



EL ALGORITMO DE LA VIOLENCIA

Cristina Marcano

¿Por qué sigues viva? Envuelta en la nebulosa de la más temprana adolescencia, Rebecca Sedwick no supo qué hacer, ni cómo buscar ayuda. Su respuesta fue un salto al vacío desde la torre de una vieja fábrica de cemento. Durante meses, las redes sociales habían replicado mensajes similares enviados por niñas de su misma edad: doce años. ¿Por qué no tomas cloro? Todo estuvo al alcance de su pantalla: los mensajes violentos y puntuales, instrucciones precisas para no seguir viviendo. Rebecca quedó atrapada en la noria digital como una pequeña mosca en una telaraña. Y ningún algoritmo pudo ayudarla a salir de otra manera.

Su historia era excepcional en 2013, cuando sucedió, pero ya no resulta tan extraña. Ha venido repitiéndose puntualmente con otros nombres. Cada vez con mayor frecuencia, la violencia digital se sale de la pantalla y rebota en el mundo *offline*. Y viceversa. Tampoco parecen raros los ataques y las extorsiones virtuales. Ni los linchamientos en línea, aunque todavía no deja de sorprenderme que un rumor en las redes pueda tener un desenlace violento en la vida real. Como el asesinato de dos hombres inocentes golpeados por una turba en Acatlán (Puebla), tras ser acusados de robar niños en un mensaje de WhatsApp que circuló —sin fundamento, ni dudas— cuando no había ni un solo caso de secuestro en el pueblo.

Pero no todas las agresiones en línea terminan tan trágicamente. De hecho, la mayoría se reproduce como si nada, de la manera más rutinaria. Hay millones de víctimas anónimas de un odio silvestre, menos

mortífero pero igualmente visceral. Un odio mediado por la tecnología, más difuso, amplio y generalizado. La ciberviolencia nuestra de cada día, algo común y corriente como el pronóstico del tiempo. Diariamente, en México se difunden a través de las redes sociales entre quince y veinte mil mensajes violentos por razones de género, apariencia física y orientación sexual, según estimaciones del Consejo para Prevenir y Eliminar la Discriminación en la Ciudad de México (COPRED). Las principales víctimas son niños, jóvenes y mujeres.¹

Las redes sociales son un ring colectivo donde muchos boxeadores se golpean simultáneamente, sin árbitro ni reglas. Y, a la vuelta de un clic, podemos convertirnos en espectadores involuntarios de la pelea aunque no nos guste el boxeo. ¿Cuántas veces en la rutina del *scroll* nos han salpicado el sudor y la sangre? El odio virtual se puede ejercer fácilmente y sin mayores riesgos. No requiere audacia ni coraje. Es parte del ecosistema informático.

La violencia cibernética es, también, amplia y horizontal. Se ensaña particularmente con ciertos blancos por las diferencias de siempre: origen étnico o nacional, el sexo, las preferencias sexuales y la religión. Las clásicas víctimas del odio atávico. O, simplemente, por alguna de sus características físicas o su forma de vida, condición social o económica, preferencia política o gustos. No es un fenómeno marginal. Casi dieciocho millones de mexicanos (cerca del 21.4 por ciento de los internautas) sufren ataques digitales. La mitad, ofensas y críticas por alguna de sus características físicas o su forma de vida, condición social o económica.²

Pero, en realidad, nadie está a salvo. Los mensajes de odio pueden dispararse en cualquier momento contra quien sea por razones inesperadas: una opinión, una fotografía, un video, un audio. No es necesario que roce algún tema polémico. En las redes sociales, más que en la vida real, es posible odiar y ser odiado por perfectos desconocidos.

¿Es la violencia en línea un reflejo de lo que sucede en la sociedad? Algunos sostienen que el problema no es la tecnología, sino los usuarios. La naturaleza humana. Como si las plataformas fueran cooperativas que dependieran únicamente de sus miembros. Ha habido incluso propuestas serias para educar



©Butcher Billy, cartel para *Black Mirror*, *Nosedive*, 2016

¹ COPRED, Boletín 161/2020, publicado el 16 de noviembre de 2020. Disponible en <https://bit.ly/3caBF9R>

² INEGI, "Comunicado de prensa 364/22", 2022. Disponible en: <https://bit.ly/3QXPTJQ>

Si vemos contenido tóxico es porque resulta rentable. Porque vende. Nos atrapa.

a la “ciudadanía digital”. Algo que, por sí solo, aisladamente, libra de toda responsabilidad a las empresas tecnológicas y luce imposible ante una masa informe —casi cinco mil millones de personas (el 64.5 por ciento de la población mundial)— cuyo único punto en común es el acceso a internet.

