

INVESTIGACIÓN

FAKE NEWS: CÓMO LAS PLATAFORMAS ENFRENTAN LA DESINFORMACIÓN

AUTORAS Y AUTORES

Bia Barbosa, Helena Martins e Jonas Valente

GRAPHIC DESIGN AND DIAGRAMMING

Oficina Sal

INTERVOZES - COLECTIVO BRASIL DE COMUNICACIÓN SOCIAL CONSEJO DIRECTOR

Alex Hercog, André Pasti, Bruno Marinoni, Marcos Urupá, Mônica Mourão, Paulo Victor Melo e Tâmara Terso

COORDINACIÓN EJECUTIVA

Gyssele Mendes, Iara Moura, Maria Mello, Marina Pita, Olívia Bandeira e Pedro Ekman

APOYO (MARCA FINANCIADORA)



intervozes.org.br | intervozes@intervozes.org.br







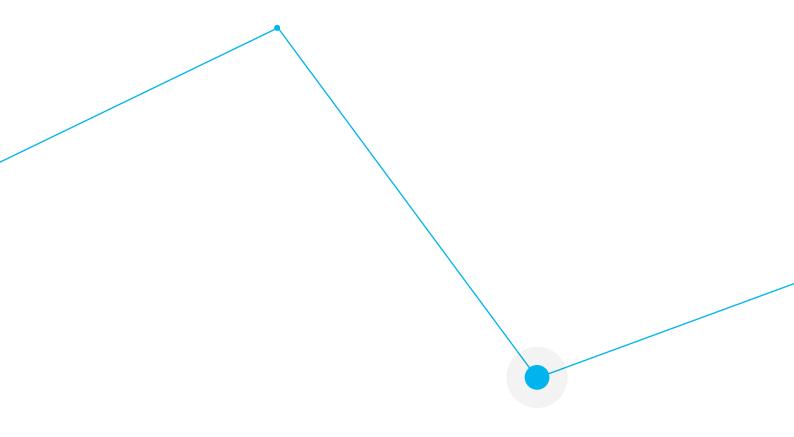








fake cómo las plataformas enfrentan la desinformación **News**



resumen ejecutivo_

Síntesis de la investigación

Esta investigación analiza la desinformación como estrategia de obtención de ventajas políticas y económicas.

Relaciona el fenómeno a la actual configuración de la Internet, marcada por la presencia de monopolios digitales que operan con el objetivo de capturar la atención de los usuarios, recolectar y usar datos personales para la construcción de perfiles y para el direccionamiento de mensajes publicitarios o de propaganda política, por ejemplo.

Considera que el modelo de negocio de las plataformas favorece la ocurrencia de desinformación que, si bien no se inaugura con ellas, ya que también está presente en la historia de los medios tradicionales, como la radio y la TV, se vuelve más constante, amplia, penetrante y de rápida circulación en virtud de las formas de producción, circulación, mediación algorítmica y acceso a la información propias de las plataformas digitales.

Tomando en cuenta lo importante de esto, se estudian las principales plataformas de redes sociales que operan en Brasil: Facebook, Instagram, WhatsApp, YouTube y Twitter.

Se describen las medidas adoptadas por cada una de estas empresas para combatir la desinformación desde 2018, cuando el fenómeno ya era considerado un problema mundial, hasta finales de 2020. Por lo tanto, se cubren también acciones justificadas en el contexto de la pandemia de Covid-19.

Para la descripción y análisis de dichas medidas, se definieron cuatro categorías: 1. Abordaje del fenómeno; 2. Moderación de contenido; 3. Promoción de información y transparencia; y 4. Medidas correlacionadas.

Las medidas se analizan teniendo en cuenta el contexto y las referencias expresadas en las normas de las plataformas existentes sobre desinformación, documentos elaborados por autoridades internacionales y especialistas en derechos humanos, así como propuestas ya elaboradas por el propio Intervozes.

Los cuadros abajo presentan las principales medidas adoptadas por las plataformas digitales en este asunto, incluyendo medidas correlacionadas, que inciden indirectamente sobre la desinformación.

