

Image not found or type unknown



Núm 43 - NOVIEMBRE 2025

Paz en la era
digital

ICIP
✓

SUMARI

Introducción

- Paz en la Era Digital

Artículos centrales

- Desinformación, manipulación y discurso del odio en los conflictos
- Riesgos de la IA para una paz sostenible
- Algoritmos puente: herramientas para despolarizar
- La mediación en la era de los algoritmos: riesgos y oportunidades para los procesos de paz
- Democracia digital y ciberactivismo en África Oriental: el caso de Siasa Place
- Documentar la cultura antes, durante y después del conflicto

Entrevista

- Entrevista amb Stephanie Williams

INTRODUCCIÓN

Paz en la Era Digital

Helena Puig Larrauri

Cofundadora y directora de estrategia de Build Up

En 2022, #BookTok en TikTok impulsó las ventas literarias en Estados Unidos, el comercio local a través de WhatsApp expandió la presencia de pequeñas empresas por toda América Latina, y el uso de Facebook para llegar a los votantes hizo que aumentara la participación de las personas jóvenes y las que votaban por primera vez en las elecciones regionales de la India. Ese mismo año, las redes sociales amplificaron las tensiones étnicas durante las elecciones de Kenia, difundieron desinformación sobre la guerra en Ucrania e incitaron a la violencia durante las protestas en Irán. Al igual que en años anteriores y posteriores, las redes sociales moldearon la cultura, la política y los conflictos de múltiples formas, a menudo contradictorias. Pero 2022 también marcó un punto de inflexión a nivel global, el tiempo dedicado a las plataformas sociales comenzó a disminuir, especialmente entre los jóvenes. ¿Puede esto indicar que su influencia en la sociedad está empezando a menguar?

Para responder a esa pregunta es necesario comprender los mecanismos mediante los que las redes sociales influyen en las sociedades y, de mayor importancia para este monográfico, cómo repercuten en los conflictos. Las tendencias en las redes sociales son una manifestación del consenso alcanzado por un grupo concreto. En el ámbito de la construcción de paz, nos preocupa la manera en que la desinformación erosiona la verdad, daña las instituciones y conduce a la violencia. Sin embargo, resulta esencial profundizar un poco más en las causas de esta erosión de la verdad.

La verdad, en un sentido posmoderno, no es estática, sino dialógica o, en otras palabras, el reflejo de un consenso social que proporciona la estabilidad y la base común necesarias para el desacuerdo, hasta el momento en que este consenso se ve

cuestionado de forma constructiva. Las redes sociales han influido en la verdad y la estabilidad al minar el consenso social, y la característica clave de este consenso minado no es tanto la desinformación — ésta es una señal o, como mucho, una táctica — como la polarización afectiva. La polarización afectiva (o polarización tóxica) se distingue de la polarización basada en cuestiones o ideas y tiene que ver con situaciones en las que se cree en ciertas personas por lo que son, no por lo que dicen. Cuanto más poderosas son las tendencias de las redes sociales y los subgrupos digitales, mayor es la polarización afectiva y más se fragmenta el amplio consenso social sobre la verdad en múltiples consensos de subgrupos. No es tanto que la verdad se haya erosionado, sino que se ha fragmentado.

“ No es tanto que la verdad se haya erosionado, sino que se ha fragmentado ”

En este monográfico, Ahmad Qadi analiza en profundidad cómo esta fragmentación está afectando el conflicto, y examina cómo la «tríada de desinformación, discurso de odio y polarización» en las redes sociales refleja y alimenta activamente las tensiones, agudiza las divisiones y puede prolongar el conflicto. Él lo denomina — y yo estoy de acuerdo — «un componente esencial de la guerra de la información moderna». Describe cómo Israel ha librado esta guerra psicológica contra la población palestina, utilizando contenidos generados por inteligencia artificial y granjas de bots para amplificar el sentimiento antipalestino y promover las acciones militares israelíes en los espacios digitales occidentales. Asimismo, el artículo de Sanjana Hattotuwa estudia más a fondo el impacto de estas opiniones en línea manipuladas, en un contexto específico de importancia para la construcción de paz: las mesas de negociación a puerta cerrada y los procesos de mediación. Y la entrevista con Stephanie Williams nos habla de un caso parecido, al analizar cómo el discurso de odio en línea y la injerencia extranjera durante los esfuerzos de paz de la ONU en Libia entre 2019 y 2020 no solo hicieron que las divisiones sobre el terreno se hicieran más profundas, sino que también pusieron en peligro a las mujeres que trabajaban y negociaban por la paz (incluida la propia Williams).

Me encantaría pensar que el descenso en el uso de las redes sociales que comenzó en 2022 también puede significar un lento retorno a un mundo menos fragmentado y menos polarizado, en el que los conflictos sean menos propensos a escalar y descontrolarse; un mundo en el que tengamos que preocuparnos menos por los riesgos que Qadi y Hattotuwa exploran en sus artículos. Sin embargo, el auge de la inteligencia artificial (IA) en los últimos años no solo mantiene la tendencia iniciada por el uso de las redes sociales, sino que también la complica de formas que intensifican aún más los conflictos.

En su artículo sobre los riesgos de la IA para la paz sostenible, Evelyne Tauchnitz analiza esta cuestión desde la perspectiva de la paz relacional, es decir, cómo la IA disminuye la capacidad de las personas y sus redes (sociedades) para convertir los desacuerdos en discusiones constructivas en lugar de conflictos destructivos o violentos. Tauchnitz pone de relieve cómo la IA afecta a la dignidad —a través de la extracción y el colonialismo de datos, al explotar y responder a nuestros estados emocionales, mediante la vigilancia digital y al desgastar nuestra capacidad de acción práctica y moral— y cómo socava la confianza en las instituciones. En su opinión, los sistemas de IA erosionan las condiciones que permiten a las sociedades ser resilientes ante los conflictos.

“ Me encantaría pensar que el descenso en el uso de las redes sociales puede significar un lento retorno a un mundo menos fragmentado y menos polarizado, en el que los conflictos sean menos propensos a escalar y descontrolarse ”

Más allá de lo que explica Tauchnitz, hay dos riesgos concretos de la IA para la paz a los que vale la pena prestar especial atención. En primer lugar, la IA puede menoscabar la verdad entendida como consenso al empobrecer el periodismo y la investigación. Las máquinas no pueden proporcionar los matices y el instinto profundamente contextual y humano que ofrecen los periodistas y los investigadores, por ejemplo, en entrevistas a

testigos de los hechos.

En relación con esta cuestión, Wahbi Abdelrahman teje una hermosa y triste historia sobre cómo un archivo digital está evitando el genocidio cultural en la guerra de Sudán, y cómo la incorporación de habilidades de digitalización en las instituciones locales es una forma de resistencia a ser borrados del mapa. Su artículo destaca que el archivo digital es un proceso sociotécnico, una colaboración entre humanos y máquinas que la IA no puede sustituir. Irónicamente, esta disminución de la información humana amenaza el futuro desarrollo de la propia IA. Los grandes modelos de lenguaje actuales (LLM) se basan en enormes cantidades de datos, y existe cierta preocupación (en debate) de que si la IA agota el contenido de calidad generado por humanos que necesita para su entrenamiento —como el periodismo de calidad, el archivo y la investigación— puede llegar a un callejón sin salida en materia de datos.

En segundo lugar, la IA puede debilitar los procesos de diálogo, aunque también existen algunos casos excelentes de uso de la IA en favor de la deliberación. El artículo de Luke Thorburn explora cómo podemos diseñar algoritmos de clasificación —es decir, los programas informáticos que deciden qué contenido vemos primero o qué contenido vemos más, en cualquier medio digital —para fomentar la conexión. Por ejemplo, Thorburn explica que se pueden usar algoritmos para encontrar declaraciones en las que quienes de otro modo estarían en desacuerdo pueden ponerse de acuerdo, y relata la manera en que la Alianza para la Paz en Oriente Medio ha utilizado esta herramienta. Stephanie Williams describe cómo el uso de la IA para resumir las opiniones en los diálogos digitales sobre Libia organizados por la ONU ayudó a presionar a la clase política para que ésta reconociera la opinión popular sobre cuestiones difíciles. Al igual que en el caso de la investigación, estos casos de uso están integrados en procesos sociotécnicos que mantienen la conexión humana. Cuando la IA se utiliza para representar a grupos sin participantes humanos, erosiona los procesos humanos esenciales de escucha y diálogo, fundamentales para la democracia y la paz. La deliberación es primordial para el debate público y los procesos legislativos, ya que permite a los grupos cuestionar los prejuicios, exigir información y forjar compromisos que conduzcan a decisiones más defendibles y duraderas. La dependencia excesiva de los resúmenes o modelos basados en la IA no permite captar esta experiencia vital,

compleja y viva.

“ Existe cierta preocupación (en debate) de que si la IA agota el contenido de calidad generado por humanos que necesita para su entrenamiento —como el periodismo de calidad, el archivo y la investigación— puede llegar a un callejón sin salida en materia de datos ”

Ni las redes sociales ni la IA son fuerzas inevitables. Son iniciativas empresariales. Lo que hace que las redes sociales sean peores para la sociedad es que su modelo de negocio se basa en captar la atención. Durante un tiempo, pareció que el modelo de negocio de la IA podría centrarse más en la infraestructura de la nube que en la publicidad o en captar la atención, pero a principios de 2025 quedó claro que esto está cambiando: OpenAI, uno de los actores más destacados en el ámbito de la IA, ha elegido la publicidad como su modelo principal de negocio, y otros les seguirán. Volvemos a lo de siempre. Mientras el diseño de Internet, las tecnologías digitales y la IA se centren firmemente en maximizar la atención, tendrán un impacto negativo en los conflictos, por medio de la fragmentación, la polarización y la erosión de las oportunidades para que nuestra sociedad llegue mediante la deliberación a un consenso que defina una paz sostenible y pluriversal.

Es más, si en el diseño y la gobernanza de la tecnología no aplicamos principios que tengan en cuenta el contexto de los conflictos (*conflict sensitive principles*), la tecnología puede ser —y será— utilizada como arma de control por parte de actores poderosos. Nerima Wako explica vivamente este riesgo en su apasionado artículo sobre el ciberactivismo en África Oriental —donde, según ella, Internet no es solo el medio, sino el movimiento— y la reacción digital represiva contra esta nueva actividad democrática (principalmente juvenil). Su artículo es una advertencia sobre cómo la regulación puede significar represión, pero también un llamamiento esperanzador y potente a defender un Internet libre y fiable que ofrezca espacios para el diálogo cuando los espacios *offline*

están siendo atacados. Según Wako, «una sociedad pacífica digitalmente es aquella en la que los activistas no necesitan teléfonos de prepago, en la que las leyes protegen la libertad de expresión y en la que podemos discrepar de forma fuerte y apasionada y con seguridad».

Es posible que la gente esté harta de los espacios digitales que son a la vez represivos y polarizantes; tal vez el descenso en el uso de las redes sociales en 2022 sea un presagio de lo que está por venir. Ninguna de estas tecnologías es una fuerza imparable de la naturaleza. Podemos dejar de usarlas. Podemos usarlas de manera diferente. Podemos regularlas. Podemos rediseñarlas. Los seis artículos de este monográfico también apuntan a la esperanza.

Tras pintar un panorama sombrío sobre cómo se ha librado la guerra digital contra la población palestina y cómo esto es un ejemplo de tendencias más extendidas en la guerra de la información moderna, el artículo de Ahmed Qadi enumera una gran cantidad de vías para pasar de la división al diálogo. Del mismo modo, el artículo de Evelyn Tauchnitz termina con un mensaje de esperanza, argumentando que la IA puede contribuir a la paz si alineamos sus principios de diseño no solo con la mejora de la eficiencia y la solidez, sino también con la salvaguarda de la dignidad y la libertad humanas. Añadiendo a estos enfoques más amplios de esperanza, los artículos de Wahbi Abdelrahman y Luke Thorburn aportan ejemplos concretos y prácticos de cómo las tecnologías digitales pueden contribuir a la paz. Stephanie Williams describe cómo el impacto negativo de las redes sociales en el proceso de paz libio provocó un debate en las Naciones Unidas sobre cómo se podrían utilizar los diálogos digitales para aumentar la transparencia en los procesos de mediación.

“ Internet, la tecnología digital y la inteligencia artificial plantean tantos retos para la paz como oportunidades. Abordarlos es una cuestión moral y política que concierne a cualquiera que trabaje por la paz en la era digital ”

¿Cómo podemos lograr que el diseño de tecnología prosocial y de fomento de la paz, tal y como lo describe Thorburn, se convierta en la norma? He aquí una posible respuesta: el artículo de Sanjana Hattotuwa termina con una llamada a los mediadores y constructores de paz para que no solo se vuelvan más hábiles a la hora de navegar por «los comunes digitales adversarios de hoy en día», sino que también influyan en cómo se diseñan.

La conclusión de Hattotuwa es también la dirección que creo que debemos seguir. Internet, la tecnología digital y la inteligencia artificial plantean tantos retos para la paz como oportunidades: los artículos que siguen ofrecen una visión general de lo que nos debe preocupar en un mundo mediado digitalmente y de cómo podemos cambiar el rumbo hacia la paz. Pero lo más importante es que este monográfico defiende que abordar estos retos no es solo una cuestión técnica reservada a quienes sabemos cómo llevar a cabo campañas en las redes sociales para contrarrestar el odio en línea o aplicar la IA en un proceso consultivo. Abordarlos es sobre todo una cuestión moral y política que concierne a cualquiera que trabaje por la paz en la era digital.

Este monográfico de la revista “Por la Paz” (número 43) es una coedición del ICIP y Build Up. La colaboración parte de la coorganización de Build Peace 2025 en Cataluña, la conferencia anual de Build Up sobre tecnología, innovación y construcción de paz. El evento se ha celebrado del 21 al 23 de noviembre en Santa Coloma de Gramenet (Barcelona).

Esta es una versión traducida del artículo publicado originalmente en inglés.

Sobre la autora

Helena Puig Larrauri

Helena es la cofundadora de Build Up, un colectivo de construcción de paz dedicado a innovar en la prevención de conflictos y a abordar la polarización en contextos frágiles. Con más de 15 años de experiencia, ha trabajado junto a organizaciones de la sociedad civil e instituciones multilaterales para diseñar e implementar programas de diálogo,

consulta y análisis de conflictos en entornos divididos y afectados por la violencia. Su experiencia se centra en el diseño de procesos de paz digitales, la inclusión digital en la mediación y la construcción de paz, y en abordar los factores que impulsan los conflictos en espacios en línea e híbridos.

Helena es asesora sénior en Tecnologías Digitales y Mediación para la Unidad de Apoyo a la Mediación de las Naciones Unidas, miembro del registro del Grupo de Estabilización Civil del FCDO y Fellow de Ashoka. Es licenciada en Política, Filosofía y Economía por la Universidad de Oxford y posee un máster en Políticas Públicas (Economía) por la Universidad de Princeton.

Fotografía

Representación simbólica de la colaboración entre humanos y tecnologías digitales.
Autor: Rawpixel.com (Shutterstock).

ARTÍCULOS CENTRALES

Desinformación, manipulación y discurso del odio en los conflictos

Ahmad Qadi

Director de supervisión y documentación de Zamleh

En la era digital actual, los conflictos y las guerras se han extendido cada vez más a la esfera online, amplificando la propagación de la desinformación y la incitación al odio. Las plataformas de las redes sociales se han convertido en espejos (y lupas) de las tensiones del mundo real, sirviendo tanto de campos de batalla como de caldo de cultivo para la hostilidad digital. Si bien los usuarios individuales pueden promover espontáneamente contenidos alineados con sus creencias políticas o ideológicas, los Estados y los actores no estatales suelen participar en campañas coordinadas para difundir deliberadamente información falsa o engañosa con la intención de engañar, manipular o perjudicar. Esta manipulación intencionada de la información se denomina desinformación: la difusión estratégica de falsedades con el fin de engañar a la población, dañar reputaciones y sembrar la división.^[1]

Al mismo tiempo, el discurso del odio prolifera en estas plataformas, a menudo incitado o legitimado por las élites políticas y las estructuras de poder. Este se manifiesta como una comunicación que expresa o incita al odio contra individuos o grupos basándose en características intrínsecas como la etnia, la religión, la nacionalidad o el género.^[2] Con frecuencia, en tiempos de conflicto, el discurso del odio se convierte en un arma para unir a las poblaciones contra un enemigo común, justificar la violencia o intimidar a las voces disidentes.

En el centro de este panorama digital se encuentra la polarización, un fenómeno en el que las sociedades están fracturadas en bandos opuestos con visiones del mundo

rígidas. La polarización se va intensificando a medida que los individuos se atrincheran cada vez más en el apoyo a un bando al tiempo que rechazan por completo o deshumanizan al bando contrario. Esta tríada de desinformación, discurso del odio y polarización no se limita a reflejar las tensiones existentes, sino que las alimenta de manera activa. Juntos, forman un ciclo que se perpetúa y que agrava las divisiones sociales, aumenta la hostilidad y mantiene el conflicto.^[3]

“ La tríada de desinformación, discurso del odio y polarización forma un ciclo que se perpetúa y que agrava las divisiones sociales y aumenta la hostilidad ”

La espiral de polarización en tiempos de guerra

Los estudios empíricos en numerosas zonas de conflicto revelan un patrón recurrente: la desinformación y el discurso de odio surgen en los momentos de mayor tensión política y estallan cuando los conflictos se intensifican y derivan en violencia. En este contexto, las redes sociales se convierten en una herramienta esencial para las enfrentar narrativas. Las partes enfrentadas explotan las plataformas a fin de movilizar apoyos, justificar acciones militares, desacreditar a sus oponentes y manipular la opinión pública. A menudo, las campañas de desinformación buscan minar la moral del enemigo, infundir miedo o desestabilizar la cohesión interna. Estas campañas digitales no se limitan al ciberespacio, sino que tienen consecuencias tangibles, a menudo mortales. Naciones Unidas, por ejemplo, concluyó que Facebook desempeñó un «papel determinante» en el genocidio contra los rohingya en Myanmar. El odio no se quedó en el ámbito virtual, sino que se convirtió en un catalizador de la violencia en el mundo real.^[4]

En contextos de conflicto, la desinformación constituye un sofisticado ecosistema diseñado no solo para confundir, sino también para controlar. Distorsiona la verdad, minimiza el sufrimiento, erosiona la empatía y ahonda las divisiones sociales. En el

fondo, se trata de un arma emocional y psicológica, un componente esencial de la guerra moderna de la información. Esta militarización de los espacios digitales es evidente en numerosos conflictos actuales.

