretar planes de acción para la alfabetización digital. En esta línea se está trabajando con diversos investigadores en el ámbito nacional e internacional.

Finalmente, ¿Cuáles son las perspectivas éticas? ¿Cuáles son los principios éticos o la ética que habremos de poner en marcha frente a la inteligencia artificial?

Uno de los grandes desafíos éticos que se presentan al respecto se propone a partir de un análisis sobre las características fundamentales que se extraen de la teoría de la ética del cuidado, de Carol Gilligan: cuidado, responsabilidad, comunidad, atención e interdependencia. La fuerza que las mueve es la cooperación mutua que se lleva a cabo por medio de aptitudes como la empatía y la capacidad para sostener relaciones interpersonales. Esto está planteado así, a diferencia de la ética de justicia, cuyos valores iban más encaminados hacia la individualidad, la independencia, la objetividad, la libertad, la igualdad y la justicia, y son impulsadas por la razón y el cumplimiento de las normas estipuladas. Las mujeres, según Gilligan, se centran menos en la justicia abstracta y la equidad, y piensan más en las responsabilidades que tienen hacia personas específicas de su entorno. En su modelo, Gilligan postula que las preferencias de las mujeres hacia el cuidado de los demás como responsabilidad moral se fundamenta en el hecho de ser responsables, en mayor medida, del cuidado de los recién nacidos. Las mujeres piensan más en términos relacionados con el cuidado, en un plano contextual, y por ello se requiere plantear los principales desafíos para la política social y de desarrollo de la sociedad tecnológica en esa dirección.

Por lo tanto, valores como equidad, justicia, intimidad y privacidad, desde sus abordajes, así como la privacidad del norteamericano y la intimidad en el enfoque europeo, deben de ser base inmutable para que las decisiones algorítmicas tengan

que pasar necesariamente por una supervisión humana de los sistemas de inteligencia artificial.

Reviste especial importancia aportar un enfoque similar en la región que se relaciona con el desarrollo ético de la inteligencia artificial. Existe la necesidad de avanzar y profundizar en garantizar este desarrollo ético. ¿Qué significa eso? ¿Qué significa la transparencia algorítmica? ¿Qué significa que los chatbots no produzcan contenidos ilegales? ¿Qué significa velar por derechos humanos tan significativos como la libertad de expresión y el derecho a la información?

Observatorio Desinformación Chile. Investigaciones y estudios

El Observatorio Desinformación Chile¹⁰ es una iniciativa académica que nace a partir del proyecto “Garantías institucionales y regulatorias, autoridades electorales y de supervisión digital ante interferencias y narrativas hostiles”, de publicidad segmentada y polarización, que llevan a cabo 69 investigadores de todo el mundo, con un fondo adjudicado por el Ministerio de Ciencia e Innovación de España a la Facultad de Ciencias de Información de la Universidad Complutense de Madrid.

El Observatorio Desinformación Chile está anclado al Observatorio de la Desinformación Complutense. Las investigaciones que se abordan se centran en la regulación de la desinformación y en estudios sobre la desinformación en redes sociales. Un aspecto central de Estudios sobre la Regulación de la Desinformación es la investigación de todas las regulaciones que aborden tanto los aspectos procedimentales e institucionales (quién debe conocer y resolver los conflictos) como la regulación de las plataformas tecnológicas. Se ha puesto especial atención en el estudio de las estructuras de red y el flujo informacional de hashtags desinformativos. Asimismo, se ha realizado seguimiento y monitoreo de cuentas que realizan desinformación durante los procesos electorales.

Chile focaliza en el proceso electoral del plebiscito constitucional de 2023, que fue interesantísimo para el análisis en la temática de las desinformaciones. Algunos resultados de interés son los siguientes: un 68,8 % de las desinformaciones provienen de la red social X, un 12,5 % de la TV y un 6,3 % de otras plata-

formas como TikTok, Instagram y Facebook. El estudio concluyó también que el autor causante de la desinformación, en un 12,5 % de los casos, es atribuible a partidos políticos, mientras que en un 87,5 % a un actor de origen desconocido. Respecto de la temática de las desinformaciones, un 58,8 % de ellas están vinculadas con deficiencias en el sistema electoral, seguidas de un 11,8 % sobre seguridad y otro 11,8 % sobre legislación y funciones del Congreso.

Cabe destacar que el análisis contempló los desórdenes informativos encontrados en los meses de noviembre y diciembre (antes de la fecha del plebiscito de 2023), para los cuales se aplicaron criterios de selección relacionados con el rango de interés público, el potencial de daño que generan, el nivel de viralización y las solicitudes de la audiencia. En total, se investigaron 17 desórdenes informativos.