Plataforma

Abordaje del fenómeno

Moderación de contenido

Facebook

No posee una política específica ni tampoco trabaja con una definición propia de desinformación. Presenta las estrategias adoptadas contra el fenómeno de manera resumida en las "Normas Comunitarias". La elaboración de medidas a adoptarse sobre el tema es realizada por el equipo de "Políticas Globales de Contenido". En 2020, creó el Consejo Asesor de Contenido, con miembros externos a la empresa, responsable por la revisión profunda de las denuncias (no solamente sobre desinformación).

Recibe denuncias y utiliza sistemas automatizados para analizar las publicaciones. Reenvía contenidos para agencias de chequeo asociadas y, si el contenido es clasificado como "falso" o "parcialmente falso", recibe una etiqueta específica y pasa a ser acompañado de "artículos relacionados" producidos por chequeadores. Declaraciones y publicaciones de líderes políticos son excepciones y no pasan por procesos de verificación de agencias de chequeo asociadas. Videos e imágenes pueden ser etiquetados como "manipulados", "fuera de contexto" o "falso". No se las elimina de la plataforma, pero se reduce la circulación de publicaciones falsas en los feeds y en los mecanismos de recomendación. La desinformación es removida en casos que motive violencia o daño, que contenga información falsa sobre vacunas o que comprometa procesos electorales. Anuncios con mensajes falsos no pueden ser vehiculados y anunciantes reincidentes pueden sufrir sanciones.

Instagram

Sigue la línea general de abordaje del fenómeno adoptada por Facebook, su controlador. Las Condiciones de Uso determinan que el usuario no puede realizar "actividades ilegales, engañosas o fraudulentas, ni actividades con fines ilegales o no autorizados". No posee estructura institucional o procesos específicos para tratar la desinformación. Son válidos los flujos y procedimientos adoptados por Facebook, incluso los equipos de revisión de contenido y el Consejo Asesor.

Realiza procesos de análisis de contenido a partir de agencias de chequeo que clasifican una publicación utilizando diferentes categorías. Los contenidos son marcados cuando son verificados como desinformativos y pueden tener su circulación reducida (menor visualización en el feed y en las historias, eliminación de la página "explorar" y de las páginas de hashtags). Contenidos relacionados a falsos tratamientos para la Covid-19, sobre teorías conspirativas o falsas afirmaciones registradas como dañinas por las autoridades sanitarias son eliminados, así como anuncios y hashtags que puedan promover desinformación sobre el tema. En la búsqueda sobre el virus, se prioriza información proveniente de autoridades sanitarias.. Durante las elecciones, si un contenido es clasificado como falso, un filtro gris es exhibido sobre la imagen y se alerta al usuario antes que comparta la publicación. Se viene adoptando herramientas de transparencia para anuncios en diferentes países.

WhatsApp

Se presenta como plataforma de mensajes privados criptografiados que no accede, modera, juzga, verifica, bloquea o elimina cualquier contenido, lo que es contradictorio con la existencia de grupos y listas de transmisión que caracteriza el funcionamiento de una red social. No reconoce la frecuente utilización de la plataforma para propagación de desinformación. Coherente con dicho entendimiento, no posee política, procesos y/o estructura institucional ni concepto de desinformación definidos.

Informa oficialmente que no modera contenido. No etiqueta contenidos como desinformativos, por ejemplo. En el campo de la moderación, señalamos solamente que, desde 2018, viene procurando la reducción de la circulación de mensajes masivos, a través del establecimiento de límites para reenvíos. Inicialmente, el límite de envío de una vez fue fijado en 20 chats. Después, 5. Desde la pandemia, los mensajes muy reenviados solo pueden ser reenviados a un contacto a la vez. Sin embargo esta política no está, por lo general, asociada al combate a la desinformación. Esta relación solo fue expresada en el contexto de la pandemia.

Promoción de informaciones y transparencia

Ofrece al usuario un ícono para buscar más información sobre la fuente de la publicación. En el contexto de la Covid-19, presenta información de las autoridades sanitarias en los resultados de la búsqueda (direccionando a los usuarios para el sitio web de la OMS); lanzó el Centro de Información sobre el Coronavirus, ubicado en la parte superior del feed de noticias. Personas que interactúen con contenidos "falsos o dañinos" reciben mensajes. Financia iniciativas de educación mediática, desarrolló materiales y campañas sobre el tema y se asocia con agencias de chequeo para que orientaciones sobre verificación de contenidos estén disponibles. Apoya experiencias periodísticas en pos de ampliar la circulación de noticias profesionales en la plataforma. Pone a la disposición datos para 60 investigadores en todo el mundo.