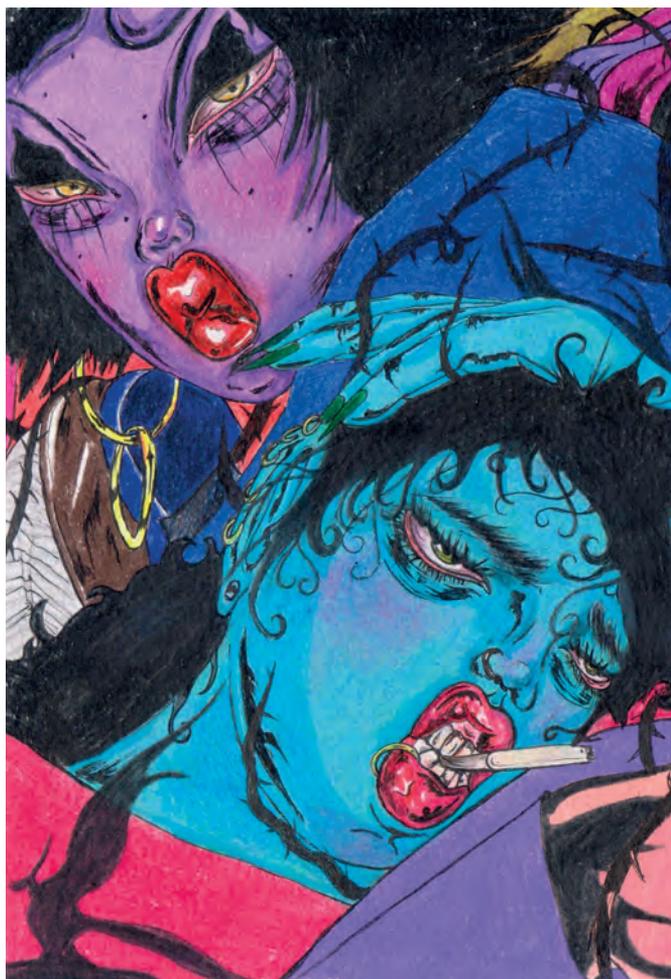
¿En realidad nos odiamos tanto y con tanta pasión? ¿Nos exaltamos automáticamente al conectarnos? ¿La violencia en línea se reproduce de manera espontánea? ¿Cómo aterrizan en nuestras pantallas los mensajes más agresivos y los contenidos más tóxicos sin buscarlos? ¿Estamos viendo lo que queremos ver o lo que un algoritmo decide mostrarnos? Una cosa es la violencia, su origen. Y otra, su difusión.

Fue en 1994 cuando por primera vez escuché el término “Superautopista de la información” en boca de uno de los profetas más activos de Internet, el senador estadounidense Al Gore, en una sala de conferencias de Buenos Aires. Entonces aquello sonaba más a relato de ciencia ficción. Parecía imposible que estuviera a la vuelta de la esquina y que pronto nos atraparía para convertirse en algo imprescindible. Ahora nos parece imposible vivir sin conectarnos. Si no estás en las redes, no existes, no cuentas.

En aquella época, había límites claros entre el periodismo y cualquier cosa que no lo fuera. No existía el término *viral* para designar un contenido muy popular. Las noticias falsas eran simplemente mentiras impublicables y el amarillismo estaba confinado a los panfletos de crónica roja. No había tribuna para una criatura imposible de concebir sin las redes: el ha-

ter. Tres décadas después, no es fácil distinguir. Todo, bueno o malo, cabe en la web, junto y mezclado en ese enorme globo gratuito de contenidos.

Internet puede ser también un espejismo. Uno piensa que está asomándose al mundo a través de la pantalla cuando, en realidad, lo más probable es que estemos entrando en un túnel. Basta que me detenga unos segundos en un video de gatos en Instagram para que en mi cuenta lluevan animales de todo tipo. Vivos,



©Fabiola Espiga, *Sepect 20*, 2020. Cortesía de la artista

heridos o muertos. Lo mismo sucede con la violencia.

Las redes sociales apuestan a la impulsividad, a los reflejos inmediatos, a la reacción más instintiva. Y les va de maravilla. Si vemos contenido tóxico es porque resulta rentable. Porque vende. Nos atrapa. Ya sea por morbo o indignación. Produce lo que las redes buscan: interacción entre los usuarios. No es inusual que se compartan encabezados sensacionalistas sin antes haberse leído el contenido.