En Palestina, las campañas de desinformación han contribuido a vilipendiar a las víctimas, a negar las atrocidades documentadas y a legitimar los ataques contra periodistas y trabajadores humanitarios. Las narrativas que presentan a los niños y a los supervivientes como «actores de crisis» o que escenifican su sufrimiento sirven para deslegitimar las experiencias reales y proteger a los perpetradores de su responsabilidad. Estas estrategias no solo niegan justicia a las víctimas, sino que también manipulan la opinión pública mundial y configuran el discurso político. Aunque algunos elementos pueden surgir de forma espontánea, especialmente durante las crisis, la tendencia dominante de estas campañas es deliberada y organizada. A través de redes coordinadas, las partes en conflicto elaboran y difunden contenidos destinados a manipular el consenso, atribuir culpas y construir visiones binarias del mundo que dejan poco espacio para la complejidad, la empatía o el diálogo.^[5]

“ La desinformación es una arma emocional y psicológica. En contextos de conflicto, distorsiona la verdad, minimiza el sufrimiento, erosiona la empatía y ahonda las divisiones sociales ”

Si bien se supone que es cierto que esas formas de contenido se exacerban durante los conflictos, la incitación al odio y la desinformación suelen tener su raíz en narrativas sociales más amplias, ya sean nacionalistas, religiosas o ideológicas. Estas narrativas, arraigadas a lo largo del tiempo, resurgen en el discurso digital como contenidos violentos o engañosos. Los actores políticos suelen amplificar estas historias, presentando a su propio bando como el moralmente justo y agraviado. Si bien estas narrativas pueden simplificar realidades complejas en oposiciones binarias entre el bien y el mal, los algoritmos de las redes sociales refuerzan este pensamiento

reduccionista al privilegiar los contenidos con carga emocional y desalentar los matices. Este entorno es terreno fértil para la propagación de la desinformación y la incitación.^[6]

El discurso del odio y la desinformación, entre otras formas de contenido violento, son el resultado directo de la polarización tóxica, que va más allá del desacuerdo ideológico normal y surge cuando los individuos desarrollan un profundo desprecio por quienes tienen creencias opuestas a las suyas, al tiempo que expresan una intensa lealtad y apego a las opiniones de su propio grupo. Esta forma de polarización transforma las diferencias políticas o ideológicas en divisiones basadas en la identidad, fomentando la percepción de que el bando contrario no sólo está equivocado, sino que es un enemigo irreconciliable. La investigación psicológica identifica tres factores fundamentales que impulsan esta dinámica: la deshumanización, la aversión y el desacuerdo. Cuando los miembros de un grupo creen que la otra parte los detesta, los deshumaniza o se opone a ellos de manera fundamental, la polarización se intensifica.^[7]

Esto es exactamente lo que ocurrió en el caso del genocidio que se está produciendo en Gaza. Desde que estalló la guerra de Israel contra Gaza el 7 de octubre de 2023, una enorme guerra de desinformación se ha extendido por los espacios digitales, y muchos la describen como algo sin precedentes. Israel lanzó una amplia campaña integral de desinformación e influencia diseñada para justificar su ataque militar contra Gaza y deslegitimar los derechos de los palestinos. Un elemento central de esta iniciativa fue una operación encubierta de 2 millones de dólares, financiada por el Ministerio de Asuntos de la Diáspora de Israel, que utilizó contenido generado por inteligencia artificial y bots para manipular la opinión pública y promover narrativas deshumanizantes sobre los palestinos. Esta campaña estuvo dirigida estratégicamente a los legisladores estadounidenses, especialmente a los demócratas, a través de plataformas como Facebook, Instagram y X, con propaganda personalizada producida por la empresa israelí STOIC. OpenAI desbarató posteriormente algunos de estos esfuerzos, revelando una sofisticada operación destinada a amplificar el sentimiento islamófobo y antipalestino en los espacios digitales occidentales.

“ La guerra de Gaza ha demostrado cómo los conflictos *offline* pueden intensificar drásticamente la violencia y la desinformación *online*, aumentando narrativas falsas ”

Al mismo tiempo, el Ministerio de Asuntos Exteriores israelí lanzó una campaña publicitaria gráfica y emocionalmente manipuladora en YouTube, en Europa y Norteamérica, diseñada para suscitar el apoyo a las acciones militares de Israel. Según el análisis de 7amleh, estos anuncios, que incluían imágenes perturbadoras y llamamientos enmarcados en narrativas centradas en los niños, violaban las normas de la plataforma y, sin embargo, seguían activos. Además, personas vinculadas a Israel se pusieron en contacto con *influencers* para ofrecerles pagos y «sesiones informativas» con el fin de difundir mensajes a favor de Israel en las redes sociales, lo que contribuyó a afianzar aún más las narrativas patrocinadas por el Estado en los espacios digitales de base. Esta manipulación orquestada pone de relieve la guerra digital desigual en la que Israel aprovecha herramientas avanzadas y vastos recursos para dominar el discurso y suprimir las voces palestinas.^[8]

La guerra de Gaza ha demostrado, de forma inequívoca, cómo los conflictos *offline* pueden intensificar drásticamente la violencia y la desinformación *online*. La existencia de toda una página de Wikipedia dedicada a rastrear la enorme cantidad de desinformación desatada durante el genocidio subraya la magnitud de este fenómeno. Destaca no solo cómo el conflicto fuera de línea alimentó un aumento de narrativas falsas y dañinas en línea, sino también cómo gran parte de esta desinformación fue impulsada sistemáticamente por actores políticos, más allá de la difusión inconsciente de información errónea por parte de usuarios comunes.^[9]

Al mismo tiempo, los oponentes de Israel, como Irán, han intensificado sus esfuerzos de desinformación, especialmente durante el recrudecimiento de julio de 2025 tras la ofensiva israelí. Una campaña iraní coordinada que utilizó más de 100 cuentas bot en X (anteriormente Twitter) se difundió más de 240.000 veces, con el objetivo de influir en

la opinión pública estadounidense y disuadir de posibles ataques contra las instalaciones nucleares de Irán. La operación promovió al líder supremo de Irán, difundió afirmaciones falsas sobre los fracasos militares de Israel y describió a Israel como un Estado terrorista. Las publicaciones utilizaron imágenes y hashtags incendiarios para maximizar la viralidad y socavar la alianza entre los Estados Unidos e Israel. Los analistas consideran que la campaña forma parte de una estrategia más amplia de Irán para manipular el discurso global y prevenir la acción militar mediante la guerra psicológica e informativa.^[10]

“ La naturaleza emocional de los mensajes relacionados con el conflicto, marcados por el miedo, la ira o el dolor, aumenta aún más la susceptibilidad a la manipulación y la confusión

”

Varios factores contribuyen a la propagación de la desinformación y otras formas de contenido polarizador durante los conflictos. Entre ellos se encuentran la falta generalizada de verificación y pensamiento crítico a la hora de evaluar la información recibida, así como la tendencia a considerar creíble un contenido simplemente porque proviene de una fuente interna del grupo. La desinformación suele percibirse como factual desde una cosmovisión específica, especialmente cuando coincide con estereotipos profundamente arraigados o narrativas tradicionales.

Los sesgos cognitivos también desempeñan un papel fundamental: los individuos tienden a buscar información que confirme sus creencias existentes, evitan los contenidos que cuestionan esas creencias y se resisten a actualizar sus puntos de vista incluso cuando se enfrentan a pruebas contrarias creíbles. La naturaleza emocional de los mensajes relacionados con el conflicto (marcados por el miedo, la ira o el dolor) aumenta aún más la susceptibilidad a la manipulación y la confusión. La gente puede reconocer la existencia de la desinformación, pero creer que solo los demás son vulnerables a ella, lo que refuerza los puntos ciegos de su propio juicio. El

efecto de cámara de eco profundiza estas dinámicas al limitar la exposición a perspectivas alternativas, mientras que la desinformación encubierta bajo el lenguaje de la autoridad pública, institucional o científica adquiere una legitimidad injustificada. En algunos casos, paradójicamente, la exposición a información precisa puede afianzar la creencia en falsedades, ya que los individuos reinterpretan o rechazan la verdad de formas que refuerzan sus convicciones originales. En conjunto, estas dinámicas crean un entorno fértil para que prospere la desinformación y se intensifique la polarización.^[11]

El papel de las redes sociales

Las plataformas de redes sociales han desempeñado un papel significativo en la escalada de conflictos al incentivar discursos de división y potencialmente violentos, a menudo amplificando contenidos que fomentan la polarización y el acoso masivo a través de algoritmos optimizados para métricas de participación como comparticiones, comentarios y tiempo de permanencia. Aunque las plataformas han enfrentado los conflictos violentos principalmente mediante la moderación reactiva de los contenidos (respondiendo a incidentes o brotes específicos), este enfoque no aborda los problemas de diseño más profundos y sistémicos que alimentan la dinámica del conflicto mucho antes de que estalle la violencia. Investigaciones independientes, como la relativa al papel de Facebook en Myanmar, han demostrado cómo las plataformas pueden facilitar la incitación a la violencia fuera de Internet al fomentar la división. La investigación respalda un panorama complejo: las redes sociales se correlacionan tanto con la polarización como con un mayor conocimiento político, pero la primera —especialmente la polarización tóxica, en la que la gente demoniza a los grupos opuestos— es un preludio más peligroso de la violencia que los meros desacuerdos políticos.^[12]

“ Las redes sociales desempeñan un papel significativo en la escalada de conflictos al incentivar discursos violentos. El impacto no se limita al ámbito digital, opera dentro de un ciclo

de refuerzo que incita a la polarización y a la violencia *offline* ”

El informe de la Stern School of Business de la Universidad de Nueva York titulado «Fueling the Fire: How social media intensifies U.S. political polarization and what can be done about it» concluye que, si bien las redes sociales no son la causa principal de la polarización política en Estados Unidos, desempeñan un papel fundamental en la intensificación de la polarización afectiva (la hostilidad y el desprecio profundamente arraigados entre los grupos políticos) lo que, a su vez, erosiona las normas democráticas, debilita la confianza en las instituciones y alimenta la violencia en el mundo real. El informe destaca cómo los algoritmos basados en la participación amplifican sistemáticamente los contenidos divisivos y, a pesar de las ocasionales medidas internas, las plataformas han fracasado en gran medida a la hora de autorregularse.^[13]

La violencia *offline* y *online* funciona como un ciclo de refuerzo, a través del cual cualquier incidente *offline* podría alimentar la violencia *online* y viceversa. La organización 7amleh documentó un ejemplo destacado de Palestina que ilustra el papel de las redes sociales: analizó la escalada digital que condujo al ataque de febrero de 2023 contra la aldea de Huwara, en Cisjordania, donde cientos de colonos israelíes llevaron a cabo un violento asalto que se saldó con el asesinato de un palestino, la destrucción generalizada de propiedades (incluyendo la quema de cultivos y vehículos), ataques a viviendas y el terror infligido a los residentes. En los meses previos al ataque, 7amleh identificó más de 15.000 contenidos violentos en lengua hebrea en plataformas de redes sociales, que iban dirigidos directamente contra la aldea y sus habitantes. Este contenido sirvió para deslegitimar, calumniar y deshumanizar a la población palestina local, retratándola en términos inhumanos y amenazadores. Esa incitación digital sentó las bases de la violencia en el mundo real al normalizar la hostilidad y justificar la agresión contra la comunidad.^[14]

De la división al diálogo: polarización inversa

Se puede hacer mucho para acabar con los prejuicios y la polarización, y una de las hipótesis es, simplemente, la comunicación. El contacto entre grupos puede tener un impacto significativo entre las partes en conflicto, pero solo si se hace correctamente, ya que a veces, también puede agravar la polarización. Por ejemplo, seguir a los oponentes en X podría consolidar las propias ideas extremas o los prejuicios contra los demás. Si el contacto se lleva a cabo de forma sostenible, con un intercambio respetuoso de ideas entre personas del mismo rango o edad, por ejemplo, eso podría reducir la polarización. La intervención puede ser realizada por gobiernos, organizaciones de la sociedad civil, medios de comunicación, plataformas de redes sociales y otros actores, con objeto de difundir las perspectivas de los demás. Escuchar historias desde la perspectiva de otros podría ayudar a reducir gran parte de los prejuicios y la polarización. Sin embargo, las redes sociales han hecho más mal que bien al permitir que las personas vean y se relacionen con personas o narrativas similares, en lugar de abrirles a otras perspectivas.^[15]

“ Escuchar historias desde la perspectiva de otros podría ayudar a reducir gran parte de los prejuicios y la polarización ”

Para contrarrestar esto, iniciativas como Beyond Conflict recomiendan poner en marcha campañas de concienciación pública destinadas a exponer las percepciones erróneas generalizadas que los partidarios tienen unos de otros y piden que las figuras influyentes asuman sus responsabilidades cuando difunden información falsa. Entre las estrategias de orden práctico para mitigar la polarización se incluyen fomentar el diálogo entre grupos, promover la empatía a través de la toma de perspectivas y destacar los desacuerdos internos dentro de los grupos políticos para poner fin a las narrativas rígidas de «nosotros contra ellos». Además, promover la amabilidad en las redes sociales puede ayudar a dejar de normalizar la deshumanización, del mismo modo que evitar la repetición de información errónea y abstenerse de bromas prejuiciosas puede debilitar la aceptabilidad social de la retórica de división.^[16]

Otros enfoques abogan por pasar de una moderación reactiva a intervenciones proactivas, a largo plazo y escalables, basadas en el diseño de las plataformas. La propuesta es que las redes sociales se alejen de la clasificación de contenidos basada en la participación en contextos sensibles, limiten la capacidad de difusión masiva e introduzcan características de diseño que fomenten interacciones significativas y de conexión. Las plataformas también deben integrar el apoyo a los esfuerzos de consolidación de la paz, reconociendo que la paz no es solo la ausencia de violencia, sino la presencia de unas condiciones sociales en las que todas las comunidades puedan prosperar. En última instancia, los conflictos se intensifican mediante ciclos de refuerzo, y para romper esta espiral es preciso acabar con los incentivos que premian la división y la manipulación en línea.^[17]

Otras recomendaciones abogan por reformas estructurales integrales, instando a las plataformas de redes sociales a rediseñar de forma transparente sus sistemas algorítmicos a fin de reducir la difusión de contenidos incendiarios. Por ejemplo, el informe de la Stern School of Business antes mencionado subraya la importancia de invertir en equipos internos de moderación de contenidos y fomentar una colaboración más estrecha con las organizaciones de la sociedad civil. También destaca el papel fundamental de la intervención gubernamental, abogando por una legislación que exija transparencia, faculte a los organismos reguladores para hacer cumplir las normas de conducta y apoye el desarrollo de plataformas digitales alternativas que promuevan la participación democrática. Finalmente, el informe apunta a la polarización descontrolada impulsada por las redes sociales como una amenaza directa para la estabilidad democrática, lo que exige una acción urgente y coordinada entre todos los sectores.^[18]

Conclusión

Las redes sociales y las herramientas digitales desempeñan un papel significativo y, a menudo, perjudicial a la hora de alimentar la violencia, la división y la polarización, tanto en tiempos de paz como, de forma más marcada, durante la guerra. Las partes en conflicto explotan activamente estas plataformas para amplificar sus narrativas, recurriendo con frecuencia a la desinformación y al discurso del odio como estrategias para deslegitimar al bando contrario. Las redes sociales ofrecen un medio rentable y de

gran alcance para la difusión de esos contenidos, lo que las convierte en un poderoso instrumento en la guerra de la información moderna. Resulta fundamental señalar que el impacto de este contenido no se limita al ámbito digital. Opera dentro de un ciclo de refuerzo que intensifica la polarización, ahonda las fracturas sociales y, en última instancia, incita a la violencia *offline*. Esta dinámica subraya la urgente necesidad de que las plataformas rindan cuentas y de replantearse, de forma fundamental, el modo en que las infraestructuras digitales interactúan con la dinámica de los conflictos.

- [1] Stavros, A., S. Phalen, S. Almakki, M. Nacionales-Tafoya, y R. A. García. 2023. “Disinformation in Conflict Environments in Asia.” Gerald R. Ford School of Public Policy, University of Michigan.
- [2] 7amleh (Centro Árabe para el Avance de las Redes Sociales) 2022. “A Guide to Combating Online Hate Speech.”
- [3] Polarization Research Lab. 2022. “How Do You Study and Reverse Political Animosity? These Researchers Are Working to Answer That Question.” Charles Koch Foundation.
- [4] Stavros et al. 2023.
- [5] Barforoush, S., y S. Plaut. 2024. “Information Disorder in Times of Conflict.” Canadian Museum for Human Rights.
- [6] Cobb, S., S. Kaplan, A. Marc, y G. Milante. 2021. “The Role of Narrative in Managing Conflict and Supporting Peace.” Discussion paper, IFIT, Institute for Integrated Transitions, Barcelona.
- [7] Equipo editorial de Psico-smart. 2024. “The Role of Technology in Conflict Mediation: Exploring Digital Tools and Platforms.” Psico-smart Blog..
- [8] Qadi, A. 2025. “From Bot Farms to Censorship: Israel’s Disinformation Warfare against Palestinians.” *Palestine Chronicle*.
- [9] “Misinformation in the Gaza War.” 2025.

- [10] Kahana, A. 2025. ["Iran Deployed Bots to Post 240K Times to Block US Strikes on Nuclear Facilities."](#) NYP.
- [11] Lewandowski, P. 2024. ["Psychological Mechanisms of Disinformation and Their Impact on Social Polarization."](#) *Studia Polityczne* 2: 85-104. Łukasiewicz Research Network - ITECH Institute of Innovation and Technology.
- [12] Stray, J., R. Iyer, y H. Puig Larrauri. 2023. ["The Algorithmic Management of Polarization and Violence on Social Media."](#) Knight First Amendment Institute.
- [13] Barrett, P. M., J. Hendrix, y J. G. Sims. 2021. ["Fueling the Fire: How Social Media Intensifies U.S. Political Polarization and What Can Be Done about It."](#) Center for Business and Human Rights, Stern School of Business, New York University.
- [14] 7amleh (Centro Árabe para el Avance de las Redes Sociales). 2023. ["An Analysis of the Israeli Inciteful Speech against the Village of Huwara on Twitter."](#)
- [15] Moskalenko, S. 2023. ["What Are the Solutions to Political Polarization?"](#) *Greater Good Magazine*, Greater Good Science Center, University of California.
- [16] Equipo editorial de Psico-smart. 2024. ["The Role of Technology in Conflict Mediation: Exploring Digital Tools and Platforms."](#) Psico-smart Blog.
- [17] Stray, J., R. Iyer, y H. Puig Larrauri. 2023. ["The Algorithmic Management of Polarization and Violence on Social Media."](#) Knight First Amendment Institute
- [18] Barrett, P. M., J. Hendrix, y J. G. Sims. 2021. ["Fueling the Fire: How Social Media Intensifies U.S. Political Polarization and What Can Be Done about It."](#) Center for Business and Human Rights, Stern School of Business, New York University.

Esta es una versión traducida del artículo publicado originalmente en [inglés](#).

Sobre el autor

Ahmad Qadi

Ahmad Qadi es director de supervisión y documentación de Zamleh. Trabaja en moderación de contenidos, políticas de redes sociales y censura. Ahmad es defensor de los derechos digitales, investigador y estudiante de doctorado que examina el impacto de la tecnología en las sociedades.

Fotografía

Metáfora visual sobre la desinformación y la polarización en el entorno digital. Autor: Evan Huang (Shutterstock).

ARTÍCULOS CENTRALES

Riesgos de la IA para una paz sostenible

Evelyne Tauchnitz

Investigadora sénior y profesora en el Instituto de Ética Social, Universidad de Lucerna

A menudo se malinterpreta la paz como la mera ausencia de guerra o violencia

manifiesta: un estado de orden y estabilidad. Pero esta concepción limitada de lo que Galtung ha denominado «paz negativa»^[1] pasa por alto los fundamentos relacionales, sistémicos y éticos más profundos que hacen que la paz sea sostenible en el tiempo. Por el contrario, la paz sostenible no es un logro estático o un estado final institucional, sino una condición dinámica que surge de la calidad de las relaciones, la resiliencia de las redes sociales y la presencia de la dignidad y la libertad humanas.