Medidas correlacionadas

Adoptó una política de prohibición del discurso de odio en 2018 y mejoró los sistemas automatizados para identificar videos violentos, restringiendo ciertas transmisiones en vivo. Las búsquedas relacionadas con la supremacía blanca, como la exaltación del nazismo, reciben sugerencias de materiales en contra del discurso de odio. En 2020, comenzó a eliminar contenido que negaba el holocausto. Ese mismo año, también se prohibió la actividad del grupo QAnon, propagador de teorías conspirativas. La plataforma también elimina cuentas falsas, propagadoras de spam o consideradas de "comportamiento coordinado no auténtico". Busca verificar la autenticidad de las cuentas en el momento de su apertura y mapea señales de conductas maliciosas en cuentas recién abiertas. También reduce la circulación de cuentas que difunden contenido de baja calidad (clickbait o que llevan para sitios con anuncios maliciosos e impactantes). Pasó a restringir el acceso a la recolección de datos por aplicaciones de terceros. Como medida de protección a la integridad electoral, proporciona un archivo de anuncios políticos con información sobre la identidad de los anunciantes, pagos y alcance de los contenidos, además de etiquetar estas piezas como "propaganda electoral". Usuarios pueden elegir si desean o no recibir este tipo de anuncio.

Sistemas informáticos buscan "correspondencia de imágenes" para etiquetar contenidos ya clasificados como falsos, incluso cuando fueron inicialmente publicados en Facebook (y viceversa). Proporciona textos que ofrecen un contrapunto elaborados por verificadores, con "desmentidos" o informaciones oficiales sobre el tema. Búsquedas sobre vacunas direccionan a los usuarios al sitio web de la OMS. En el contexto de la Covid-19, hizo lo mismo y pasó a proporcionar información del órgano en la parte superior del feed de los usuarios, además de ofrecer imágenes sobre medidas de prevención para compartir en las historias.

Las Normas Comunitarias condenan el discurso de odio, los mensajes no deseados enviados repetidamente y los ataques a las personas con la intención de generar incomodidad. Las cuentas falsas se eliminan, sobre todo en el momento de su apertura. Las cuentas cuya identidad de la persona responsable se puede verificar, pueden recibir una señal de su verificación. Las cuentas o redes de cuentas que entran en la categoría de "comportamiento no auténtico", se eliminan. Las cuentas que registran cierto número de infracciones en un período de tiempo pueden desactivarse, con derecho a apelar. Se prohíbe la promoción, oferta o intercambio de evaluaciones de usuarios falsas; y se eliminan los "me gusta", los seguidores y los comentarios no auténticos para "aumentar la popularidad". Se ha ampliado las medidas contra la intimidación y el acoso, brindando un mayor control sobre las interacciones de los usuarios, y sin mostrar el número de "me gusta" en las publicaciones.

En la aplicación o en el blog oficial, es posible encontrar referencias a la desinformación, pero las informaciones son dispersas y poco accesibles a un usuario común, aunque la plataforma tiene la posibilidad de contacto directo, a través de notificaciones. Después de las elecciones de 2018, se establecieron alianzas institucionales para la verificación de contenidos y la realización de investigaciones, pero el alcance y la efectividad de dichas medidas son cuestionables. Asimismo, las acciones de sensibilización del público en general han sido restringidas. Las medidas de promoción de información fueron ampliadas en la pandemia, con la creación de canales de mensajes. En este sentido, para las elecciones de 2020, fue firmado un acuerdo con el Tribunal Superior Electoral (TSE) y fue creado un chatbot para denuncias y figuritas sobre voto consciente. El usuario tiene que añadir el número de la cuenta y entrar en contacto. Falta transparencia respecto a las prácticas adoptadas por WhatsApp, como la prohibición de creación o eliminación de cuentas, además de informes sobre sus medidas.