Adicionalmente, el Observatorio elaboró un análisis comparado de la cobertura mediática de diversos procesos políticos y/o electorales tanto en medios tradicionales como independientes y extrasistémicos o los llamados pseudomedios. El objetivo es caracterizar con detalle los enfoques y encuadres noticiosos de cada uno de estos soportes, para determinar el grado de injerencia que tienen en la validación de información por parte de las audiencias y su posterior asimilación y difusión. Esto implica indagar tanto en el consumo informativo en diversos grupos etarios como en las rutinas y los métodos de los periodistas e informadores empíricos para concebir información, la responsabilidad profesional que emana de dicho proceso y los mecanismos de media accountability que utilizan los medios de comunicación en Chile.

NOTAS

¹ Artificial Intelligence for Information Accessibility.

² Programa Información para Todos (PIPT) o Information for All Programme (IFAP) es un programa intergubernamental único de la UNESCO que apoya a los Estados miembros en el fomento de sociedades equitativas mediante la promoción del acceso universal a la información y al conocimiento para el desarrollo sostenible. <https://www.unesco.org/en/ifap>

³ Esta ley entró en vigor el 16 de noviembre de 2022. Su objetivo es crear un entorno digital más seguro y abierto y proteger los derechos fundamentales de los usuarios en línea. La DSA establece normas para com-

batir la oferta de contenidos ilícitos en línea, empoderar a los usuarios y a la sociedad civil, evaluar y paliar los riesgos y supervisar y controlar el cumplimiento de las plataformas en línea.

⁴ Esta ley tiene como objetivo garantizar la equidad y la competencia en los mercados digitales. La DMA impone normas a las plataformas que actúan como “guardianes” en el sector digital. Su objetivo es proteger la privacidad de los usuarios y fomentar la competencia leal, la transparencia y la innovación en el mercado digital.

⁵ Facebook Instant Articles ofrece a los editores una manera rápida y atractiva de presentar sus contenidos a una gran audiencia.

⁶ Massachusetts Institute of Technology.

⁷ Facebook se halla sumida en una tormenta política global por la fuga de datos de 50 millones de usuarios estadounidenses, que habría utilizado la consultora Cambridge Analytics para afinar con perfiles psicológicos las estrategias de atracción de votos de la campaña de Donald Trump en 2016.

⁸ Inter Press Service (IPS) es una agencia mundial de noticias comprometida con el ejercicio del periodismo independiente y especializada en reportajes analíticos sobre procesos y acontecimientos económicos, políticos, sociales, artísticos y culturales. Su cobertura se caracteriza por incluir a las organizaciones que conforman la sociedad civil como fuente informativa, y los procesos de la era de la globalización como temas de especial interés.

⁹ El Consejo de la Unión Europea señala la importancia de la libertad de los medios de comunicación en la UE en un momento en el que la libertad de los medios de comunicación se está cuestionando en muchos países del mundo. Se han presentado nuevas propuestas con el fin de seguir reforzando la libertad de los medios de comunicación en la UE. <https://www.consilium.europa.eu/es/politicas/media-freedom-eu/>

¹⁰ <https://observatoriodesinformacion.cl/>

REFERENCIAS

Dahlgren, P. (2018). La participación en línea en la esfera pública. Las ambigüedades del afecto. *In Mediaciones de la Comunicación*, 13. <https://doi.org/10.18861/ic.2018.13.1.2824>

Graves, L. (2016). *Deciding what's True: The Rise of Political Fact-checking in American Journalism*. Columbia University Press.

López, M. y Rodríguez, J. (2019). El papel de las plataformas de verificación en la lucha contra la desinformación: un análisis de su impacto en redes sociales. *Revista internacional de Comunicación Digital* Nro. 23, 87-110.

- Mena, P. (2018). Los desórdenes informativos y su impacto en el proceso democrático: Herramientas de lucha contra la desinformación. *Revista de Estudios Políticos y Sociales*, 15(2), 45-65.
- Moreno Gil, V. y Salgado de Dios, F. (2023). El cumplimiento del código de principios de la International Fact-Checking Network en las plataformas de verificación españolas. Un análisis cualitativo. *Revista de Comunicación*, 293-307.
- UNESCO (2023). *The challenge of disinformation in electoral processes*. Obtenido de <https://www.unesco.org/es/articles/el-desafio-de-la-desinformacion-en-procesos-electorales>.
- Vizoso, A. y Vázquez-Herrero, J. (2019). Plataformas de fact-checking en español. Características, organización y método. *Communication & Society*.
- Wardle, C., y Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Council of Europe. <https://rm.coe.int/information-disorder-report-november-2017/1680764666>