Por lo tanto, para comprender mejor la paz es necesario ir más allá de los ideales abstractos o los marcos de los tratados y prestar atención a cómo se relacionan entre sí las personas, los grupos y las instituciones. La paz no es una sustancia o una estructura, sino un patrón positivo de interacción, algo que está estrechamente ligado al comportamiento humano. Es un fenómeno que ocurre «entre los nodos»: en las conexiones, los intercambios y los reconocimientos mutuos que unen a las personas a través de experiencias e intercambios humanos positivos. Estos vínculos pueden trazarse como redes, en las que cada nodo $\otimes$ ya sea una persona, un grupo o una institución $\otimes$ es portador de capacidades y vulnerabilidades, y cada eslabón que conecta los diferentes nodos refleja una cualidad particular de la relación, una combinación única de características como la confianza y el cuidado, pero también el conflicto o el abandono. Esta visión de una paz interconectada desplaza el foco de atención desde los indicadores estáticos hacia la dinámica de la conexión: cómo se construyen, tensan o rompen las relaciones, cómo se distribuye el poder y cómo coexisten la paz y la violencia en una línea continua. Las redes pueden ser robustas, salvando divisiones y

apoyando la cooperación; o pueden ser frágiles, marcadas por el miedo, la exclusión y la fragmentación. Así pues, la paz no es algo dado, sino, al igual que la violencia, una propiedad emergente de las condiciones que definen las características del sistema, siempre en riesgo de venirse abajo, especialmente cuando se ve perturbado por crisis, injusticias o también transformaciones tecnológicas (incluida la Inteligencia Artificial, IA).

“ La dignidad y la libertad son los fundamentos éticos de la paz, y la resiliencia es su columna vertebral: saber gestionar las crisis y transformar los conflictos de manera sostenible depende del nivel de resiliencia de las comunidades ”

En el centro de las redes pacíficas se encuentran los principios éticos de la dignidad humana y la libertad. La dignidad no es solo un valor intrínseco, sino una cualidad relacional, que se mantiene cuando las personas son reconocidas en su singularidad, y tratadas como personas con voz y valor. La libertad no es solamente estar libre de coerción, sino la capacidad de participar de forma significativa en la configuración de la propia vida y las relaciones. Como han argumentado Hannah Arendt[2] y Amartya Sen [3], la libertad es la capacidad real de decidir y actuar dentro de un contexto, en estructuras que la pueden impulsar o limitar. Cuando se niegan sistemáticamente la dignidad y la libertad mediante la vigilancia, la exclusión o el desinterés la infraestructura relacional de la paz comienza a deshacerse. La confianza se erosiona, la participación se desvanece y la violencia surge en su mayor parte no del odio, sino del colapso sistémico: la «banalidad del mal», como advertía Arendt, donde los actores ordinarios siguen normas perjudiciales en sistemas disfuncionales. El genocidio en Gaza y el aumento de la violencia por parte de los colonos en Cisjordania tras el ataque de Hamás del 7 de octubre de 2023 constituyen un trágico ejemplo.

Mientras que la dignidad y la libertad son los fundamentos éticos de la paz, la resiliencia es su columna vertebral operativa: el conjunto de capacidades que permiten

a los individuos, las comunidades y los sistemas resistir a las perturbaciones y recuperarse o transformarse sin caer en la violencia. Las crisis e incluso los conflictos siempre ocurrirán, pero el hecho de que estalle la violencia o que, por el contrario, los individuos, las comunidades y las instituciones consigan gestionar las crisis de forma sostenible y transformar los conflictos por medios no violentos en estados de coexistencia pacífica, depende de sus niveles de resiliencia. La resiliencia no consiste únicamente en sobrevivir o recuperarse. Se trata de afrontar las crisis de forma que se preserven los valores y las relaciones fundamentales, adaptarse a las condiciones cambiantes y abrir espacio para el cambio positivo y la renovación.

Se pueden identificar cuatro pilares interdependientes de la resiliencia:

1. **Recursos:** Acceso a alimentos, refugio, atención sanitaria, conocimientos, sentido emocional: todas las herramientas materiales y simbólicas necesarias para hacer frente a las crisis.
2. **Capital social:** Confianza, solidaridad y redes de apoyo mutuo que permitan a las personas compartir responsabilidades, acceder a recursos y coordinar respuestas.
3. **Capacidad de adaptación:** Capacidad de aprender, innovar y ajustar estrategias cuando las antiguas ya no funcionan.
4. **Entornos propicios:** Instituciones y políticas que garanticen la equidad, protejan los derechos humanos y ofrezcan vías para el cambio pacífico.

Juntos, estos factores determinan si la paz puede perdurar bajo presión. Cuando la resiliencia es fuerte, las redes se flexionan pero no se rompen; cuando es débil, las crisis pueden llevar a los sistemas más allá de puntos de inflexión en los que las espirales de violencia se autorrefuerzan (véanse, por ejemplo, los recientes brotes de violencia tribal en Siria, en 2025). Estos puntos de inflexión se producen cuando las normas sociales cambian hacia el miedo o la agresión, los recursos escasean, la confianza se erosiona o los actores oportunistas se aprovechan de los conflictos. Como en la teoría de redes, una pequeña ruptura en una parte del sistema puede producir una reacción en cascada, especialmente si los lazos relacionales ya están tensos. En la siguiente sección se examina la fuerza disruptiva de la IA y su impacto potencial en la resiliencia que configura la capacidad de las redes para salvaguardar la paz.

La IA como riesgo para la resiliencia de las redes y su capacidad para salvaguardar la paz

La inteligencia artificial está reconfigurando la forma de trabajar, comunicarse y gobernarse de las personas, a menudo enmarcada en términos de innovación, eficiencia u optimización. Pero cuando se mira a través de la lente de la paz relacional, la IA revela un riesgo más profundo: reduce la resiliencia de los individuos y las sociedades a la hora de gestionar las crisis y los conflictos de forma pacífica.

Comprender el impacto de la IA en la paz requiere, por tanto, ir más allá de amenazas manifiestas como los sistemas de armas autónomas o las nuevas amenazas cibernéticas. En lo que respecta a la resiliencia, deben examinarse los efectos de la IA sobre los recursos, el capital social, la capacidad de adaptación y los entornos propicios, elementos que forman la infraestructura relacional de la paz sostenible.

La IA y la interrupción de los recursos: dignidad a través del trabajo y la integridad emocional

Los recursos son la base de la resiliencia en cualquier red, no solo los bienes materiales como la comida o el refugio, sino también las capacidades simbólicas, emocionales y de información. Las tecnologías de la IA controlan cada vez más el flujo y la distribución de estos recursos y, al hacerlo, remodelan las condiciones en las que se afirma o se niega una vida en dignidad.

“ Desde la lente de paz relacional, la IA revela un riesgo profundo: reduce la resiliencia de los individuos y las sociedades para gestionar las crisis y los conflictos de manera pacífica ”

Desde un punto de vista estructural, las herramientas potenciadas por la IA se están ocupando cada vez más de tareas administrativas repetitivas que tradicionalmente

realizaban los empleados de oficina: por ejemplo, procesar facturas, actualizar registros de nóminas o verificar formularios electrónicos. Aunque es posible que estos sistemas no sustituyan funciones enteras de la noche a la mañana, reducen significativamente la demanda de mano de obra, lo que se traduce en menos oportunidades laborales. En otros sectores se observan tendencias similares: Los *chatbots* de IA gestionan ahora una gran parte de las consultas de atención al cliente; los sistemas de visión por ordenador en los almacenes guían a los robots autónomos en la recogida, clasificación y embalaje de mercancías; en las finanzas, los sistemas de *trading* algorítmico ejecutan operaciones y optimizan carteras que antes eran gestionadas por equipos de analistas junior.

Aunque surgen algunos puestos de trabajo nuevos —por ejemplo, en la formación de modelos de IA, el etiquetado de datos o el mantenimiento y la supervisión de sistemas automatizados— la naturaleza de estas oportunidades suele definirse de forma limitada. Muchos están vinculados a funciones específicas dentro del proceso de desarrollo de la IA, como la anotación de conjuntos de datos para el aprendizaje automático, el ajuste fino de modelos lingüísticos, la supervisión de resultados algorítmicos en busca de errores o el mantenimiento de maquinaria autónoma. Otros conllevan funciones de supervisión en las que las personas examinan los procesos automatizados, interviniendo solamente cuando el sistema encuentra excepciones o fallos. Además, muchas de estas nuevas oportunidades de empleo suelen ser temporales, estar concentradas geográficamente o requerir conocimientos avanzados que muchos trabajadores desplazados no pueden adquirir fácilmente.

Por lo tanto, es probable que el efecto neto de la IA sobre el empleo sea desigual y perturbador por tres razones. En primer lugar, la automatización tiende a beneficiar a los actores altamente cualificados o con gran poder adquisitivo, mientras que los trabajadores del Sur Global, los mercados laborales precarios y los trabajos rutinarios se enfrentan al mayor riesgo de desposesión. En segundo lugar, los periodos de transición entre la pérdida del trabajo y un nuevo empleo pueden ser largos y desestabilizadores, erosionando así la seguridad económica. En tercer lugar, incluso cuando la IA aumenta la eficiencia y reduce el coste de los bienes, esto no se traduce automáticamente en mejores condiciones de vida.

“ La IA contribuye al «colonialismo de datos»: Se extraen ingentes cantidades de datos y las personas pasan a ser leíbles y explotables siguiendo patrones históricos de extracción y dominio de los recursos ”

Las ganancias se distribuyen a menudo de forma desigual: mientras que algunos individuos y regiones se benefician de productos más baratos, mejoras de los servicios y nuevas oportunidades de mercado, otros experimentan la pérdida de empleo, el estancamiento salarial o el debilitamiento de las protecciones laborales. Estas disparidades pueden aumentar las desigualdades existentes, tanto dentro de los países como entre ellos, ya que los trabajadores altamente cualificados y los propietarios de capital se llevan una parte desproporcionada de los beneficios. El resultado para muchos es una pérdida de dignidad económica: la capacidad de mantenerse a sí mismo y a su familia con propósito y seguridad. Estos cambios tecnológicos proporcionan un sólido argumento a favor de una renta básica incondicional: un pago en metálico garantizado y regular a todos los ciudadanos, basado en el derecho a una vida digna en la que se satisfagan las necesidades básicas de todos.

Además, los sistemas algorítmicos de contratación y rendimiento  como, por ejemplo, las herramientas automatizadas de selección de CV o los programas informáticos de seguimiento de la productividad basados en la IA  reducen a las personas a métricas, lo que da lugar a la discriminación (por ejemplo, rechazando candidatos debido a brechas en el empleo) y la infravaloración de las dimensiones humanas del trabajo como la creatividad, la atención y la colaboración.

A escala mundial, la IA contribuye a lo que los eruditos han denominado «colonialismo de datos».[4] Se extraen ingentes cantidades de datos  a menudo sin consentimiento  de individuos, comunidades y dispositivos de todo el mundo, solo para ser analizados, monetizados y controlados por un puñado de empresas y estados poderosos. Este proceso refleja los patrones históricos de extracción y dominio de recursos, con el giro

inesperado de que el «recurso» en cuestión es la huella relacional y de comportamiento de las vidas humanas. La dignidad se ve comprometida no solo porque se elude el consentimiento, sino porque las personas se vuelven legibles y explotables sin reciprocidad ni recurso.

Incluso los recursos emocionales y psicológicos se ven afectados. Las plataformas mejoradas con IA como el feed Para ti de TikTok, la página Explorar de Instagram, el News Feed de Facebook, el algoritmo de línea temporal de X (antes Twitter) o la reproducción automática y los vídeos sugeridos de YouTube utilizan el aprendizaje automático para predecir y amplificar los contenidos con más probabilidades de captar y mantener la atención del usuario.

“ La IA puede mejorar el acceso a la información o a determinados servicios, pero también socava la dignidad humana ”

Especialmente en las redes sociales, estos mecanismos intensifican la comparación, la competición y la performatividad, ya que los usuarios están continuamente expuestos a representaciones seleccionadas de los logros, estilos de vida y opiniones de los demás. Mediante la captación de la atención y las emociones con fines lucrativos, estos sistemas erosionan la estabilidad interior y autoestima, dimensiones esenciales de la resiliencia. Los sistemas de vigilancia degradan aún más estos recursos al producir una ansiedad constante de bajo nivel, especialmente entre las poblaciones marginadas y los actores de la sociedad civil que ya se enfrentan a una vigilancia desproporcionada, ya sea mediante algoritmos policiales predictivos, la supervisión en el lugar de trabajo, el reconocimiento facial automatizado en espacios públicos o la vigilancia del activismo en línea impulsada por la IA.

En conjunto, la IA reconfigura los flujos de recursos de forma que puede mejorar el acceso a la información o a determinados servicios, pero también socava la dignidad que se deriva del hecho de que se valore el propio trabajo, se satisfagan las propias necesidades y se respeten los propios límites.

La IA y la erosión del capital social: fracturar la confianza que sustenta la paz

El capital social es la red de confianza, normas y relaciones informales que unen a las personas. Es un pilar básico de la resiliencia en cualquier comunidad. Permite la coordinación, la ayuda mutua y la acción colectiva. Cuando la confianza y la empatía circulan libremente, las redes tienden a doblarse en lugar de romperse bajo tensión. Cuando son débiles o frágiles, las crisis se inclinan más fácilmente hacia la culpa, el miedo y la fragmentación.

Los sistemas de IA intervienen cada vez más en el entramado relacional que sustenta la paz. De manera más visible, los algoritmos de las redes sociales, optimizados para el compromiso, dan prioridad a los contenidos que provocan indignación, afirmación u otras emociones fuertes. Esto cambia el equilibrio de lo que circula dentro de las redes y entre ellas, dando mayor visibilidad al material divisivo o con carga emocional y desplazando al mismo tiempo el contenido que fomenta la deliberación o el matiz. Con el tiempo, este filtrado algorítmico fomenta la formación de cámaras de eco polarizadas: agrupaciones muy unidas dentro de redes más amplias en las que la información fluye principalmente entre miembros de ideas afines. Estos conglomerados fragmentan la red general, debilitando el capital social que hace de puente las conexiones que unen a las personas de distintos orígenes o perspectivas mientras refuerzan el capital social de unión dentro de grupos homogéneos. A medida que se fortalecen las identidades dentro del grupo, también lo hace la polarización tóxica (afectiva), una forma de división en la que los individuos no solo discrepan, sino que desarrollan un profundo desprecio por quienes tienen opiniones opuestas, junto con una intensa lealtad a su propio grupo. Esta mentalidad de «nosotros contra ellos» reduce el abanico de perspectivas que encuentra la gente y menoscaba la confianza, la reciprocidad y las normas compartidas que sustentan la cooperación. Aunque la polarización no es nueva, la IA la acelera y la amplifica a menudo de forma invisible cambiando la dinámica de la red que favorece el entendimiento y el reconocimiento mutuo.

“ Los algoritmos de las redes sociales dan prioridad a los contenidos que generan división e indignación. Así, la IA acelera y amplifica afirmación la polarización ”

Más allá de dar forma a lo que se ve en Internet, la IA moldea cada vez más la forma en que sentimos y nos comportamos. Las tecnologías de IA implementadas en las redes sociales, el comercio minorista, la enseñanza, los lugares de trabajo y la administración pública utilizan a menudo el análisis facial, la detección del tono de voz y los datos biométricos para inferir emociones en tiempo real. Una vez detectados, estos estados emocionales pueden utilizarse para adaptar el contenido o las respuestas: un cliente frustrado podría recibir un tono tranquilizador de un *chatbot*; un estudiante señalado como desmotivado por una plataforma de aprendizaje de IA podría recibir material más estimulante; se podría solicitar a un empleado cuya voz no se considere lo bastante «positiva» que adoptara un tono más optimista en las llamadas.

Aunque a menudo se presentan como herramientas para mejorar la calidad de los servicios o la eficacia en el lugar de trabajo, las mismas técnicas para leer las emociones e influir en ellas pueden reutilizarse en contextos políticos o sociales. Por ejemplo, las campañas políticas o de activistas podrían usar la IA emocional para identificar y dirigirse a personas ya preparadas para la ira o el miedo, impulsando contenidos que refuercen los agravios y fortalezcan la solidaridad dentro del grupo. Ambos efectos intensifican las divisiones existentes, haciendo más difícil que los individuos interactúen de manera constructiva por encima de las diferencias. Cuando se combinan con la vigilancia impulsada por la IA, los efectos se magnifican. La conciencia permanente de estar siendo vigilados o perfilados fomenta la autocensura y reprime la disidencia. Esto mina la autenticidad, la confianza y la apertura necesarias para una paz resiliente, sustituyéndolas por interacciones cautelosas y calculadas que debilitan las propias redes de las que depende la cohesión social.

En regímenes autoritarios como, por ejemplo, Rusia o China, la vigilancia impulsada por la IA y los sistemas de crédito social formalizan esta ruptura de la confianza. Cuando

las personas temen que cualquier vínculo social pueda hacerlas vulnerables al control del Estado o a daños en su reputación, comienzan a retirarse de la vida cívica y comunitaria. El tejido social se deshilacha no por la violencia directa, sino por la corrosión relacional, un lento colapso de las instituciones intermediarias y de las solidaridades cotidianas.

La IA socava la capacidad de adaptación: encerrar a las sociedades en prácticas perjudiciales del pasado

La capacidad de adaptación es quizás la forma más sutil pero esencial de resiliencia. Es lo que permite a los individuos y a las instituciones afrontar las novedades, responder a la retroalimentación y cambiar de rumbo cuando las viejas estrategias ya no funcionan. Sin capacidad de adaptación, los sistemas se vuelven frágiles, incapaces de recuperarse o transformarse bajo presión. En un mundo moldeado por la IA, esta capacidad es cada vez más limitada.

“ Donde antes las personas confiaban en su propio juicio y sentido común, ahora se espera que cumplan protocolos algorítmicos ”

Muchos sistemas de IA están diseñados para optimizar resultados específicos basándose en datos históricos. Pero al hacerlo, a menudo se aferran a ciertos supuestos, métricas o patrones que se resisten a la contestación o a la adaptación. Por ejemplo, los sistemas policiales predictivos pueden perpetuar prejuicios históricos basados en el origen étnico o socioeconómico. Además, las personas afectadas por las decisiones de la IA a menudo carecen de la capacidad para comprender, cuestionar o resistirse a esos sistemas. Esto es especialmente cierto en el caso de las poblaciones marginadas que pueden verse sometidas a evaluaciones algorítmicas de bienestar, denegaciones automatizadas de visados o vigilancia digital, todo ello sin acceso a explicaciones, reparación o posibilidades de participación. Esto genera no solo injusticia epistémica, sino también desempoderamiento práctico, frustración y

eventual resignación: la negación de la propia capacidad para aprender de los sistemas que conforman la propia vida y responder a ellos.

Quizá lo más importante es que las IA pueden desalentar la agencia moral y la capacidad de hacer excepciones, ambas fundamentales para la capacidad de adaptación. Donde antes las personas confiaban en su propio juicio y sentido común, ahora se espera a menudo que cumplan protocolos algorítmicos, incluso cuando esos protocolos no se ajustan a la realidad de la situación vivida. Por ejemplo, en el sistema de justicia penal, los algoritmos de evaluación de riesgos pueden recomendar sentencias severas para los jóvenes desempleados de barrios marginados sin tener en cuenta las causas profundas de la delincuencia, como el desempleo ☒potencialmente vinculado a la automatización☒ o la ausencia de perspectivas significativas para el futuro. La capacidad de hacer excepciones, de reconocer el potencial de rehabilitación, de ofrecer segundas oportunidades y de sopesar factores no recogidos en los datos es crucial no solo para las perspectivas a largo plazo de la vida de los individuos, sino también para la cohesión social de las comunidades. Del mismo modo, en los sistemas de asistencia social, las comprobaciones automatizadas de elegibilidad pueden denegar la ayuda a las familias con grandes necesidades si sus circunstancias no coinciden con las categorías preestablecidas, dejando a los trabajadores de primera línea con poca discreción para anular la decisión. En ambos casos, la adhesión rígida a los resultados algorítmicos reduce el espacio para el juicio moral sensible al contexto y la resolución pragmática de problemas, las mismas cualidades que permiten a las sociedades adaptarse de forma constructiva a las circunstancias cambiantes a lo largo del tiempo.