Las redes no son plazas públicas donde todos podemos encontrarnos. Son un negocio que opera básicamente explotando las emociones, incluidos el odio, la burla, la violencia. Un gran negocio que se rige por una simple ecuación comercial: a más usuarios y más tráfico, mayor publicidad y más ganancias. Son las empresas tecnológicas —no nosotros— las que determinan la dinámica de las redes sociales en base a sus intereses económicos.

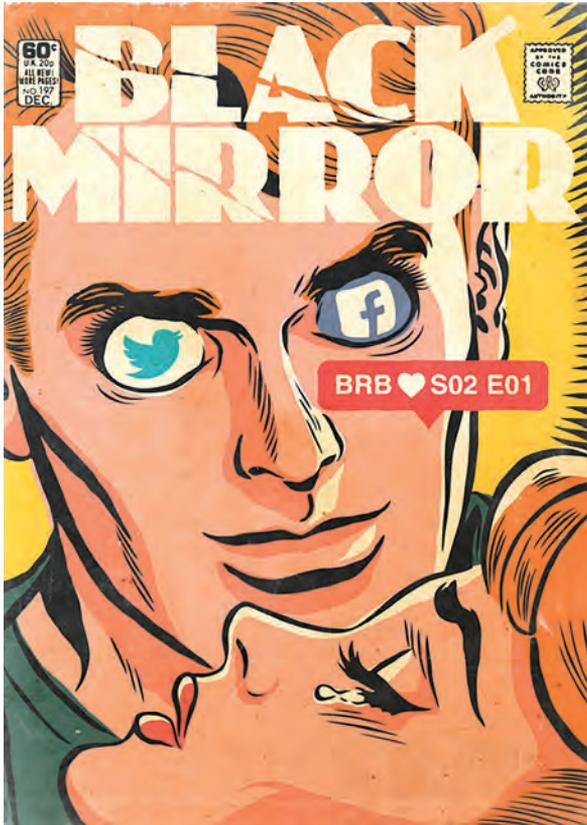
La clave es precisamente que la masa se enganche impulsivamente a la red. El sistema está diseñado para propagar los contenidos más atractivos: bromas inocentes o crueles, disputas triviales o importantes, chismes o teorías absurdas, memes y mentiras disfrazadas de noticias. O meras estupideces que, en no pocas ocasiones, son replicadas por periódicos rendidos a la banalidad: "Critican a enfermera curvilínea por usar uniforme muy ajustado y ella se defiende".³

Las fronteras se diluyen cada vez más. No es fácil distinguir la información periodística. En lugar de marcar la diferencia, muchos medios tradicionales ceden a la tentación de imitar a las redes para ganar más clics. Contribuyen a la confusión como si se hubieran extraviado al salir de sus formatos. Algunos han adoptado incluso la ligereza digital. "Pelean mujeres a golpes por agua", un video de más de un minuto donde la noticia no es la sequía sino la trifulca.⁴ O: "Fue al médico por una infección estomacal, pero era un cáncer colorrectal. Esta es la historia".⁵ ¿Por qué no, si sobra el espacio y es preciso atraer más usuarios?

³ Disponible en <https://www.milenio.com/virales/enfermera-es-criticada-por-usar-uniforme-ajustado-y-se-defiende>

⁴ Disponible en <https://www.youtube.com/watch?v=xLaOwEKrnqM>

⁵ Disponible en https://twitter.com/eltiempo/with_replies?lang=es



©Butcher Billy, cartel para *Black Mirror*, *Be Right Back*, 2016

En la jungla virtual vale todo. Se puede ver el video de una “joven madre de cuatro niños que decidió cortarse las venas” en Facebook, juegos que incluyen violencia en Youtube, un dúo que romantiza el suicidio en tono de hip hop o un clip que sublima la violencia doméstica. La fauna digital incluye *haters* y *trolls*, cuentas falsas y *bots*. No es secreto que en Twitter los *bots* distorsionan la discusión y manipulan las tendencias y las etiquetas (*Hashtags*). Según un análisis de doscientos millones de tweets de la Universidad Carnegie Mellon, el 45 por ciento de los mensajes sobre el Covid-19 provenía de *bots*, al menos durante los primeros seis meses de la pandemia.⁶

⁶ “Casi la mitad de lo tuiteado sobre Covid-19 lo harían cuentas *robots*”. *Voz de América*, 2020. Disponible en https://www.vozdeamerica.com/a/coronavirus_eeuu-estudios-twitter-cuentas-robots-coronavirus/6063285.html

“La ira y el odio son la forma más fácil de crecer en Facebook”, denunció Frances Haugen, exgerente de la empresa que filtró documentos a la prensa en 2021, en lo que se conoce como *Facebook Papers*. La red ha causado daños sin remordimiento. A plena conciencia.