La gran parte de las acciones desarrolladas por WhatsApp se inscriben como "medidas correlacionadas" empezando con la definición de usos que involucran declaraciones falsas, incorrectas o engañosas como violaciones de sus Términos de Servicio. Además de la ya mencionada limitación de reenvíos de mensajes, la plataforma marca los mensajes altamente reenviados con una o dos flechas, señalando una mayor viralización. Fue eliminado también el botón que permitía reenvíos rápidos de contenidos y se añadió una lupa al lado de los mensajes marcados con dos flechas. Utiliza herramientas para el manejo de spam y aprendizaje avanzado de máquinas para eliminar mensajes masivos automatizados y prohibir cuentas de usuarios por comportamiento inadecuado, como enviar mensajes masivos y crear múltiples cuentas.

Plataforma

Abordaje del fenómeno

Moderación de contenido

YouTube

No trabaja con una única definición de desinformación y no posee una política específica para el tema. Sin embargo, restricciones a la difusión de desinformación aparecen en diferentes políticas, con respecto a daños y materiales digitales manipulados. Contenidos que violen las Normas de la Comunidad son eliminados y reciben una notificación. Casos reincidentes sufren restricciones de publicación y monetización y el canal puede ser excluido. Se anima a la comunidad que denuncie contenidos inapropiados y equipos internos analizan los casos. Inauguró recientemente la página "Mitos y Hechos sobre Desinformación", que suma las medidas que adopta contra el problema.

Saca del aire contenidos que violan sus directrices y reduce el alcance de "desinformación dañina" y contenido limítrofe", dejando de recomendarlos. Videos editados y adulterados son prohibidos por la política de "material manipulado", así como canales que falsean su identidad. Elimina contenidos y canales que no respeten la política con respecto a las elecciones, incluyendo aquellos videos que busquen socavar la confianza en las votaciones. Pasó a eliminar informaciones engañosas sobre la Covid-19, sobre todo desinformación médica, promoción de medicamentos peligrosos e información sobre el origen del virus. Adoptó una fuerte política contra videos anti-vacuna. Recomienda a los productores que chequeen los hechos y que utilicen como fuentes a las autoridades sanitarias. No toda la publicidad relacionada a la pandemia fue restringida y los videos que respeten las normas de la comunidad, pueden recibir anuncios. Decisiones con respecto al cumplimiento de las directrices son tomadas a partir del análisis del contenido e independientemente de quien habla o publica dichos materiales, incluso altas autoridades. Posee un Grupo de Análisis de Amenazas para identificar prácticas de desinformación patrocinadas por gobiernos.

Twitter

No posee una política para validar la autenticidad de contenidos, tampoco trabaja con una definición del término desinformación. Su foco es ofrecer contexto para ayudar a las personas a decidir el contenido de un post o publicación denunciado, actuando directamente solamente en casos de posibles daños causados por materiales manipulados, informaciones engañosas sobre procesos electorales y, recientemente, relacionadas a la Covid-19. Prohíbe la utilización de robots manipuladores de la red. Tiene un Consejo de Confianza y Seguridad que también analiza el tema.

Materiales manipulados, incluyendo los deepfakes, son etiquetados, y pueden tener su visibilidad reducida, recibir un enlace con más explicaciones o, en caso de riesgo de daños graves e inmediatos, pueden ser eliminados. Contenidos que pueden engañar a las personas sobre cuándo, dónde y cómo votar también están prohibidos como parte de la política de integridad en las elecciones, que en los EEUU etiqueta las publicaciones de candidatos. Informaciones falsas con el objetivo de socavar la confianza del público en las elecciones pueden ser etiquetadas o eliminadas. Anuncios políticos y de medios de comunicación estatales ya no son permitidos. Monitorean tendencias y picos en las conversaciones sobre elecciones a fin de detectar actividades de manipulación y anular cuentas falsas. Publicaciones de líderes globales, aunque sean desinformativas, pueden seguir en el aire si son de evidente interés público. La publicación es exhibida por detrás de un aviso y no es posible la interacción. Tras la Covid-19, se amplió la definición de "dañino" para que se incluya contenidos contrarios a las recomendaciones de las autoridades sanitarias y eliminan publicaciones que de forma clara inciten a acciones riesgosas. Bajo riesgo de daño menos severo, la publicación recibe una etiqueta o alerta. La política de anuncios solo se autoriza para servicios públicos y privados que cambiaron con la Covid-19.