La IA y el desmoronamiento de los entornos propicios: el deterioro de las instituciones democráticas

El último pilar de la resiliencia ☒el entorno propicio☒ incluye los sistemas jurídicos, las instituciones políticas y los marcos de gobernanza que gestionan el riesgo, protegen los derechos y garantizan la equidad. Estas instituciones desempeñan un papel fundamental en la estructuración de las relaciones, la mediación en los conflictos y la defensa de la dignidad y la libertad de todos los miembros de una comunidad. Sin

embargo, la IA se está desarrollando de formas que cada vez más superan, eluden o socavan estas estructuras.

“ En muchos contextos, el uso de la IA en la vigilancia, el control de fronteras o la supervisión de protestas erosiona las libertades civiles, especialmente de grupos vulnerables ”

La opacidad de muchos sistemas de IA a menudo protegidos por la ley de propiedad intelectual o el secreto comercial dificulta que las personas afectadas o incluso los reguladores comprendan cómo se toman las decisiones administrativas y de gobernanza. Esto menoscaba la rendición de cuentas y la justicia procesal. El hecho de que la IA a menudo obtenga conocimientos a partir de datos sesgados o incompletos agrava aún más este reto: cuando las injusticias sistémicas se codifican en lógicas algorítmicas, la injusticia se naturaliza y se hace invisible.

Además, el desarrollo y el despliegue de la IA están actualmente dominados por un puñado de poderosos actores corporativos y organismos estatales que operan a menudo a escala transnacional, con una supervisión democrática mínima. En muchos contextos, el uso de la IA en la vigilancia, el control de fronteras o la supervisión de protestas ha erosionado activamente las libertades civiles, especialmente en el caso de grupos ya de por sí vulnerables. Por ejemplo, en la región china de Xinjiang se han desplegado avanzados sistemas de reconocimiento facial para vigilar y detener a los musulmanes uigures, una clara violación de la libertad de movimiento y de culto. En Grecia, los drones, las cámaras y los sistemas basados en inteligencia artificial financiados por la UE vigilan los movimientos de los migrantes en las fronteras terrestres con Turquía, a menudo de forma poco transparente, lo que plantea graves problemas de derechos humanos en torno al desplazamiento forzoso y la criminalización de la migración. En los territorios palestinos ocupados, las autoridades israelíes han implementado sistemas de reconocimiento facial basados en la IA, como los programas «Lobo azul» y «Lobo rojo», para identificar y rastrear a los palestinos a

través de los puestos de control y en los espacios públicos. En otros lugares, durante las protestas en países como Rusia, India e Irán, se han utilizado herramientas de vigilancia basadas en la IA ☐ incluidos sistemas automatizados de identificación como el reconocimiento facial y de la forma de andar, que analiza los patrones de marcha de un individuo ☐ para identificar y seleccionar a los manifestantes, vulnerando gravemente los derechos a la libertad de expresión y de reunión.

El uso de sistemas algorítmicos para vigilar, perfilar o seleccionar grupos sin contexto o recursos adecuados socava tanto los derechos humanos como la resiliencia social. A medida que el entorno flaquea bajo estas presiones, las redes que antaño sostenían la interacción pacífica se vuelven cada vez más propensas a la ruptura. La confianza en las instituciones democráticas disminuye. El recurso a la reparación se vuelve difícil de alcanzar. Las personas pierden la confianza en que el sistema pueda protegerlas y, en ausencia de vías pacíficas para el cambio, quizás recurran en su lugar a la resistencia, el retraimiento o la violencia.

Conclusión: restablecer la paz en la era de la IA

Aunque no son intrínsecamente violentos, los sistemas de IA pueden erosionar las propias condiciones ☐ recursos, capital social, capacidad de adaptación y entornos propicios ☐ que hacen que las sociedades sean resilientes a los conflictos y las crisis. Al interrumpir el acceso a los recursos, fragmentar las redes, limitar la capacidad de resolución de problemas y debilitar las instituciones democráticas, la IA corre el riesgo de deteriorar el tejido relacional que mantiene unidas a las comunidades. Estos riesgos rara vez se dan a conocer por medio de sacudidas repentinas; surgen silenciosamente mediante la desconexión, el desempoderamiento y la deshumanización. Pero no son inevitables.

La IA no es una fuerza única y determinista. Sus repercusiones dependen de los fines a los que sirve, de los valores integrados en su diseño y de los contextos en los que se implementa. Como demuestran los siguientes ejemplos positivos, la IA también puede utilizarse a lo largo de los diferentes pilares de la resiliencia para promover la paz en lugar de minarla:

Acceso equitativo a los recursos: En contextos en los que las tensiones entre comunidades son mayores ☒ como, por ejemplo, entre pastores y agricultores que compiten por unas tierras escasas ☒ las herramientas de previsión climática y planificación del uso de la tierra impulsadas por la IA pueden ayudar a anticipar las sequías, optimizar las rotaciones de pastoreo y reducir las disputas por los recursos. En Kenia, por ejemplo, plataformas como *Virtual Agronomist* y *PlantVillage* ofrecen asesoramiento agrícola localizado en varios idiomas, lo que permite a los pequeños agricultores mejorar sus rendimientos.

Creación de capital social en todas las comunidades: Las plataformas de diálogo asistidas por IA, como *Pol.is*, que se ha utilizado en Taiwán, Canadá, Singapur, Filipinas, Finlandia, España y otros países, han demostrado cómo la tecnología puede superar las divisiones trazando áreas de acuerdo entre grupos grandes y diversos. En contextos profundamente polarizados, estas herramientas pueden reforzar el capital social puente fomentando las conexiones entre personas que, de otro modo, nunca llegarían a conocerse mediante líneas étnicas, religiosas o socioeconómicas.

Mejorar la capacidad de adaptación mediante una planificación inclusiva: En ciudades brasileñas como Porto Alegre y Belo Horizonte, una elaboración de presupuestos participativa mejorada con IA permite a los residentes ☒ incluidos los de zonas marginadas ☒ proponer, debatir y votar las prioridades de inversión pública. Esto no solo refuerza la capacidad local de resolución de problemas, sino que además genera la confianza y las aptitudes necesarias para que las comunidades se adapten a los retos futuros.

Reforzar los entornos propicios mediante la democracia digital: Plataformas cívicas como *Decidim* en Barcelona ofrecen sistemas transparentes y de código abierto para que los ciudadanos puedan crear conjuntamente políticas, elaborar propuestas y pedir cuentas a las instituciones. Estas herramientas, cuando se gestionan correctamente y de forma inclusiva, pueden revitalizar la confianza en la gobernanza y permitir formas de democracia más directas y participativas.

“ La tarea no es simplemente hacer que los sistemas de IA sean más seguros y eficientes, sino de afianzarlos en una visión de paz que dé prioridad a la dignidad humana y a la libertad, especialmente de los más vulnerables ”

Alineando el uso de la IA con los principios que salvaguardan la dignidad humana y la libertad, podemos ayudar a transformar las redes hacia la confianza en lugar de la desconfianza, y hacia la cooperación en lugar de la polarización. De este modo, la IA podría convertirse no en una amenaza para la paz sostenible, sino en algo que contribuya activamente a la misma. Por lo tanto, la tarea no es simplemente hacer que los sistemas de IA sean más seguros, robustos y eficientes. Se trata de volver a afianzarlos en una visión de paz: una visión que dé prioridad a la dignidad humana y a la libertad, especialmente de los más vulnerables. Esto puede incluir enfoques de diseño ético y cocreación, pero también definir espacios libres de IA que estén reservados sólo a los humanos. En cualquier caso, dos marcos éticos resultan indispensables:

En primer lugar, la ética del cuidado nos recuerda que la paz no se mantiene mediante el control, sino mediante la atención a la necesidad, la vulnerabilidad y el contexto. Los enfoques orientados a los cuidados fomentan el diseño de sistemas inclusivos, receptivos y basados en las relaciones humanas. Nos invitan a valorar no solamente los resultados, sino la calidad de las interacciones que nos llevan hasta ellos.

En segundo lugar, un enfoque basado en los derechos humanos garantiza que la paz no se deje únicamente en manos de una intención benévola. Los marcos de derechos integran la dignidad y la libertad en las normas jurídicas y las prácticas institucionales, protegiendo a las personas de los abusos sistémicos y permitiendo una participación significativa en la configuración del futuro tecnológico. Proporcionan barreras de protección estructurales, tales como la transparencia, la responsabilidad y la no discriminación, que son esenciales para preservar la paz en condiciones de cambios rápidos y crisis.

En este sentido, la paz en la era de la IA no es solo un problema técnico o normativo. Es un proyecto moral y político. Requiere replantearse no solo las herramientas que construimos, sino los valores que codificamos y los futuros que hacemos posibles. Si queremos preservar la paz en una era de máquinas inteligentes, debemos empezar por defender y diseñar para lo que nos hace humanos.

[1] Galtung, J. 1969. "Violence, Peace and Peace Research." *Journal of Peace Research* 6 (3): 167–91.

[2] Arendt, H. 2006 [1963]. *Eichmann in Jerusalem: A Report on the Banality of Evil*. Penguin Publishing Group.

[3] Sen, A. 2014 [1999]. "Development as Freedom." A *The Globalization and Development Reader: Perspectives on Development and Global Change*, 525.

[4] Couldry, N., y U. A. Mejias. 2019. "The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism." A *The Costs of Connection*. Stanford University Press.

Esta es una versión traducida del artículo publicado originalmente en [inglés](#).

Sobre la autora

Evelyne Tauchnitz

Investigadora sénior y profesora en el Instituto de Ética Social, Universidad de Lucerna, Suiza. Su experiencia se sitúa en la intersección entre la ética, la tecnología, la IA, la paz y los derechos humanos. Ejerce como vicepresidenta del Panel sobre IA y Construcción de la Paz del Panel Internacional sobre el Entorno de la Información (IPIE), investigadora asociada en el Centro de Tecnología y Asuntos Globales (CTGA) en la Universidad de Oxford, y miembro sénior del Centro para la Innovación en la Gobernanza Internacional (CIGI). Evelyne también es miembro del Consejo de la Fundación Peacenexus y antigua

miembro del Grupo Asesor Multistakeholder (MAG) del Foro de Gobernanza de Internet de las Naciones Unidas. Tiene un doctorado en Relaciones Internacionales, con especialización en Ciencia Política, por el Instituto de Posgrado de Estudios Internacionales y de Desarrollo (IHEID) de Ginebra.

Fotografía

Representación visual del concepto de inteligencia artificial y conexión digital. Autor: Pixel-Shot (Shutterstock).

ARTÍCULOS CENTRALES

Algoritmos puente: herramientas para despolarizar

Luke Thorburn

Investigador de la AI & Democracy Foundation

Es ampliamente conocido que, al optimizar el *engagement*, los algoritmos de las redes sociales tienden a dar una visibilidad desproporcionada a las voces más extremas e incendiarias^[1], lo que eleva la temperatura emocional de las conversaciones en línea y contribuye a la polarización política^[2]. Pero ¿podríamos diseñar algoritmos de clasificación que ayuden a crear puntos en común? Esta es la motivación que subyace a un área emergente de investigación y de orden práctico en torno a los «algoritmos puente» y la «clasificación basada en puentes». En este artículo, ofrezco una introducción concisa a los algoritmos puente, te oriento sobre dónde y cómo han sido utilizados hasta ahora, tanto por organizaciones de la sociedad civil como por plataformas de redes sociales, y resumo sus limitaciones y retos actuales.

El término «clasificación basada en puentes» puede describir cualquier método de clasificación de alternativas —ya sean publicaciones en redes sociales, propuestas políticas, respuestas a encuestas o candidatos a cargos electos— que ayude a crear entendimiento mutuo y confianza entre las partes enfrentadas^[3]. Este objetivo cualitativo puede ponerse en práctica de muchas maneras diferentes. Explicaré los enfoques más comunes, pero para ayudar a entenderlo, imagina que estás convocando — en la comunidad o región que mejor conozcas — un proceso en línea en el que los miembros de grupos que han estado en conflicto entre sí presentan sus propuestas por escrito sobre lo que debería suceder a continuación. Como persona facilitadora del proceso debes determinar las propuestas — que podrían ser miles — que tendrán mayor visibilidad cuando comuniqués los resultados a la comunidad, y debes hacerlo de una

manera que sirva de puente, es decir, que ayude a fomentar el entendimiento y la confianza mutuos.

“ Los algoritmos de las redes sociales tienden a dar una visibilidad desproporcionada a las voces más incendiarias. ¿Podríamos diseñar algoritmos de clasificación que ayuden a crear puntos en común? ”

El enfoque más común consiste en identificar las propuestas que representan alguna forma de terreno común o «aprobación diversa», es decir, aquellas que cuentan con la aprobación de personas que normalmente no están de acuerdo entre sí. Si se dispone de datos sobre los votos de las personas de la comunidad acerca de si apoyan o no cada propuesta, es posible cuantificar el nivel de apoyo a cada propuesta dentro de cada grupo y, a continuación, identificar aquellas propuestas que cuentan con un alto nivel de apoyo en varios grupos. Encontrar esas propuestas y darles visibilidad tiene — al menos — dos efectos importantes.

En primer lugar, demuestra que, de hecho, sí existe un terreno común, a menudo bastante amplio, que no suele ser frecuentemente aceptado en situaciones de conflicto. En segundo lugar, dirigir la atención hacia las ideas que se encuentran en la intersección entre diferentes grupos, incluso si dichas ideas no cuentan con un apoyo universal, ayuda a reducir la clasificación^[4], es decir, la homogeneidad de los grupos de opinión, que es un conocido factor de riesgo para la escalada de conflicto. Veamos algunos ejemplos:

Este enfoque ha sido utilizado recientemente por la Alianza para la Paz en Oriente Próximo (ALLMEP) en una serie de diálogos en línea^[5] destinados a ayudar a crear un terreno común entre constructores de paz israelíes y palestinos, lo que dio lugar a un conjunto de cinco demandas para los líderes mundiales y la comunidad internacional, las cuales contaron con un apoyo superior al 90% en cada bando.

También se ha implementado a gran escala en las Notas de la Comunidad una función de la plataforma de redes sociales X (anteriormente Twitter) en la que los usuarios pueden proponer notas que añaden un contexto importante a los tuits que consideran engañosos. Estas notas, pero, solo se muestran públicamente si un algoritmo determina que es probable que sean calificadas como útiles por personas de ambos extremos del espectro político. Este algoritmo es de código abierto y la función también se está probando en las plataformas de Meta (Facebook, Instagram, Threads), TikTok y YouTube.

Otro ejemplo es Polis, una herramienta de código abierto destinada a facilitar el diálogo colectivo en línea que se ha utilizado bastante en contextos de la sociedad civil y de consolidación de la paz, hace que este tipo de clasificación sea muy accesible. En el informe estándar de una conversación de Polis, basta con desplazarse hacia abajo hasta la sección «Todas las declaraciones» y seleccionar la opción de ordenar las declaraciones por «consenso grupal».

“ La clasificación basada en puentes puede describir cualquier método para clasificar alternativas – publicaciones en redes o propuestas políticas- que ayude a construir entendimiento mutuo y confianza entre las partes enfrentadas ”

Un segundo enfoque consiste en intentar reflejar un conjunto de propuestas que sean representativas de las opiniones de la comunidad, garantizando que se incluyan propuestas diversas, pero también que haya suficiente contexto sobre el grado de apoyo que tienen esas propuestas y de dónde proviene ese apoyo. Dicha representación y dicho contexto social son importantes porque, en contextos polarizados, suelen existir diferencias de percepción^[6]: tendemos a pensar que la otra parte tiene opiniones más extremas o intolerables de lo que realmente son. El contexto social necesario para desmontar estas falsas percepciones también suele faltar en las redes

sociales,^[7] que actúan como un «prisma^[8]» al dar una visibilidad desproporcionada a las perspectivas más incendiarias, y solo proporcionan recuentos brutos de «me gusta», «compartir», etc., sin información sobre el grado en que una publicación resuena en la población en general, o en qué subcomunidades ese sentimiento está más presente. Por ello, puede ser importante que los procesos gestionados por la sociedad civil contribuyan a proporcionar ese contexto social.

Este enfoque también se aplica en Polis, la herramienta mencionada anteriormente, que en su informe estándar proporciona tanto una visualización de los grupos de opinión más destacados como un informe sobre el grado de acuerdo de cada grupo con cada afirmación.

Esta visión es también una de las razones fundamentales que justifican el uso de encuestas de opinión pública más tradicionales en los procesos de paz^[9], como las realizadas por Colin Irwin^[10] en Irlanda del Norte, Cachemira, los Balcanes, Sri Lanka y otros lugares, cuyo objetivo, en parte, es ayudar a las personas a darse cuenta de que hay mucha más gente que comparte sus opiniones de lo que creen; es decir, reducir la «ignorancia pluralista».

“ Los algoritmos puente son parte de un ecosistema emergente de «herramientas deliberativas» en línea que pueden facilitar el diálogo en contextos polarizados ”

Un tercer tipo de enfoque consiste en utilizar clasificadores automatizados para identificar y mejorar la clasificación de las propuestas que presentan características evidentes de haber sido redactadas de buena fe. Estos clasificadores suelen ser más «opacos» que los dos enfoques anteriores, lo que significa que no siempre es posible saber por qué el clasificador ha puntuado un texto determinado de una determinada manera. Por ello, el uso de este tipo de clasificadores podría ser menos adecuado en aquellos procesos en los que la transparencia y la equidad procedimental resultan fundamentales. Sin embargo, pueden ser útiles en foros en línea más informales, por

ejemplo, para animar a las personas a reflexionar sobre cómo su contribución puede ser recibida por los demás y para empoderarlas para que filtren los tipos de contribuciones que consideran valiosas.

Como ejemplo, Google Jigsaw ha creado clasificadores automatizados, de uso gratuito, que pueden utilizarse para puntuar textos (por ejemplo, comentarios en un foro en línea) en función del grado en que muestran «atributos puente», como la compasión, la curiosidad, los matices o el respeto.

Desafíos y limitaciones

Cada uno de estos enfoques tiene sus limitaciones, y el diseño de algoritmos puente es un área de investigación muy activa. No obstante, existen tres retos que es importante tener en cuenta.

En primer lugar, la clasificación de los elementos de contenido es intrínsecamente de suma cero, y la optimización de cualquier aspecto — ya sea el terreno común, la representatividad, los matices o cualquier otra cosa — implicará que uno se enfrente a concesiones respecto de otros objetivos que tiene como facilitador del proceso. A modo de ejemplo, consideremos el objetivo del terreno común. Es mucho más fácil que las personas se pongan de acuerdo en vaguedades sobre la creación de un futuro más armonioso que, por ejemplo, en cláusulas redactadas de forma precisa que podrían incluirse en un acuerdo de paz sustantivo. Y muchas personas del ámbito de la construcción de paz serán conscientes de la «jugada» diplomática por la que, para permitir la continuidad del diálogo a pesar de un desacuerdo irresoluble, se modifica el lenguaje a fin de que sea cada vez más abstracto o procedimental. El resultado, por ejemplo, no es un acuerdo sobre cómo se distribuirán los beneficios y las cargas, sino un acuerdo que tratará meramente sobre el hecho de que deben distribuirse «de manera justa» o el que se seguirá un determinado proceso para deliberar acerca del modo en el que se procederá a su distribución.