Según *The Washington Post*, en un momento Facebook se planteó eliminar el ícono “Me gusta” al comprobar que puede producir ansiedad y afectar la salud mental de los más jóvenes, pero decidió dejarlo porque, si lo quitaba, los usuarios interactuarían menos con los anuncios. También usó los íconos para impulsar contenidos más emotivos, indignantes incluso. A partir de 2017, para el algoritmo los emojis eran cinco veces más valiosos que los “Me gusta”. Pesaba más el enojo que el agrado. Cuando varios empleados “descubrieron que el algoritmo estaba exacerbando los daños, abogaron por ajustes que podrían ayudar. Pero esas propuestas a veces fueron anuladas”.⁷ Todas las plataformas procuran lo mismo que Facebook: que sus usuarios se mantengan enganchados el mayor tiempo posible sin importar si el contenido es agresivo.

El ideal: la adicción a las redes sociales, el *junkie* digital. La necesidad imperiosa de conectarse varias veces cada día. El miedo a perderse algo que lleva a zambullirse en la pantalla de manera compulsiva es ya tan común que en Estados Unidos se le conoce simplemente como FOMO, por las siglas de *Fear of Missing Out*.

[vozdeamerica.com/a/coronavirus_eeuu-estudios-twitter-cuentas-robots-coronavirus/6063285.html](https://www.vozdeamerica.com/a/coronavirus_eeuu-estudios-twitter-cuentas-robots-coronavirus/6063285.html)

⁷ Jeremy B. Merrill y Will Oremus. “Five points for anger, one for a ‘like’: How Facebook’s formula fortered rage and misinformation”. Disponible en <https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/>

¿Cómo frenar la violencia —artificial y real— en línea? ¿Pueden hacer algo las plataformas para impedir los linchamientos virtuales, aparte de pedir a los usuarios que denuncien? ¿Cuántos bots habrán animado, a principios de agosto, el ensañamiento contra un adolescente de quince años que encabezó por varios días las tendencias de Twitter en México? ¿Hasta dónde llega la responsabilidad de las plataformas por difundir mensajes violentos?

¿Es posible que TikTok se desentienda de haber alojado un peligroso reto que ha costado la vida de varios menores? La Organización Social Media Victim Law Center ha acusado a la red más popular entre los jóvenes de homicidio culposo por la muerte de dos niñas estadounidenses, de ocho y nueve años, que se estrangularon siguiendo el desafío Blackout Challenge.⁸

Durante mucho tiempo, las empresas se han lavado las manos y se han escudado en la libertad de expresión para no intervenir. Las críticas que esto genera han hecho que empleen moderadores y diseñen programas para eliminar cuentas falsas y contenido tóxico. Sin embargo, no es suficiente. El algoritmo de Facebook apenas borró entre el 3 y el 5 por ciento de los mensajes de odio y sólo el 0.6 por ciento de los que vulneraron las reglas de violencia, según el diario *The Wall Street Journal*.⁹

Todo indica que el *laissez faire* no durará mucho. Al menos cuarenta países discuten iniciativas para regular los contenidos violentos con sanciones a usuarios y empresas. En Ja-

pón, un insulto digital puede costar hasta un año de cárcel, mientras que en Inglaterra las plataformas podrían ser multadas con hasta el 10 por ciento de su facturación global.

No ha de ser sencillo controlar los mensajes que emiten miles de millones de usuarios. Nada más en Twitter se publican 350 mil tweets por minuto. Facebook borró mil 300 millones de cuentas falsas el último trimestre de 2021. Pero tampoco parecía fácil la fantasía de la que hablaba en 1994 el profeta Gore. Tal vez no sea posible eliminar el odio, pero sí dejar de amplificarlo intencionalmente para calentar las redes sociales. Las empresas tecnológicas tendrán que enfrentar el desafío de controlar a su propio monstruo de Frankenstein. **U**



©Fabiola Espiga, *Too close*, 2021. Cortesía de la artista

⁸ Disponible en <https://socialmediavictims.org/tiktok-lawsuit/>
⁹ Deepa Seetharaman, Jeff Horwitz y Justin Sheck. "Facebook says AI will clean up the platform. Its own engineers have doubts". Disponible en <https://www.wsj.com/articles/facebook-ai-enforce-rules-engineers-doubtful-artificial-intelligence-11634338184>