“ Para avanzar con el diálogo se modifica el lenguaje a fin de que sea cada más abstracto; el objetivo de maximizar los puntos en común entra en tensión con el objetivo de minimizar la ambigüedad. ”

Aunque a veces sea la única forma posible de avanzar, esa vaguedad no hace sino posponer el problema – el desacuerdo que inevitablemente habrá que resolver – al jugar con la ambigüedad del lenguaje – y, por tanto, el fondo – de lo que se acuerda. De este modo, el objetivo de maximizar los puntos en común entra en tensión con el objetivo de minimizar la ambigüedad, por lo que la clasificación por aprobación diversa puede dar lugar a que las declaraciones relativamente vagas o ambiguas ocupen los primeros puestos. Esto se puede apreciar en los resultados del proceso ALLMEP mencionado anteriormente, en el que algunos de los «valores» y «visiones» más integradores (puente) para el futuro tienen este carácter. En relación con esta cuestión, está en curso un proyecto, en el que colaboro, que busca desarrollar un método que permita identificar declaraciones, a la vez integradoras (puente) e inequívocas.

En segundo lugar, el hecho de que se pueda utilizar un algoritmo para reconocer puntos en común o matices no garantiza que se vayan a encontrar muchos. Las aplicaciones con más éxito en estos métodos se han dado en contextos en los que también se realizó un esfuerzo considerable para crear las condiciones en las que las personas participan de buena fe y aportan ideas que consideran que pueden tener una amplia resonancia. Una limitación de las Notas de la Comunidad^[11] ampliamente conocida, por ejemplo, es que solo un porcentaje relativamente pequeño – menos del 10% – de todas las notas que se escriben se consideran puente -y, por lo tanto, se muestran públicamente-. Su eficacia como control de falsedades en la esfera pública se ve, por tanto, significativamente limitada por la reflexión de los colaboradores de las notas, pero es difícil crear una cultura de este tipo a la escala de una plataforma moderna de redes sociales.

Por el contrario, cuando los algoritmos puente se utilizan en procesos con una incorporación y una creación de confianza adicionales, con tiempo para la deliberación y en comunidades lo suficientemente pequeñas como para que las personas tengan un sentido compartido de la responsabilidad — por ejemplo, porque viven en el mismo lugar y reconocen que tienen que construir alguna forma de vida en común juntos —, es probable que tengan más éxito, debido a que habrá más ideas puente que encontrar.

En tercer lugar, es importante reconocer que la accesibilidad de la clasificación basada en puentes como herramienta para facilitadores y constructores de paz aún no es la que debería ser. Quizás la implementación más accesible sea la herramienta de código abierto Polis. Y las personas familiarizadas con utilizar una hoja de cálculo, también verán sencillo utilizar la clasificación basada en puentes para seleccionar los resultados de los procesos ejecutados en Remesh — la herramienta utilizada por ALLMEP — o en las votaciones recopiladas mediante cualquier herramienta de encuesta estándar. Los algoritmos puente son solo una de las características que suelen estar presentes en un ecosistema emergente de «herramientas deliberativas» en línea que pueden facilitar la deliberación en contextos polarizados.^[12]

“ El impacto de los algoritmos puente probablemente provenga de su uso estratégico por parte de organizaciones de construcción de paz y de la sociedad civil en contextos polarizados ”

¿Y ahora qué?

Los algoritmos puente siguen siendo un área de investigación activa^[13]. En el contexto de las redes sociales, un amplio estudio académico ha probado recientemente varios algoritmos puente con aproximadamente 6.000 participantes, durante cuatro meses, en múltiples plataformas (Facebook, X y Reddit), y pronto sabremos si ha tenido éxito en la reducción de la polarización afectiva. Un estudio reciente más pequeño^[14], que

probó una forma de algoritmo puente en X, ya ha demostrado que es posible. Pero, aunque gran parte de la atención se ha centrado en las redes sociales, el impacto de los algoritmos puente – al menos a corto plazo – probablemente provenga de su uso estratégico por parte de organizaciones de construcción de paz y de la sociedad civil en contextos polarizados a una escala más local o regional.

*Si estás interesado en utilizar estos métodos puedes contactar con el autor o con el centro Deliberative Tech del [Consejo de Tecnología y Cohesión Social](#), una comunidad de práctica emergente para organizaciones de la sociedad civil que utilizan estas herramientas.

[1] Smitha Milli, Micah Carroll, Yike Wang, Sashrika Pandey, Sebastian Zhao y Anca D. Dragan. Marzo de 2025. «[Engagement, user satisfaction, and the amplification of divisive content on social media](#)». En PNAS Nexus, volumen 4, número 3.

[2] Tiziano Piccardi, Martin Saveski, Chenyan Jia, Jeffrey T. Hancock, Jeanne L. Tsai y Michael Bernstein. 2024. «[Social Media Algorithms Can Shape Affective Polarization via Exposure to Antidemocratic Attitudes and Partisan Animosity](#)». Preprint.

[3] Aviv Ovadya y Luke Thorburn. 2023. «[Bridging Systems: Open problems for countering destructive divisiveness across ranking, recommenders, and governance](#)», ArXiv.

[4] Luke Thorburn, Maria Polukarov y Carmine Ventre. 2024. «[Societal Sorting as a Systemic Risk of Recommenders](#)». En la 18.a Conferencia de la ACM sobre Sistemas de Recomendación (RecSys '24), Nueva York: Association for Computing Machinery, 2024, p. 951-956.

[5] Andrew Konya, Luke Thorburn, Wasim Almasri, Oded Adomi Leshem, Ariel Procaccia, Lisa Schirch y Michiel Bakker. 2025. «[Using collective dialogues and AI to find common ground between Israeli and Palestinian peacebuilders](#)». En la Conferencia de 2025 de la ACM sobre equidad, rendición de cuentas y transparencia (FAcCT '25), Nueva York: Association for Computing Machinery, 2025, p. 312-333.

[6] Jonathan Stray. 2025. «[Why Do They Think We're Extreme?](#)». En el *Better Conflict Bulletin*.

[7] E. Glen Weyl, Luke Thorburn, Emillie de Keulenaar, Jacob Mchangama, Divya Siddarth y Audrey Tang. 2025. «[Prosocial Media](#)». ArXiv.

[8] Chris Bail. 2021. *Breaking the Social Media Prism: How to Make Our Platforms Less Polarizing*. Prince University Press.

[9] Colin Irwin. 2025. «[The People's Peace Third AI and UN Edition. Public Opinion, Public Diplomacy and World Peace](#)». *Peace Polls Publication*.

[10] Ver la página web de [Peace Polls](#).

[11] Madison Czopek. 2023. «[Why Community Notes mostly fails to combat misinformation](#)». En *Poynter*.

[12] La AI & Democracy Foundation facilita el acceso gratuito a probar herramientas de deliberación a través de su [Deliberative Tools Access Program](#).

[13] Madon Revel, Smitha Milli, Tyler Lu, Jamelle Watson-Daniels y Max Nivkel. 2025. «[Representative Ranking for Deliberation in the Public Sphere](#)», ArXiv.

[14] Tiziano Piccardi, Martin Saaveski, Chenyan Jia, Jeffrey T. Hancock, Jeanne L. Tsai y Michael Bernstein. 2024. «[Social Media Algorithms Can Shape Affective Polarization via Exposure to Antidemocratic Attitudes and Partisan Animosity](#)», ArXiv.

Esta es una versión traducida del artículo publicado originalmente en [inglés](#).

Sobre el autor

Luke Thorburn

Investigador de l'AI & Democracy Foundation i estudiant d'últim any del doctorat en informàtica al King's College de Londres. Actualment, la seva recerca se centra en el

disseny de sistemes de recomanació de les xarxes socials que permetin reduir els riscos de conflicte, un concepte conegut també com a «classificació basada en ponts». És coautor del projecte «Understanding Recommenders» del Center for Human-Compatible AI de la Universitat de Califòrnia a Berkeley, membre de la xarxa de recerca GETTING-Plurality de l'Allen Lab for Democracy Renovation de la Universitat de Harvard i investigador afiliat del Machine Intelligence and Normative Theory Lab de la Universitat Nacional d'Austràlia. Thorburn també ha assessorat Ofcom, la newDemocracy Foundation i Meta. Té formació en probabilitat i estadística.

Fotografía

Representación simbólica de los algoritmos y su papel en la polarización en el ámbito digital. Author: Achira22 (Shutterstock).

ARTÍCULOS CENTRALES

La mediación en la era de los algoritmos: riesgos y oportunidades para los procesos de paz

Sanjana Hattotuwa

Doctor en Políticas y Redes Sociales (Universidad de Otago)

«La mediación en línea tiene un enorme potencial para ayudar a resolver conflictos.

En la región de Asia-Pacífico, ocupa un lugar especial al facilitar el acceso a la justicia a grandes poblaciones que tienen un acceso limitado a la resolución de disputas por otros medios».[1] Esta observación, realizada hace dos décadas, sigue siendo profundamente relevante hoy en día. Mi argumento entonces era que un paradigma de resolución de conflictos centrado en Occidente y basado en la tecnología era inadecuado para el contexto asiático. Mucho antes de la omnipresencia de las redes sociales y los teléfonos inteligentes, nos centramos en desarrollar arquitecturas de resolución de conflictos en línea que aprovecharan tecnologías incipientes como la telefonía móvil y la radio comunitaria por Internet.

Un buen ejemplo de ello fue en 2002, como parte integral de las negociaciones oficiales de alto el fuego en Sri Lanka. En este caso, de forma inédita, se utilizó un software comercial especialmente adaptado (COTS), llamado *Groove Virtual Office*, para apoyar los procesos de mediación activa, basados en un proceso de un *One-Text*[2]. Esta plataforma, que codiseñé y de la que dirigí el desarrollo de herramientas de apoyo a las mediaciones, permitía comunicaciones cifradas, asíncronas y multilingües, creando repositorios de documentación con función de búsqueda, herramientas de apoyo a la toma de decisiones y mapeo de posiciones de múltiples partes interesadas. Fundamentalmente, se diseñó para integrar las aportaciones, no solo de las principales partes negociadoras, sino también de la diáspora y de los participantes en la Vía 2 y Vía

3[3] de los procesos de paz, lo que demuestra una práctica arraigada de utilizar las tecnologías de la información para mediaciones complejas lejos del Norte Global.

En 2025, dos décadas después del proceso de *One-Text* en Sri Lanka, el papel y la relevancia de la tecnología en la consolidación de la paz están mucho más establecidos. Sin embargo, los años transcurridos revelan un reto fundamental para las personas mediadoras. Las mismas plataformas que ofrecen oportunidades sin precedentes para el diálogo, también sirven como vectores de dinámicas saboteadoras que pueden hacer descarrilar las frágiles conversaciones de paz. Este dilema es grave porque estos nuevos vectores de producción de información se encuentran totalmente fuera del ámbito de las reglas de Chatham House y del entorno controlado necesario para fomentar la confianza. Ahora, los mediadores se ven bombardeados con información, y sus procesos, cuidadosamente gestionados, se enfrentan a la amenaza constante de campañas en línea de desinformación e información errónea diseñadas para avivar el odio, difundir falsedades incendiarias y erosionar la confianza del público en un posible acuerdo. Este efecto de «contagio emocional», en el que el sentimiento en línea influye directamente en cómo se siente la gente fuera de línea, pone en evidencia que la seguridad de una sala de negociaciones a puerta cerrada ya no es suficiente para proteger la integridad del proceso. ¿Qué marcos, de haberlos, pueden conciliar la capacidad de la tecnología para arruinar y reconstruir la confianza en el contexto específico de una negociación mediada?

“ Las mismas plataformas tecnológicas que ofrecen oportunidades para el diálogo también también sirven como vectores de dinámicas saboteadoras que pueden hacer descarrilar las frágiles conversaciones de paz ”

Lo que salta a la vista de inmediato es el lado oscuro (o más oscuro) de la tecnología y las redes sociales, que suponen una amenaza directa y creciente no solo para la cohesión social en general, sino también para los propios mecanismos de la mediación

para la paz. La amplificación algorítmica de contenidos incendiarios en las redes sociales puede endurecer las posiciones de las partes negociadoras y sus electores, haciendo que el compromiso, que es la piedra angular de la mediación, sea políticamente insostenible. La difusión de desinformación puede utilizarse como arma para socavar la credibilidad de los mediadores, descarrilar conversaciones o violar el «entorno controlado» esencial para negociaciones frágiles. Para un mediador, esto supone un reto sin precedentes: los aspectos clave del entorno del conflicto están siendo moldeados en tiempo real por fuerzas que escapan por completo al ámbito de aplicación de las reglas de Chatham House o de los códigos de conducta establecidos. La dinámica de los sabotadores ya no se limita a actos físicos de violencia, sino que se manifiesta en forma de campañas virales de odio y falsedad que pueden deshacer los delicados avances logrados en la mesa de negociaciones.

Por el contrario, y aunque se destaca con menos frecuencia, el potencial transformador de la tecnología en la mediación es más profundo cuando va más allá de los procesos elitistas y empodera a las comunidades de base. Tal y como concebí en 2006, el futuro de la Resolución de Conflictos en Línea (*Online Dispute Resolution*, ODR) en el Sur Global no residía en replicar los sistemas basados en ordenadores personales, sino en aprovechar las tecnologías que ya estaban en manos de la gente. El crecimiento explosivo de la telefonía móvil presentó una oportunidad para diseñar sistemas ODR con un «rostro humano». Entre las aplicaciones concretas para la mediación comunitaria se incluían el uso de SMS para enviar notificaciones en lengua vernácula de los acuerdos a las partes en conflicto; la posibilidad de que los equipos de mediación sobre el terreno recopilaran testimonios de audio y vídeo con sus teléfonos móviles; la creación de sistemas expertos que pudieran proporcionar a los mediadores opciones en tiempo real para resolver disputas comunes, como las relacionadas con la tierra; y el uso de videoconferencias móviles para conectar a las partes en zonas remotas con los centros de Resolución Alternativa de Disputas (*Alternative Dispute Resolution*, ADR). Este enfoque lleva la mediación a los arrozales, a la oficina de correos y a la residencia del jefe de la aldea, haciendo que pase de un proceso centralizado e inaccesible a un servicio omnipresente, fácil de usar y con resonancia cultural para la resolución no violenta de conflictos.

“ Los sabotajes a la mediación ya no se limitan a actos físicos de violencia, sino que se manifiesta en forma de campañas virales de odio y falsedad. La desinformación se puede utilizar como arma para socavar la credibilidad de los mediadores ”

Este dualismo refleja el trabajo de la profesora Miriyam Aouragh, que comparó las redes sociales con la espada de Damocles y argumentó «que aquellos que se empoderan al tomar asiento bajo la espada lo hacen acechados por la amenaza constante de ser heridos por la misma espada, porque la matanza podría producirse ante la más mínima perturbación»[4]. Como afirma el profesor Admire Mare, la tesis central de Aouragh es que «las redes sociales son una tecnología abierta y sin límites, lo que permite a los actores estatales y no estatales aprovecharlas para fines buenos y nefastos». En otras palabras, aunque es raro que se reconozca de este modo en los debates políticos, e incluso en la literatura académica del Norte Global, las redes sociales son a la vez buenas y malas, útiles y perjudiciales, conciliadoras y divisivas, pacíficas y violentas.

Esta doble naturaleza de las herramientas digitales se manifiesta directamente en los propios procesos de mediación. WhatsApp, por ejemplo, ha revolucionado la forma en que los mediadores se relacionan con las partes en conflicto: la Oficina del Enviado Especial de las Naciones Unidas para Yemen utiliza grupos de WhatsApp para mantener una comunicación en tiempo real con los negociadores[5], al tiempo que se enfrenta al problema de cómo la misma plataforma difunde desinformación sobre el proceso de paz. En Libia, la UNSMIL estableció relaciones de «socio de confianza» con Facebook para eliminar contenidos perjudiciales dirigidos a los miembros del Foro de Diálogo Político Libio, en particular a las mujeres participantes, al tiempo que utilizaba la misma plataforma para promover narrativas de paz[6]. Esta paradoja —en la que la misma aplicación, tecnología, plataforma, producto o herramienta que permite una consulta inclusiva puede convertirse en un arma para sabotear el proceso— exige que las personas mediadoras desarrollen estrategias digitales sofisticadas para aprovechar las oportunidades y mitigar los riesgos.

Más allá de las plataformas web tradicionales, han surgido enfoques innovadores de mediación digital en distintos contextos. En 2021, las consultas de Build Up a través de WhatsApp llegaron a 93 mujeres yemeníes de 11 provincias, creando espacios de diálogo allí donde las reuniones físicas eran imposibles[7]. El Espacio de Apoyo a la Sociedad Civil para Siria creó un sitio web interactivo que permite a los actores de la sociedad civil siria comunicar sus aportaciones directamente al Enviado Especial de las Naciones Unidas, salvando la brecha entre las negociaciones de paz de la Vía 1 y las perspectivas de las bases[8]. Estas iniciativas revelan cómo las herramientas digitales, cuando se diseñan con objetivos de mediación específicos, pueden superar las barreras tradicionales de la geografía, el género, la seguridad y el accesibilidad.

“ Los equipos de mediación deben desarrollar estrategias digitales sofisticadas para aprovechar las oportunidades de las plataformas tecnológicas y mitigar los riesgos que éstas conllevan ”

La evolución del diálogo en línea se extiende más allá de los procesos de paz formales a las sociedades digitales emergentes. En un artículo sobre el futuro de la Resolución de Conflictos en Línea[9], señalé las complejas transacciones sociales y comerciales que se producen en mundos virtuales como «Second Life». Estos entornos, con sus propias economías, derechos de propiedad y normas sociales, ya estaban generando nuevas disputas que se extendían al mundo real, incluido el primer asesinato inducido por un conflicto sobre un artefacto virtual. Esto planteó una pregunta crucial para el futuro de la mediación: ¿necesitamos sistemas de ODR diseñados explícitamente para conflictos que surgen y existen íntegramente en ámbitos virtuales? El «metaverso» actual se enfrenta a los mismos retos de gobernanza, daño y resolución. Esto demuestra que la necesidad de marcos de mediación innovadores, que abarquen diversos productos tecnológicos y plataformas, está en constante expansión, lo que nos obliga a tender puentes no solo entre el mundo físico, sino también en la interacción cada vez más compleja, fluida y dinámica entre los ámbitos online, virtual y real, donde se forman las

comunidades y surgen los conflictos.

Los mediadores de hoy en día reconocen cada vez más que una mediación digital eficaz requiere reunirse con las partes allí donde se comunican, ya sea a través de aplicaciones de mensajería cifrada, plataformas de redes sociales o interfaces móviles, en lugar de esperar un acceso universal a las plataformas web tradicionales. Sin embargo, las reformas de las plataformas son esenciales para apoyar este prometedor potencial. Los algoritmos deben priorizar el diálogo constructivo frente a las métricas de participación (*engagement*), con mecanismos que frenen la propagación viral de comentarios incendiarios y falsos durante períodos de crisis intensas y repentinas o en contextos definidos por déficits o retrocesos democráticos^[10]. La asignación de recursos debe lograr la paridad entre las operaciones del Norte y del Sur, contratando a hablantes de idiomas locales y expertos culturales, e invertir en sistemas de inteligencia artificial diseñados para contextos no anglófonos y no occidentales.

Las intervenciones dirigidas por la comunidad y a nivel comunitario requieren una inversión sostenida y el desarrollo de capacidades. Las y los constructores de paz tradicionales necesitan formación sobre la adopción de plataformas digitales y su adaptación de forma fundamentada, con perspectiva de género, sostenible, accesible y equitativa, al tiempo que debe reforzarse la constelación de marcos políticos, normativos, mediáticos y jurídicos nacionales en torno a la integridad de la información. De hecho, la alfabetización en medios digitales debería convertirse en una prioridad de la educación pública, desde la infancia hasta la edad adulta. Los grupos de mujeres y las organizaciones juveniles que utilizan las redes sociales para la paz requieren apoyo y recursos específicos, reconociendo su papel como poblaciones vulnerables y agentes innovadores de paz.

“ Los algoritmos deben priorizar el diálogo constructivo frente a las métricas de participación, *engagement*, con mecanismos que frenen la propagación viral de desinformación ”

El impacto de los medios sociales en la consolidación de la paz en los contextos de la Mayoría Global desafía las categorizaciones simplistas y binarias como beneficioso o perjudicial[11]. Como sostiene el profesor Mare, las plataformas representan «una tecnología abierta y sin cierre» que requiere una comprensión sofisticada de los contextos locales, las dinámicas de poder y las relaciones sociotécnicas[12]. Mi propia investigación doctoral en 2021 estableció que, en Sri Lanka, las redes sociales «contribuyeron simultáneamente al afianzamiento autoritario y a la resistencia a la erosión democrática», que «las diferentes motivaciones de los usuarios de Facebook y Twitter apoyaron simultáneamente marcos prosociales y violentos durante momentos de importantes disturbios fuera de línea», y que Facebook y Twitter «amplificaron simultáneamente el odio y produjeron contenidos prosociales, no violentos y conciliadores que pedían civismo, defendían las instituciones democráticas y celebraban la diversidad».

El camino a seguir para la mediación en la era digital requiere un replanteamiento fundamental del entorno en línea en el que se desarrollan los procesos de paz. Los modelos de negocio de las plataformas que se benefician de una «guerra de historias» socavan directamente la tarea fundamental del mediador de reducir la tensión, al amplificar las narrativas divisivas y acelerar la decadencia epistémica[13]. La solución no radica en modificar las plataformas perjudiciales, sino en diseñar arquitecturas de mediación digital a medida, basadas en una colaboración genuina con la sociedad civil local. Esto garantiza que quienes tienen experiencia en la transformación de conflictos sean fundamentales para crear herramientas que no solo sean técnicamente funcionales, sino que también tengan una base contextual. Para un mediador, es fundamental reconocer que el código no es neutral, sino que refleja sesgos inherentes que pueden apoyar el delicado equilibrio de una negociación o sabotearla por completo.

Para el mediador contemporáneo, la autoridad persuasiva, proyectada y percibida de las sofisticadas herramientas de Inteligencia Artificial (IA) debe ser recibida con un profundo escepticismo profesional. Su función no es suplantarse el oficio del mediador, sino aumentar su capacidad. Como demostró la plataforma One-Text durante el proceso de paz de Sri Lanka hace décadas, el verdadero valor de la tecnología reside, aparte de en su desarrollo endógeno, en su capacidad para ayudar a los mediadores a gestionar

grandes cantidades de información, facilitar el diálogo estructurado a través de las distancias y modelar escenarios complejos para mejorar el juicio humano. Debemos cuestionar constantemente los sesgos de las fuentes de datos y los análisis erróneos de la IA, recordando que ningún algoritmo puede replicar las tareas esenciales, dirigidas por humanos, de establecer relaciones, demostrar empatía y realizar juicios críticos y matizados en una mesa de negociación.

“ Debemos cuestionar c los sesgos de las fuentes de los análisis de la IA, recordando que ningún algoritmo puede replicar las tareas humanas de establecer relaciones, demostrar empatía y realizar juicios críticos en una mesa de negociación ”

En última instancia, hacer que las plataformas digitales pasen de escenarios de conflicto a herramientas prácticas para la mediación exige que nos centremos en la sabiduría de quienes negocian la paz a diario en los lugares más fracturados del mundo. Son las comunidades locales y los y las profesionales de la paz, incluidos los de las Primeras Naciones, los Adivasi y las comunidades indígenas, quienes comprenden las particularidades de la creación de confianza y la gestión de la narrativa, esenciales para cualquier mediación exitosa. Sus conocimientos, experiencia y sabiduría deben servir de base para el diseño de una nueva generación de entornos de mediación digital. Para ello es necesario abandonar la lógica extractiva, las plataformas rapaces y los esfuerzos mercenarios de Silicon Valley y, en su lugar, crear espacios en línea diseñados para apoyar el diálogo, la confidencialidad y el consenso, que son los pilares del proceso de mediación que los algoritmos actuales están incentivados a destruir.

Es decir, la tarea de la mediación moderna ya no se limita a una mesa de negociación física, sino que se extiende a los contenidos, comentarios y corrientes en línea que alimentan el conflicto. Por lo tanto, el éxito no se encontrará solo en la habilidad para navegar por los adversos bienes comunes digitales de hoy en día, sino en su

remodelación activa. El arte de la mediación, con su énfasis en la desescalada, la confidencialidad y el diálogo mesurado, debe ahora inspirar la lógica misma del código. Este es el próximo reto fronterizo: no adaptar los procesos de paz a los caprichos del algoritmo, sino adaptar el algoritmo a los principios perdurables de una paz justa.

[1] Hattotuwa, S., y M. C. Tyler. 2005. An Asian Perspective on Online Mediation. *Asian Journal on Mediation* 1 (1): 1-24. U of Melbourne Legal Studies Research Paper No. 158.

[2] ICT4Peace Foundation. 2013. The Janus Effect: Social Media in Peace Mediation. Zürich: ICT4Peace Foundation.

[3] Women's Peace and Humanitarian Fund. Definiciones de «proceso de paz», «vía 1» y «vía 2», y «implementación de un acuerdo de paz».

[4] Mare, A. 2024. Social Media, Conflict, and Peacebuilding in Southern Africa: A Primer. Kujenga Amani, Social Science Research Council. 17 de diciembre de 2024.

[5] UNITAR. 2021. 'WhatsApp Diplomacy': The Future of Multilateralism in a Post-COVID-19 World? United Nations Institute for Training and Research. 19 de mayo de 2021.

[6] Sustaining Peace Select. 2023. Platform Engagement.

[7] Build Up. 2023. Feminist Approaches to Online Consultations and What They Reveal. Blog de Build Up. 18 de mayo de 2023.

[8] Oficina del Enviado Especial de Naciones Unidas per Siria (OSE-Syria). 2016. Civil Society Support Room. 26 de septiembre de 2025.

[9] Hattotuwa, S. 2006. The Future of ODR: One Brief Glimpse. ICT for Peacebuilding (blog). 22 de febrero de 2006.

[10] Bunse, S. 2021. Social Media: A Tool for Peace or Conflict?. SIPRI Commentary. 20 de agosto de 2021.

[11] Reuss, A., y S. Stetter (eds.). 2023. Social Media and Peacebuilding: How Digital Spaces Shape Conflict and Peace. Palgrave Mcmillan.

[12] Ibíd.

[13] Hattotuwa, S. 2024. A 'War of Stories': Humanitarianism in the Disinformation Age. ICT for Peacebuilding (blog). 23 de diciembre de 2024.

Esta es una versión traducida del artículo publicado originalmente en inglés.

Sobre el autor

Sanjana Hattotuwa

Realizó su investigación doctoral sobre la intersección de las redes sociales, la comunicación política, la propaganda y los trastornos de la información en Sri Lanka, así como sobre cómo se representó en Twitter la masacre de Christchurch en Nueva Zelanda en marzo de 2019. Especializado y asesor sobre estrategia de comunicación en redes sociales, seguridad digital para periodistas y defensores de los derechos humanos, activismo en redes sociales, defensa en línea y fundamentada, investigación sobre redes sociales con métodos mixtos, la experiencia de Hattotuwa en el estudio, la negociación y el desarrollo de políticas contra los trastornos de la información abarca dos décadas, con trabajos en el sur de Asia, el sudeste asiático, el norte de África, Estados Unidos, Europa y los Balcanes.

Fundó en 2006 y hasta junio de 2020 dirigió Groundviews, el primer sitio web de medios cívicos de Sri Lanka, galardonado con varios premios. Desde 2021 hasta 2024 fue Director de Investigación en The Disinformation Project, con sede en Nueva Zelanda. Continúa siendo Asesor Especial en la ICT4Peace Foundation, con sede en Suiza.

Fotografía

Representación simbólica de la mediación en un contexto digital y algorítmico. Autor: Anbu-Creations (Shutterstock).

ARTÍCULOS CENTRALES

Democracia digital y ciberactivismo en África Oriental: el caso de Siasa Place

Nerima Wako

Directora ejecutiva de Siasa Place

En 2015, armados únicamente con nuestra pasión, el wifi de una cafetería del centro de la ciudad que por las noches se convertía en bar y la corazonada de que Twitter podría llegar a ser importante algún día, lanzamos Siasa Place, una organización cívica tecnológica dirigida por jóvenes con sede en Nairobi, Kenia. Nosotros, un grupo de jóvenes ambiciosos que, por alguna razón, coincidimos en la manera de ver la democracia, nos conocimos en diversos eventos por la ciudad; nos unió el conocimiento general sobre el país y la actualidad, así como la pasión por cambiar las cosas. Yo acababa de terminar mis estudios en el extranjero, así que cuando volví a casa tras haber estado siete años fuera, me alegró encontrar esta nueva comunidad, con la que conecté intrínseca e inmediatamente. Éramos patriotas y no teníamos dinero, sólo una sensación de urgencia. El espacio cívico se estaba reduciendo, el ambiente estaba cargado de desilusión y los jóvenes ansiaban un espacio para expresarse, ser escuchados y organizarse. Diez años después, esa apuesta por la democracia digital nos ha situado en la vanguardia del movimiento cibernético de África Oriental.

Por aquel entonces, era como predicar en el desierto. Organizábamos una tertulia semanal llamada #SiasaWednesday, en la que debatíamos sobre la asignación presupuestaria, los escándalos de corrupción, el liderazgo y la participación de los jóvenes en la gobernanza. Nos reuníamos virtualmente a una hora concreta y, a veces, teníamos la impresión de estar hablando solos.

Durante tres años, nos reunimos cada semana sin falta a la misma hora, para debatir diversos temas y movilizar a otras personas a fin de que se unieran a nosotros. Y, poco a poco, la gente se fue sumando. Los hilos fueron haciéndose más largos. Los seguidores se multiplicaron. ¿Te puedes creer que teníamos que escribir las conversaciones y que a veces nos quedábamos hablando de un tema tres horas seguidas? Esto fue antes de los espacios, cuando teníamos hilos de conversación. A a veces no tenías ni idea de quién estaba leyendo o las preguntas se quedaban a mitad de camino, mientras que otras preguntas llegaban días después, cuando alguien se topaba por casualidad con la conversación. Por aquel entonces tenías que responder por escrito. Hoy en día, damos por sentados los espacios X, en los que miles de personas pueden participar en debates en directo, hablar y escuchar, con tan sólo pulsar un botón, desafiando al poder en tiempo real. Incluso los vídeos en directo de TikTok, donde la gente puede comentar en tiempo real.

“ Internet no solo se ha convertido en el medio, sino también en el movimiento. Los *hashtags* son demandas; los *influencers* son analistas políticos; los memes son pancartas de protesta ”

Hoy en día, los memes son pancartas de protesta, graciosas, pero con posturas políticas y mensajes fuertes. Un lenguaje político inteligente y diferente con el que a los jóvenes les encanta interactuar. Los *hashtags* son demandas. Los *influencers* son analistas políticos. Internet no solo se ha convertido en el medio, sino también en el movimiento. Pero, a medida que nuestras voces iban ganando fuerza, también lo hacían las repercusiones.

La espada de doble filo del empoderamiento digital

Durante la última década, las plataformas digitales se han convertido en el escenario principal del compromiso cívico en África Oriental. De Nairobi a Kampala, de Dar es Salaam a Kigali, los y las jóvenes han creado poderosas comunidades en línea,

impulsando la transparencia, la rendición de cuentas y la justicia.

Siasa Place existe para fortalecer ese movimiento. Desarrollamos marcos institucionales, ofrecemos educación cívica y formamos a la próxima generación de líderes sobre cómo utilizar las herramientas tecnológicas para el compromiso político. Iniciativas como «Política, tecnología y derechos» abordan los matices del trabajo digital, la moderación de contenidos y la violencia de género en los espacios en línea. También apoyamos plataformas como zKE, que recogen la opinión de la ciudadanía, esencialmente de los jóvenes, sobre los proyectos de ley, solicitando sugerencias o recomendaciones acerca de dichos proyectos de una manera más ágil y eficaz, porque recabar la opinión pública forma parte de nuestra constitución. La manera en que se envían los comentarios cambiará considerablemente, especialmente a medida que las generaciones más jóvenes vayan acostumbrándose y familiarizándose con el uso de las plataformas digitales.

Pero las mismas herramientas que utilizamos para organizarnos se están utilizando ahora en nuestra contra. En Kenia, el Parlamento ha estado trabajando recientemente en la aprobación de un controvertido proyecto de ley sobre las redes sociales que otorgaría a las autoridades acceso a los dispositivos de la ciudadanía. Bajo el pretexto de la regulación, esto allana el camino para el control y la represión. En Uganda y Tanzania se han cerrado plataformas enteras durante las elecciones. Etiopía sufrió un apagón total de Internet en 2023. En la región, la democracia digital no solo es desigual, sino que está bajo asedio.

“ Las plataformas digitales se han convertido en el escenario principal del compromiso cívico en África Oriental. Los jóvenes han creado poderosas comunidades en línea en favor de la transparencia, la rendición de cuentas y la justicia ”

La democracia digital en África Oriental

Hablar de democracia digital en África Oriental es hablar de una contradicción. La conectividad es desigual. El acceso es costoso. Kenia dispone de Internet de alta velocidad, pero, a pesar de que nuestro alcance es cada vez mayor, la conectividad sigue siendo problemática, pues en las zonas muy rurales sigue siendo difícil acceder a Internet. En términos de asequibilidad, los paquetes de datos resultan caros para el ciudadano medio. Incluso tener un teléfono inteligente es caro. Esto hace que, si un hogar puede permitirse adquirir un teléfono inteligente, éste suele pertenecer al hombre de la casa. Eso también se traduce en la predominancia masculina de los usuarios; la mayoría de los usuarios en línea son hombres, no mujeres. La violencia en línea ya es elevada, pero la violencia contra las mujeres asistida por la tecnología es aún mayor. Cuando se es una minoría, las posibilidades de sufrir acoso en línea son elevadas, por lo que muchas personas reducen su participación en algunos temas, especialmente en asuntos de orden político, que suelen despertar bastantes pasiones.

También hay que tener en cuenta que Kenia es uno de los países más activos del mundo, de hecho, se encuentra entre los cinco primeros en cuanto al número de horas que los ciudadanos pasan en línea. Sin embargo, existe una contradicción: en los países donde los medios de comunicación tradicionales están dominados por propietarios afines a determinados partidos políticos, Internet se convierte en el único espacio sin filtros para la disidencia. Esto explica el crecimiento de los medios alternativos, desde los podcasters hasta los YouTubers. No obstante, ni siquiera en este caso podemos confiar únicamente en el mundo online.

**“ La democracia digital implica crear sistemas
alternativos de confianza y rendición de cuentas,
porque las instituciones nos fallan con
demasiada frecuencia ”**

La democracia digital significa combinar las sesiones en directo de Facebook y los tuits con los ayuntamientos. Significa comprender que el hecho que algo sea legal no significa que sea justo. Significa crear sistemas alternativos de confianza y rendición de cuentas, porque las instituciones nos fallan con demasiada frecuencia. Quién iba a pensar que seríamos testigos de la suplantación de ejércitos, como el de Irán, por ejemplo, que publicó que estaba atacando a Israel en defensa propia en la plataforma X, y en cuestión de horas tenía más de 100.000 visitas. Internet ha hecho que instituciones misteriosas, a menudo protegidas por velos de autoridad, parezcan más accesibles y, en ocasiones, más humanas. La ley puede doblegarse para servir a la represión; Internet no debe hacerlo.

Historias de resistencia y riesgo

En Siasa Place hemos apoyado campañas que exigían la liberación de presos políticos en países como Mozambique. Hemos trabajado con Africtiviste para hacer visible la represión en Senegal. A nivel local, hemos visto cómo las herramientas digitales permiten a los jóvenes responder rápidamente movilizándose contra injusticias como la brutalidad policial o la corrupción. Hemos mantenido conversaciones con organizaciones similares como Yiaga, en Nigeria, en las que hemos comparado nuestras elecciones y conocido de primera mano los procesos, hemos hablado entre nosotras, hemos colaborado y elaborado informes que puedan ser útiles para futuras recomendaciones.

Pero también hemos podido ver los riesgos. Los trolls patrocinados por el gobierno difunden desinformación para desacreditar a los activistas y esta técnica se ha vuelto cada vez más visible y agresiva. Consiste en despreciar a los activistas, por considerarlos comerciales y patrocinados por Occidente, para provocar la anarquía y destruir los países en los que vivimos. Esto es una especie de contradicción, pues son estos mismos países occidentales u orientales quienes financian la mayor parte de los presupuestos de nuestro gobierno, incluidos pilares esenciales como la policía y la sanidad. Los gobiernos reciben la mayor cantidad de ayuda en sentido figurado.

“ Las herramientas digitales permiten a los jóvenes movilizarse rápidamente contra la brutalidad policial o la corrupción. Pero Internet también tiene riegos, como la desinformación, la vigilancia y la intimidación ”

Se ha producido un aumento global de las agencias estatales que vigilan teléfonos y se infiltran en grupos de WhatsApp. Esto nos provoca terror, pues es literalmente la diferencia entre la vida y la muerte. Hay personas que han sido secuestradas e incluso asesinadas a causa de un post, y esto se ha normalizado hasta el punto de que incluso los secuestros ya no parecen ilegales. Se ha convertido en un acontecimiento semanal y se está secuestrando incluso a miembros electos que prestan servicio en las Asambleas, no sólo a activistas, sino a ciudadanos de a pie que deciden hacerse oír.

A principios de junio de 2025, un valiente profesor llamado Alfred Ojwang publicó un post sobre la corrupción en el sector policial. La policía lo detuvo y lo condujo a más de 300 km de su casa (pasando por varias comisarías) para interrogarlo. Fue trasladado a la Comisaría Central de Policía, situada en el centro de Nairobi. Horas más tarde, murió bajo custodia. La policía trató de hacerlo pasar por un suicidio, pero gracias a la presión de la opinión pública y a que los activistas se negaron a apartarse de su lado, la autopsia concluyó que Alfred había sido torturado y golpeado hasta la muerte. Rastrear el dispositivo de Alfred por una publicación suya en la que «empañaba» el historial de alguien. Su historia nos persigue porque nos recuerda que no sólo luchamos contra las narrativas; luchamos por nuestras vidas.

Aun así, lo que nos hace seguir adelante es la fe. La fe en que otra Kenia, otra África Oriental, es posible. Una en la que las voces no se silencien, sino que se intensifiquen. Una en la que no solo se hable de justicia en chats encriptados, sino en la que la justicia se practique.

El caso de Boniface Mwangi y Agather Atuhairu: Un punto álgido

En mayo de 2025, los activistas de derechos humanos Boniface Mwangi (Kenia) y Agather Atuhaire (Uganda) viajaron a Dar es Salaam (Tanzania) para apoyar al líder de la oposición, Tundu Lissu, durante una comparecencia ante el tribunal. Fueron secuestrados por agentes de seguridad tanzanos, que les vendaron los ojos, los desnudaron, agredieron y los abandonaron en sus respectivas fronteras. La brutalidad fue escalofriante. Existía una clara coordinación transfronteriza. Ahora nos preocupa no sólo que nos sigan la pista en nuestros países, sino en nuestra comunidad, en la región.

Lo que vino después fue igualmente significativo: el rugido digital. Los activistas se vieron obligados a unirse más allá de sus fronteras y a ampliar su acción: se manifestaron por toda África Oriental, inundando las cronologías con llamamientos a la justicia. Siasa Place lanzó, junto con otros, un ultimátum de 72 horas a los organismos regionales exigiendo que se tomaran medidas. Los *hashtags* fueron tendencia en distintos países y colaboramos con medios de comunicación locales e internacionales, que se hicieron eco de la historia. Por una vez, el silencio no era una opción.

“ El ciberactivismo ha demostrado que las fronteras son insignificantes frente a la represión. Con un video viral y una acción regional coordinada podemos hacer ruido para conseguir justicia ”

Este caso demostró tanto el peligro como el poder del activismo digital. Si no hubiera habido un hilo en Twitter, un vídeo viral o una indignación coordinada, lo más probable es que su historia hubiera acabado en el más absoluto silencio. Pero hicimos ruido. Y el ruido puede conducir a la justicia; aún estamos en ese camino.

El papel de la solidaridad regional

El ciberactivismo ha demostrado que, a menudo, las fronteras son insignificantes frente a la represión. Ya sea en Nairobi, Kampala o Dar es Salaam, el manual de

estrategia es familiar: vigilancia, desinformación e intimidación.

Por eso la solidaridad regional tan importante. No sólo necesitamos aliados, necesitamos sistemas. Sistemas que están en constante adaptación porque a menudo encuentran formas de intentar cerrar plataformas o la propia Internet. Necesitamos protocolos y mecanismos de protección mutua. Cuando se llevaron a Mwangi y Atuhaire, no esperamos a que los gobiernos actuaran. Los ciudadanos preocupados y la sociedad civil llenaron el vacío.

Fomentar la resiliencia: La labor de Siasa Place

La labor de Siasa Place abarca la formación, la investigación y la defensa. Abogamos por derechos que ya están contemplados en nuestra constitución conscientes que los trabajadores autónomos merecen contratos, protecciones y dignidad. Actualmente estamos cartografiando el ecosistema tecnológico para identificar dónde se violan los derechos y dónde es posible intervenir. A través de TrustLab, por ejemplo, formamos a organizaciones de base en seguridad digital, enseñamos a usar el VPN y la autenticación de dos factores, y concienciamos sobre el phishing y las estrategias de verificación de hechos. También ayudamos a las comunidades a crear una higiene digital del mismo modo que las generaciones pasadas crearon sindicatos o grupos de ahorro. E invertimos en contar historias, porque si no contamos nuestras historias, alguien lo hará, distorsionándolas.

Trabajamos con organizaciones comunitarias, asesorándolas para que se institucionalicen y se conviertan en líderes de sus propias comunidades, y para que aprendan sobre ciclos presupuestarios y la importancia de la responsabilidad social, ya que la constitución keniana otorga a los ciudadanos el derecho a participar en ese proceso. Deben aprender a exigir responsabilidades a los dirigentes locales, sobre todo utilizando los espacios y plataformas ya previstos al efecto.

“ Una sociedad digitalmente pacífica es aquella en la que los activistas no necesitan teléfonos

desechables; donde las leyes protegen la libertad de expresión, donde podemos discrepar en voz alta, con pasión y seguridad ”

A los jóvenes líderes interesados en asumir el liderazgo político les asociamos con partidos políticos y les animamos a que se afilien y desempeñen un papel activo en la política. Nuestra principal labor es la concienciación educativa, porque cuando estos procesos se dan a conocer, los jóvenes suelen estar dispuestos a emprenderlos.

La paz en la era digital

La paz ya no es sólo una cuestión de alto el fuego o de ciclos electorales. La paz tiene que ver con la protección de datos, con el derecho a publicar sin miedo, con que tu activismo haga que consigas un retuit o una pena de cárcel.

Una sociedad digitalmente pacífica es aquella en la que los activistas no necesitan teléfonos desechables; donde las leyes protegen la libertad de expresión, no la criminalizan; donde podemos discrepar en voz alta, con pasión y seguridad. Y esto no es sólo responsabilidad de los gobiernos, también las plataformas, los donantes, los aliados internacionales, todos tienen un papel que desempeñar. Porque si una voz se silencia en Internet, todos perdemos algo.

No es sólo un momento; es un movimiento. No estamos codificando aplicaciones; estamos codificando resistencia. Cada tuit es un voto; cada vista, una vigilia. Cada mensaje de WhatsApp, cada directo, cada meme, todo importa.

Conocemos los riesgos, pero también conocemos nuestro poder. Construimos la paz digital en África Oriental, sabiendo que el silencio nunca fue una opción. Internet no nos ha salvado, nos hemos salvado los unos a los otros. Y aún no hemos terminado.

Esta es una versión traducida del artículo publicado originalmente en [inglés](#).

Sobre la autora

Nerima Wako

Nerima Wako es la directora ejecutiva de Siasa Place, una organización juvenil creada en 2015, dedicada a crear de forma colaborativa un entorno que permita a los jóvenes de Kenia participar directamente en la corriente política de forma significativa. Siasa Place educa a los jóvenes sobre los procesos electorales, la constitución, las instituciones y funciones del gobierno. Está especializada en tecnología cívica, campañas políticas digitales y promoción.

Nerima es licenciada en Periodismo y Sociología por la Universidad Estatal de Jacksonville (Estados Unidos) en 2010 y obtuvo un máster en Administración Pública en 2012. Ha ocupado puestos destacados, entre ellos, vicepresidenta del Comité de Coordinación Juvenil; miembro del Kenya Eminent Peace Panel (2022); miembro del Consejo para los Medios Sociales Responsables; miembro de la Junta de Sondeo Juvenil del PNUD o miembro del Comité Asesor Juvenil de Países Bajos (2022-2023).

Fotografía

Momento de intervención durante un encuentro organizado por Siasa Place. Autor: Siasa Place (Facebook de la organización).

ARTÍCULOS CENTRALES

Documentar la cultura antes, durante y después del conflicto

Wahbi Abdalrahman

Coordinador del proyecto Sudan Memory

La localidad de Merowe linda con el río Nilo en el estado del Norte de Sudán, cerca de la presa hidroeléctrica y a más de 300 km al norte de la capital, Jartum. Cuando Wahbi Abdalrahman y su equipo de la Universidad del Valle del Nilo (NVU) llegan aquí con sus escáneres Epson y Canon, su ordenador portátil y sus cámaras digitales, la temperatura exterior suele rozar los cuarenta grados centígrados. La zona ha cambiado en los dos últimos años, desde que las Fuerzas de Apoyo Rápido (RSF) fueron expulsadas de Merowe poco después del inicio de la guerra civil, en abril de 2023. El tradicional mercado de los miércoles sigue siendo bullicioso, pero muchas de las mercancías que se venden son ahora egipcias, en lugar de sudanesas, y entre los clientes hay un gran número de desplazados que han escapado del conflicto que asola otras partes del país.

En este difícil contexto, el equipo de Wahbi crea su unidad itinerante de digitalización. Desde abril de 2023, financiados por dos becas semestrales del centro de Respuesta a Emergencias Culturales -conseguidas con la profesora Marilyn Deegan-, han escaneado más de 60.000 páginas de documentos en los estados del Norte y del Río Nilo. Muchos de estos documentos proceden de colecciones familiares. Aun así, constituyen y registran importantes acontecimientos de la historia de Sudán que, o bien estuvieron siempre ausentes de los relatos oficiales durante la dictadura de Omar al-Bashir, o bien están siendo borrados ahora por la guerra civil en curso.

Utilizando su cámara digital y su equipo de grabación de audio, el grupo de Wahbi también ha capturado secuencias de vídeo que documentan relatos de mercados

abandonados, figuras históricas olvidadas y prácticas tradicionales de la región. Él siempre viaja con un cuaderno. En él anota cualquier historia o detalle adicional que le cuenten los propietarios de la colección sobre los objetos que se están digitalizando porque, con los años, su equipo ha llegado a apreciar que cada pertenencia es más que una imagen, tiene una historia. Preservar estas historias y vestigios históricos será crucial para comprender y recordar este periodo.

“ Sudán, con 19 grupos étnicos, posee uno de los patrimonios más ricos y diversos del mundo. En poco más de una década, Sudan Memory ha digitalizado más de 400.000 objetos ”

Proyecto de digitalización

El trabajo de Wahbi contribuye al proyecto más amplio de Sudan Memory (Memoria de Sudán, en adelante MS) iniciado en 2013 para digitalizar el patrimonio cultural en riesgo de deterioro y destrucción del país a causa del cambio climático o de posibles conflictos futuros, cuya magnitud ha superado con creces lo imaginado en aquel momento. La universidad NVU es uno de los socios sudaneses e internacionales que contribuyen al proyecto, en el que también participan la Oficina Nacional de Registros de Sudán (NRO), la Sociedad Sudanesa para el Archivo del Conocimiento (SUDAAK), el Museo de la Mujer de Darfur, la Corporación de Radio y Televisión de Sudán (SRTC), la Universidad de Jartum, la Universidad de Durham y el King's College de Londres, entre otros.

Sudán posee uno de los patrimonios más ricos y diversos del mundo, con diecinueve grupos étnicos principales que hablan más de cien lenguas y dialectos. Su patrimonio arqueológico se remonta a varios milenios y el país cuenta con más de doscientas pirámides, además de ser rico en ajuares y restos funerarios, pinturas murales y artefactos, que la mayor parte del mundo, más allá de Sudán, ignora. En poco más de una década —y a pesar de un golpe de estado, cambios de régimen y conflictos— los

socios de MS han digitalizado más de 400.000 de estos objetos del patrimonio cultural sudanés.

Durante el proyecto, los socios descubrieron riquezas inesperadas, como la copia original de un telegrama enviado por Zubair Pasha desde El Cairo en 1883, oculto en el interior de un manuscrito en un pueblo al norte de Berber, y cartas originales escritas a mano por Mahdi, que Wahbi y su equipo digitalizaron cuidadosamente. La creciente colección digital comprende ahora películas, fotografías, manuscritos, objetos de museo, archivos de audio, historias orales e incluso un modelo histórico interactivo en 3D de la isla de Suakin, así como todos los metadatos asociados en árabe e inglés. Lo que todos estos elementos atestiguan es una cultura rica y resiliente que es más fuerte gracias a su larga historia de diversidad y coexistencia.

“ Borrar cualquier registro de la diversidad y coexistencia cultural es un arma de guerra probada en todo el mundo ”

Conflicto y culturicidio

Intentar borrar cualquier registro de dicha coexistencia cultural es un arma de guerra probada en todo el mundo. En años recientes, hemos sido testigos de cómo los talibanes dinamitaban las estatuas de los Budas de Bamiyán, de cómo Estado Islámico destruía los templos de Palmira y de la quema deliberada de la biblioteca de Sarajevo, un acontecimiento que el director de teatro bosnio Gradimir Gojer describió como «un triunfo de la barbarie y la muerte de la cohabitación de musulmanes, ortodoxos, católicos y judíos que había existido durante siglos en Bosnia-Herzegovina». Como espacios para «memorias e identidades compartidas», cuando dichos monumentos, archivos, museos y bibliotecas son borrados en conflictos, solo dejan tras de sí lo que la profesora de la SOAS, Dina Matar, describe como «vacíos de memoria» (Matar, 2023, citada en Khaled et al., 2023).

Desde abril de 2023, hemos visto surgir este tipo de vacíos en todo Sudán, amenazando con perforar su memoria. En Jartum, como en otros lugares, las fuerzas de la RSF saquearon, dañaron o destruyeron miles de valiosos objetos del Museo Nacional que documentaban la larga historia y cultura de Sudán y las numerosas civilizaciones que ocuparon esta región a lo largo de los siglos. La UNESCO ha advertido de una «amenaza para la cultura». Al mismo tiempo, Ikhlas Abdel Latif Ahmed, director de museos de la Corporación Nacional de Antigüedades y Museos (NCAM) de Sudán —uno de los socios fundadores de MS— afirmó con mayor rotundidad que las fuerzas de la RSF «destruyeron nuestra identidad, y nuestra historia» (Copnall, 2025).

“ La colección de Sudan Memory es ahora todo lo que queda de gran parte de la cultura sudanesa; desde el inicio de la guerra en 2023 se han destruido muchos elementos físicos, colecciones y edificios ”

Los informes y anécdotas de los socios y conocidos de MS que aún se encuentran en Sudán describen la destrucción de los edificios de la Corporación de Radio y Televisión, que anteriormente era uno de los archivos cinematográficos más extensos de África, con grabaciones de radio, vídeo y cine que se remontan a la década de 1940 y constituyen un recurso histórico único para Sudán. Entendemos que los socios de otras partes del país ciertamente han perdido colecciones y, a medida que se revele el alcance total del conflicto, prevemos noticias de muchos más vacíos de memoria.

En el estado de Darfur del Sur, en la ciudad de Nyala, se ha rescatado y colocado en un lugar secreto la colección de más de 4.000 objetos que se reunió y comisarió para el Museo de la Mujer de Darfur con el fin de contar las diferentes historias de la región. Sin embargo, en las circunstancias actuales, no es posible mostrar estos artículos. Como señaló Kate Ashley, gerente del Proyecto de Consultoría de MS, esta colección «demostró realmente cómo se pueden reunir y comprender colecciones y objetos personales y cotidianos que de una manera que realmente documenta la historia social

de un lugar y tiene tanto significado». Actualmente, solo la colección digital que MS pasó meses fotografiando en 2021, junto con una hermosa entrevista con la fundadora del museo, Fatima Mohamed Al Hassan, lamentablemente ya fallecida, es todo lo que está accesible, y solo a través del sitio web de MS.

Santuario de archivo

Sabemos que la colección Memoria de Sudán es ahora todo lo que queda de gran parte de la cultura sudanesa, ya que desde 2023 se han destruido muchos elementos físicos, colecciones y edificios. El proyecto ha adquirido, por supuesto, una nueva urgencia con este telón de fondo, tanto para garantizar la supervivencia a largo plazo de los artefactos digitales existentes como para facilitar la digitalización de otros registros patrimoniales amenazados de pérdida.

Que Wahbi y su equipo de la NVU puedan continuar su trabajo con una ayuda muy limitada del exterior de Sudán es un testimonio del enfoque descentralizado de la digitalización del proyecto. En sus inicios, los directores del proyecto, el Dr. Badreldin Elhag Musa (SUDAAK) y la profesora Marilyn Deegan (King's College de Londres), previeron llevar a cabo todas las actividades de digitalización en un *hub* central en la Ciudad Africana de la Tecnología de Jartum. Sin embargo, en 2017, en una primera reunión tras conseguir subvenciones del British Council y la Fundación Aliph por valor de 800.000 libras, los socios sudaneses dejaron claro que, en su lugar, querían que las actividades de exploración tuvieran lugar *in situ*, en ubicaciones descentralizadas. Hubo intensas negociaciones, pero los financiadores accedieron a apoyarlo.

“ Abordar los vacíos de memoria permite que la gente pueda reconstruir su cultura y sus historias con dignidad y esperanza ”

Uno de los muchos resultados positivos de esta decisión es que las capacidades de digitalización se han integrado en estas instituciones asociadas. Las personas formadas por la MS, como Wahbi, fueron capacitadas para formar a sus colegas y

desarrollar sus propios equipos de digitalización, lo que permite que el trabajo continúe incluso ahora.

Las instituciones asociadas, como la NVU, también conservaron el equipo proporcionado durante el proyecto y tomaron decisiones sobre los materiales que digitalizarían. Por supuesto, como en todas las instituciones, la arraigada burocracia y la política de poder influyeron en qué materiales se seleccionaron para su digitalización. En algunos casos, no se escogieron piezas que el equipo de MS consideraba valiosas, pero que la institución poseedora consideraba demasiado controvertidas o sensibles desde el punto de vista de la reputación. Por otro lado, los socios sudaneses, como Wahbi, que tienen un profundo conocimiento de lo culturalmente sensibles que son ciertos temas para la sociedad sudanesa, pudieron hacer valoraciones cuidadosas sobre si las colecciones privadas, como las que poseen las familias de la localidad de Merowe, debían hacerse públicas.

Este enfoque de la selección y la valoración fue esencial para generar confianza con los socios sudaneses, que comprensiblemente desconfiaban de ser explotados por el tipo de proyectos de digitalización extractiva experimentados en otros lugares del continente africano. En esos casos, los africanos habían comprometido su mano de obra y los recursos de su patrimonio cultural en proyectos de digitalización financiados por instituciones del Norte Global solo para descubrir que no tenían acceso ni control sobre los artefactos digitales que habían producido (Breckenridge, 2014; Pickover, 2014; Rassool, 2018; Chamelot, Hiribarren y Rodet, 2020).

En un esfuerzo concertado por mitigar estas preocupaciones sobre el neocolonialismo, el proyecto MS se basa en una serie de acuerdos y memorandos de entendimiento que garantizan que las instituciones asociadas mantengan el control último sobre cómo se reutilizan los objetos digitalizados, así como la conservación de copias digitales de estos. Así pues, aunque en el repositorio de datos de investigación del King's College de Londres (KORDS) se almacenan copias digitales con calidad de archivo de miles de artículos de MS, el permiso para reutilizar cualquiera de los artículos digitales requiere primero ponerse en contacto con el propietario original de los derechos. Trágicamente, el conflicto en curso significa que muchos de estos propietarios de derechos están

ahora desaparecidos, algunos dados por muertos.

“ En el postconflicto la colección digital de Sudan Memory servirá para recordar y revitalizar la identidad compartida y diversa de Sudán, y su rico patrimonio ”

En otros casos, en los que los socios sudaneses no dieron permiso para que las copias digitalizadas de los materiales salieran del país, estas también han sido destruidas junto con las versiones tangibles, de modo que no queda más que un vacío. Sin embargo, como señala Ikram Madani, director del Museo de Historia Natural de la Universidad de Jartum, «si nuestras colecciones físicas han sido destruidas como consecuencia de la guerra, al menos tendremos los registros digitales de los objetos para reconstruir nuestras colecciones a partir de ellos».

Posconflicto

La intención del proyecto MS siempre ha sido devolver el repositorio digital a una institución sudanesa y conservar otra versión como copia en algún lugar fuera del país. En el periodo inmediatamente posterior al conflicto, la colección digital de MS será un recurso importante para recordar y revitalizar la identidad compartida y diversa de Sudán y su rico patrimonio, elementos importantes para negociar un futuro basado en la coexistencia cultural para el que hay, sin duda, claros precedentes. En un reciente proyecto de historia oral dirigido por la periodista Sara El-Nager, en el que entrevistó a los participantes sobre sus funciones y reflexiones sobre MS, muchos señalaron el impacto personal positivo que el proyecto había tenido en ellos. Hablaron efusivamente de sus viajes a otras partes menos conocidas de Sudán, donde conocieron a personas de diferentes ámbitos, lo que les mostró la diversidad cultural del país pero también lo que muchos tenían en común.

La forma en que se despliegan estos recursos estará determinada por las comunidades sudanesas tanto dentro del país como en su creciente diáspora. Para Asia

Mahmoud, coordinadora local de MS e investigadora de la colección, este proyecto de digitalización es mucho más que su contenido: «Mucha gente se encontró realmente con este shock de perderlo todo [durante la guerra], y se convirtió en algo personal para todos. Proyectos como MS no solo preservan nuestra historia, cultura y patrimonio, sino que en realidad serán como un consuelo para todos; nos hace sentir que estamos aquí, que aún existimos.»

Garantizar que el recurso digital esté disponible tras el conflicto depende, sin embargo, de asegurar que Memoria del Sudán disponga, en su totalidad, un santuario en algún lugar que la salvaguarde y preserve de las amenazas de corrupción, degradación y obsolescencia a largo plazo. Su urgencia es incomparable con la necesidad de aliviar el sufrimiento que actualmente se inflige al pueblo sudanés y de garantizar el fin del conflicto. Nuestro deseo es, sin embargo, que al garantizar el futuro a largo plazo de MS, se puedan abordar los vacíos de memoria para que la gente pueda reconstruir su cultura y sus historias con dignidad y esperanza, y de forma que se salvaguarden los acuerdos de paz.

Referencias bibliográficas:

Breckenridge, Keith. 2014. “The Politics of the Parallel Archive: Digital Imperialism and the Future of Record-Keeping in the Age of Digital Reproduction.” *Journal of Southern African Studies* 40 (3): 499–519.

Chamelot, Fabienne, Vincent Hiribarren, y Marie Rodet. 2020. “Archives, the Digital Turn, and Governance in Africa.” *History in Africa* 47: 101–18.

Copnall, James. 2025. “From Prized Artworks to Bullet Shells: How War Devastated Sudan’s Museums.” *BBC News*, 26 de abril.

Khaled, Mai, Heba Saleh, Lucy Rodgers, Alexandra Heal, y Dan Clark. 2023. “How the Loss of Entire Families Is Ravaging the Social Fabric of Gaza.” *Financial Times*, 13 de diciembre, 2023.

Pickover, Michele. 2014. "Patrimony, Power and Politics: Selecting, Constructing and Preserving Digital Heritage Content in South Africa and Africa." Paper presentado en *IFLA WLIC 2014 - Libraries, Citizens, Societies: Confluence for Knowledge*, Lyon, Francia, 16-22 de agosto.

Rassool, Ciraj. 2018. "Digitisation and the Government of Collections." Paper presentado en *Postcolonial Digital Connections*, Halle, Alemania, 16-17 de mayo.

Esta es una versión traducida del artículo publicado originalmente en inglés.

Sobre los autores

Wahbi Abdalrahman

Director del Centro de Documentación y Estudios Etnográficos del Valle del Nilo en la Universidad del Valle del Nilo. Ha desempeñado funciones administrativas clave, entre ellas vicedecano de Bibliotecas y dirección del equipo del proyecto Sudan Memory (Memoria de Sudán), en el estado del Río Nilo. Su carrera abarca décadas de liderazgo en documentación, digitalización y biblioteconomía.

Sara El-Nager

Periodista sudanesa británica que apoyó el proyecto Sudan Memory (Memoria de Sudán) con actividades de digitalización mientras vivía en Jartum. Desde que abandonó Sudán en 2023, ha desarrollado una historia escrita del proyecto basada en historias orales con participantes clave.

Marilyn Deegan

Catedrática emérita de Humanidades Digitales en el King's College de Londres y directora del proyecto Sudan Memory (Memoria de Sudán). Cuenta con más de treinta años de experiencia dirigiendo proyectos de digitalización y tiene formación como

medievalista, especializada en textos médicos anglosajones.

Laura Gibson

Profesora titular del Departamento de Humanidades Digitales del King's College de Londres. Su investigación sobre la descolonización y la digitalización se nutre de varios años de trabajo en museos sudafricanos, entre ellos como directora de Colecciones y Digitalización en el proyecto de legado nacional del Museo Luthuli.

Fotografía

Hombre revisando documentos de archivo a mano, rodeado de carpetas y papeles, durante el proceso de digitalización de materiales históricos. Autor: Wahbi Abdalrahman (Sudan Memory).

ENTREVISTA

Entrevista amb Stephanie Williams

Rita Costa

Responsable de Artes y Paz Creativa, Build Up

Stephanie Williams ejerció como alta funcionaria de Naciones Unidas entre 2018 y 2022 y su último cargo fue el de Asesora Especial del Secretario General de la ONU para Libia.

En esta conversación reflexiona sobre cómo la desinformación y la manipulación digital han marcado uno de los procesos de mediación internacional más complejos de los últimos años. A partir de su responsabilidad al frente de la misión de paz de la ONU en Libia, Williams describe cómo el discurso de odio *online*, el uso de la desinformación y la injerencia extranjera no sólo agudizaron las divisiones sobre el terreno, sino que también pusieron directamente en peligro a las mujeres negociadoras y constructoras de paz.

También explica cómo la misión de la ONU respondió a la violencia digital con colaboración directa con las plataformas de redes sociales y la celebración de diálogos en línea. Una experiencia recogida en su libro: *Libya Since Qaddafi: Chaos and the Search for Peace*.

¿Cuál era el contexto político y social de Libia cuando usted llega al país como mediadora de Naciones Unidas?

Libia emergió tras 42 años de régimen autoritario en 2011, cuando Muamar el Gadafi fue destituido como consecuencia de una revuelta popular. Durante las cuatro décadas de su régimen caprichoso y violento, el Gobierno controlaba estrictamente la información. No había libertad de expresión, palabra o reunión. Tras el derribo de Gadafi, la población pasó de una libertad de expresión nula a una libertad total, sin restricciones.

Naturalmente, fue una transformación maravillosa, puesto que la población libia finalmente encontró su voz propia: surgieron muchos medios de comunicación, emisoras de radio, periódicos y canales de televisión.

Durante la Primavera Árabe, en la que Libia jugó un papel determinante, las revueltas se organizaron en gran medida a través de Internet. Yo me encontraba al otro lado del mundo árabe, en la pequeña nación insular de Baréin, donde las redes sociales también impulsaron un breve levantamiento. Por lo general, durante este período tumultuoso en la región, Internet se convirtió en un motor enorme de movilización. Lo mismo ocurrió en Libia, y la tendencia continuó tras el derribo de Gadafi, mayoritariamente por bien, pero en algunos casos también por mal, ya que cualquiera podía crear y editar noticias, o formar opiniones.

¿Cómo cambiaron las redes sociales la dinámica del conflicto en Libia?

Las redes sociales se convirtieron en un arma de doble filo a medida que evolucionaba el conflicto. La plataforma que más prosperó, y sigue prosperando, es Facebook. En Libia existen más cuentas de Facebook que personas. Tras la destitución de Gadafi, las cosas no salieron como se esperaba: el país empezó a fracturarse y luego cayó en un conflicto armado. Hay muchas razones que explican ese deterioro, pero una de ellas fue el uso de las redes sociales para dividir a la población e incitar a la violencia mediante el discurso del odio.

Después Libia vivió dos guerras civiles: una en 2014 y otra en 2019-2020, la que yo presencié. En esta última, parecía a menudo que el conflicto se manifestaba tanto en Internet como en el campo de batalla físico, con una retórica polarizante y discurso de odio y «alteridad». La deshumanización de «enemigo» nos dificultó inmensamente la promoción del diálogo como misión de Naciones Unidas. Desde nuestra posición, nos pronunciamos enérgicamente contra el discurso de odio e intentamos colaborar con varios medios de comunicación. Reunimos a algunos de los principales *influencers* para animarlos a moderar su tono, con resultados diferentes, pero no pudimos reunir en una misma sala personalidades mediáticas de bandos opuestos porque el conflicto sobre el terreno era demasiado intenso.

En ese momento, cuando ejercía el cargo de mediadora principal de Naciones Unidas es cuando la violencia digital fue más peligrosa.

“ Las redes sociales se convirtieron en un arma de doble filo en Libia. Durante el conflicto se utilizaron para dividir a la población e incitar a la violencia con discursos de odio ”

¿Qué otros obstáculos se encontraron en materia de desinformación?

Un gran obstáculo fue la injerencia externa. La mayoría de los observadores son conscientes de que algunos países desplegaron mercenarios y suministraron armas en flagrante violación del embargo de armas de la ONU, pero también es cierto que ciertas capitales extranjeras emplearon ejércitos electrónicos y coordinaron campañas de desinformación. Por ejemplo, cuando el general Haftar lanzó su ataque sorpresa contra Trípoli en abril de 2019, hubo un importante componente de (des)información asociado al mismo para hacer que su ataque pareciera un hecho consumado, cuando en realidad se enfrentó a una fuerte resistencia.

Tras el cese de las hostilidades y el alto el fuego, este último mediado por mí en octubre de 2020, pasamos a las conversaciones políticas bajo los auspicios del Proceso de Berlín, respaldado internacionalmente. Fue durante las conversaciones políticas cuando vimos el daño más inmediato y personal causado por la desinformación y el discurso de odio. Las mujeres que participaban, en particular, se convirtieron en el principal objetivo de los abusos en línea, la difamación y la intimidación.

De hecho, la violencia contra las mujeres en Libia ya se había manifestado previamente.

Antes de 2019 activistas, políticas y parlamentarias libias prominentes ya fueron señaladas. Salwa Bugaighis, activista destacada por los derechos humanos, fue asesinada en 2014 en Bengasi, durante las elecciones parlamentarias. La legisladora

Fariha Berkawi también fue asesinada en Derna, el verano del mismo año. En julio de 2019, en pleno auge de la guerra civil, Seham Sergiwa, diputada de Bengasi, fue víctima de desaparición forzada. Era muy valiente y había criticado abiertamente el uso de la fuerza militar. Una noche, después de expresar en la televisión lo que yo calificaría de críticas moderadas hacia las fuerzas del general Haftar, unos hombres enmascarados irrumpieron en su casa y la secuestraron. Nunca más se le ha vuelto a ver.

Así que ya existía un ambiente de temor. Por eso, en noviembre de 2020, cuando iniciamos las conversaciones políticas, reunimos a todos los participantes fuera del país -entre ellos 17 mujeres-, porque las condiciones en Libia eran demasiado delicadas. Naturalmente, las participantes femeninas estaban preocupadas por su seguridad.

¿Cómo se produjeron los ataques contra las mujeres a nivel digital?

Las mujeres empezaron a ser blanco de ataques en Facebook. Todo empezó con la creación de páginas falsas con el objetivo no sólo de avergonzarlas públicamente por su participación política, sino también de intimidarlas, a ellas ya sus familias.

Una de las medidas que tomamos para contrarrestar ese fenómeno fue establecer una relación de confianza con Facebook. Para entonces, era bien sabido que la desinformación había contribuido a agravar los conflictos —por ejemplo, en Myanmar—, con lo cual Facebook se mostró sensible hacia el problema. Libia no era un mercado muy grande para la empresa —sólo siete millones de personas—, pero Facebook era consciente de las implicaciones negativas de su falta de actuación en otros casos, por lo que, cuando señalamos estas cuentas falsas, las eliminaron en menos de 24 horas. Esto nos permitió mantener a las mujeres en la sala de negociación.

“ Las mujeres se convirtieron en el principal objetivo de los abusos en línea, la difamación y la intimidación a través de la creación de páginas falsas ”

Sin embargo, mientras llevábamos a cabo la mediación política, en noviembre del 2020, otra activista de Bengasi, Hanan al Barassi, fue tiroteada a plena luz del día para hablar en las redes sociales sobre los hijos del general Haftar. Era un entorno peligroso para las mujeres. Al final, perdimos a una mujer del diálogo político, porque se retiró. Era del este de Libia. Cada uno puede sacar sus propias conclusiones sobre lo que la motivó a marcharse. Mi función no consistía en cuestionarla: era evidente que todas estas mujeres estaban sometidas a una terrible presión. La violencia digital directa generó miedo entre las participantes en el diálogo político.

¿Hubo también desinformación coordinada dirigida a la propia mediación de la ONU?

Durante la primera ronda de nuestras negociaciones políticas, se llevó a cabo una amplia campaña de desinformación, cuyos orígenes e impacto desconocíamos en ese momento. Veíamos muchas cosas en las redes sociales, pero no entendíamos de dónde provenían.

No fue hasta un mes después de esa ronda de conversaciones cuando comprendimos lo sucedido. Un estudio del Observatorio de Internet de la Universidad de Stanford reveló que la Agencia de Investigación de Internet del antiguo director ruso de la empresa Wagner, Yevgeny Prigozhin, en coordinación con ejércitos electrónicos de la región de Oriente Próximo, llevó a cabo una campaña masiva de desinformación contra la mediación de la ONU en Libia.

Yo fui directamente acosada por Prigozhin. Como describo en mi libro, el incidente más grave ocurrió en enero de 2022 de visita en Moscú. El director ruso – o un doble suyo – se presentó en mi hotel con un grupo de emisarios y me siguió hasta el ascensor. No nos dimos cuenta de la operación hasta que al día siguiente se publicaron unas fotos en Internet para relatar una historia falsa de que yo había hecho un trato con Prigozhin. Yo todavía estaba en Moscú cuando salió la noticia y fue realmente intimidante para mí y para mi equipo, dado lo que sabíamos sobre su brutalidad. La ONU en Libia ya había sido víctima de ataques violentos; tres trabajadores fueron asesinados en Bengasi el 2019 en circunstancias que siguen siendo muy turbias.

El Observatorio de la Universidad de Stanford señaló que lo que ocurría en Libia formaba parte de un patrón general que también se había producido en Siria y Sudán. El estudio motivó a Facebook a eliminar cientos de páginas relacionadas con la operación.

¿Qué tecnologías promovieron más la violencia digital?

Creo que por ahora ya ha quedado demostrado que las campañas de desinformación, el uso de algoritmos y la manipulación de las redes sociales fueron factores determinantes, por ejemplo, en el Brexit o en las elecciones norteamericanas de 2016. Por tanto, no debería sorprender que, básicamente, algunas de las mismas fuerzas, o similares, aplicaran los mismos métodos en el contexto libio.

Rusia tenía un gran interés en Libia y daba pleno apoyo a una de las partes en conflicto: ya habían suministrado armas y mercenarios a Haftar. Les interesaba conseguir un resultado que no pusiera en peligro su posición o intereses sobre el terreno en Libia y, en general, en el continente africano. No fue una gran sorpresa, pero el descubrimiento reforzó nuestro compromiso y nuestro enfoque —o al menos mi enfoque— de que la única forma de combatir esto era ser completamente transparentes sobre el proceso de mediación y el objetivo final, con el fin de tranquilizar a la población libia.

**“ Para combatir la desinformación,
necesitábamos que el proceso de mediación
fuera completamente transparente, que la
población libia pudiera observar lo que ocurría
en las salas de negociación ”**

En los algoritmos se puede ver cómo la gente se adentra en estos laberintos y se cree todo tipo de teorías conspiratorias. Cuando la principal fuente de información son las redes sociales y se vive en una cámara de resonancia, es muy difícil llegar a las personas. En algún momento se convierte en un rasgo identitario: se crea una nueva identidad, o una identidad añadida, complementaria a la identidad fuera de línea. Los teléfonos, ordenadores y cualquier tecnología que haya permitido el desarrollo de las

redes sociales se han convertido en armas por ambos bandos en conflicto.

¿En qué se tradujo su voluntad de transparencia en la estrategia de mediación?

El mejor desinfectante es la luz del sol, ¿verdad? Basta con iluminar lo que se está haciendo con la luz del sol, porque si no se está abierto y transparente, y se está a la ofensiva, en lugar de la defensiva, no se está respondiendo. Por tanto, necesitábamos ser mucho más transparentes y fuimos teniendo cada vez más compromiso con una transparencia más radical.

Tuvimos una conversación bastante complicada en el contexto de Naciones Unidas, porque – y lo respeto- hay mucha cautela. Pero sentí que el entorno de las redes sociales en Libia, incluida la injerencia extranjera, era tan agresivo que, si no estábamos allí, otros contarían la historia, y no sería la correcta.

En este contexto organizamos diálogos digitales, principalmente con jóvenes, para refutar las narrativas falsas en las redes sociales. Llevamos a cabo cinco diálogos en cinco meses, y también realizamos encuestas para demostrar cuál era la opinión de los jóvenes sobre temas clave: elecciones, reforma económica, desarme de las milicias y derechos humanos. Pudimos utilizarlo para contrarrestar discursos del tipo «ya sabes, no se pueden celebrar elecciones en Libia y eso es cosa de los Hermanos Musulmanes».

A través de los diálogos digitales, constatamos rápidamente que la gran mayoría de la población quería elecciones, un cambio en el statu quo, una mayor transparencia en la gestión de los ingresos petroleros, rendición de cuentas por las violaciones de los derechos humanos y justicia transicional. Y todo esto sucedía en directo, en árabe, y no podía ser rastreado por los «Grandes Hermanos» -servicio de inteligencia- que estaban *online*.

Estos diálogos se hicieron muy populares y se convirtieron en tendencia. Utilizábamos el humor, para llegar al público joven, y la gente se dio cuenta de que sus opiniones podían ser escuchadas.

¿La transparencia también fue una característica de las negociaciones políticas?

En febrero de 2021, en nuestro segundo diálogo político, decidí retransmitir la mayor parte del mismo en directo. Era el momento postpandemia y cuando se celebraron los diálogos la gente todavía sufría las secuelas de la COVID y sus efectos psicológicos, y era aún más importante mantener la máxima transparencia en la sala, que la población libia pudiera observar lo que sucedía, ya fuera la votación o las entrevistas a los candidatos para el gobierno de unidad.

Esa experiencia me confirmó que cualquier proceso político o de paz debe ser inclusivo. Por supuesto, la preferencia de muchos países –y de la clase dirigente en Libia– es mantener las negociaciones a puerta cerrada. Se reúnen cinco o tres hombres en una sala llena de humo, y se llega a un acuerdo, ¿no es así? Pero yo creía, y sigo creyendo que, en lugares como Libia, cuanto más inclusiva sea la paz, más probabilidades tendrá de consolidarse, o de tener la oportunidad de hacerlo. Y, en cambio, cuanto más se convierta en un simple reparto del pastel, o en un acuerdo para compartir el poder, más probabilidades habrá de que vuelva a degenerar en violencia.

“ Todo proceso de paz debe ser inclusivo. Es necesaria más participación de las mujeres, la juventud y la sociedad civil. Cuanto más inclusiva sea la paz, más probabilidades tendrá de consolidarse ”

Es necesaria más inclusión de las mujeres, de la juventud y de la sociedad civil, así como de los y las representantes de los grupos vulnerables, ya sean diferentes grupos étnicos o tribus.

En este contexto de violencia digital, ¿hay lugar para la esperanza?

Mi sensación es más de preocupación que optimismo. Durante la Primavera Árabe, por ejemplo, las redes sociales fueron un gran motor de movilización, pero cuando una tecnología comienza a utilizarse de forma positiva, llegan otras influencias para sabotearla y manipularla.

Las instituciones tradicionales, ya sean gobiernos elegidos democráticamente o las Naciones Unidas, siempre están en desventaja, intentan mantenerse al día con la tecnología que cambia rápidamente. Ahora presenciamos el uso de noticias falsas, incluso por parte de líderes mundiales destacados como el presidente Trump, normalizando la desinformación.

Mi pregunta es: si los británicos hubieran sabido cómo se estaba manipulando la campaña del Brexit en Internet, ¿habrían tomado las mismas decisiones? Los gobiernos elegidos democráticamente y las instituciones responsables suelen actuar con cautela. Siempre serán vulnerables, puesto que las fuerzas malignas irán un paso por delante y estarán dispuestas a utilizar las mismas herramientas sin ética. Para ellos no existe la moralidad, no existe el bien común.

Debemos concienciar a las personas sobre cómo y por qué son objeto de manipulación en Internet.

Esta entrevista ha sido traducida del original, en inglés.

Sobre la autora

Rita Costa

Rita Costa ha trabajado con Build Up desde 2020 facilitando procesos participativos de análisis de redes sociales en África Occidental y el sur de Europa, la co-creación de una residencia artística internacional, el desarrollo de marcos para entender y abordar la violencia digital y la organización de la conferencia anual Build Peace. Tiene formación en Derecho Internacional, Ciencias Políticas y coreografía.

Fotografía

La representante especial en funciones del Secretario General de Naciones Unidas, Stephanie Williams, informa a la prensa durante la cuarta ronda de la Comisión Militar

Conjunta Libia 5+5, Palacio de las Naciones. 21 de octubre de 2020. Autoría Violaine Martin (ONU).