Promoción de informaciones y transparencia

Prioriza lo que llama "voces autorizadas", como canales periodísticos, en las búsquedas de información sobre eventos y política, y en los "próximos videos" sobre estos temas. Si el usuario busca contenido periodístico, muestra una estante con "Principales noticias". La estante "Últimas noticias" se muestra para eventos importantes. Muestra paneles de información con datos contextuales de fuentes autorizadas sobre asuntos históricos, científicamente comprobados y teorías conspirativas. Dependiendo de la búsqueda, también se muestran chequeos por editores independientes No se chequea cada video por separado. La Covid-19 obtuvo sesiones especiales con todas esas herramientas. En las elecciones de EEUU, ofrece información adicional sobre candidatos. En Brasil, hizo transmisiones en vivo sobre desinformación con el TSE. Vía Google News Initiative, apoya el periodismo digital en el combate a la desinformación y acciones de educación para los medios de comunicación.

Medidas correlacionadas

Tiene su propia política contra el discurso de odio y otra contra el acoso y las amenazas. Los videos supremacistas están prohibidos y aquellos que no se pueden caracterizar como de odio pero que se acercan a ellos tienen restricciones de participación, recomendación y monetización. Se muestra una alerta antes de que comience. En 2020, comenzó a eliminar contenido relacionado con teorías conspirativas que resultan en actos de violencia. Cualquier persona o canal que intimide, aceche, deshumanice o fomente un comportamiento violento, incluso a través de comentarios, puede ser excluido de la red. También tiene una política para atender temas como la privacidad y contra la difamación de personas, para recibir quejas. No autoriza la distribución de contenido peligroso o dañino y tiene una política contra el spam, falsificación de identidad y la participación/engagement falso. Restringe el acceso al servicio a través de cualquier medio automatizado (bots).

Posee la herramienta #ConozcaLosHechos, que se presenta en búsquedas relacionadas a vacunas y Covid-19. Información no fidedigna sobre salud no aparece en las búsquedas, las cuales priorizan respuestas de fuentes oficiales. Añadió una pestaña especial sobre la pandemia en la función #Explorar, con fuentes creíbles comunicados públicos y periodistas de salud. Aplica etiquetas en publicaciones sobre Covid-19 y 5G, alentando la verificación de la información y, en el recurso "Eventos", también selecciona informaciones seguras y las deja disponibles en la parte superior de las líneas de tiempo. Aumentó el número de cuentas verificadas en este sentido. Abrió la plataforma para que desarrolladores e investigadores acompañen en tiempo real las publicaciones sobre la Covid-19. Crea páginas especiales sobre las elecciones con información oficial y verosímil. Apoya iniciativas de periodismo independiente, chequeo de hechos y educación mediática, incluso a través de créditos en anuncios.

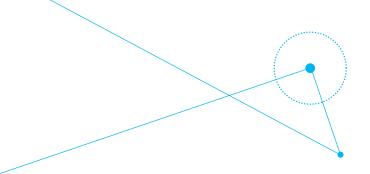
Su política contra la propagación del odio prohíbe contenidos con amenazas violentas, que promuevan o expresen el deseo de muerte o agresión, la difusión de estereotipos, el miedo y el acoso. Tampoco se permiten comportamiento que se considere abusivo. Los infractores pueden ser obligados a eliminar contenidos, pasar un tiempo sin publicar ni interactuar, o ser excluidos por completo. Combate el spam y la automatización maliciosa al monitorear y desafiar la actividad sospechosa de la cuenta, como mediante la verificación de identidad. La atención se centra en detener las cuentas falsas, el engagement no auténtico y las actividades coordinadas para influir artificialmente en las conversaciones. A los perfiles sospechosos se les reduce la visibilidad y se realizan auditorías en las cuentas existentes. Publica archivos de publicaciones de actividades coordinadas apoyadas por los gobiernos para dar transparencia a los intentos de manipular la red.

conclusiones

El estudio concluye que las plataformas digitales investigadas no presentan políticas y procesos estructurados sobre el problema de la desinformación y desarrollan acciones puntuales y reactivas en el combate al fenómeno. En general, evitan el análisis del contenido desinformativo, pero terminan haciéndolo. Ninguna de las empresas informó haber trabajado con un concepto unificado de desinformación. Además de menciones a ella, se constató la utilización de términos diversos, como noticias falsas e informaciones engañosas. A pesar de la polisemia del concepto y el debate intenso al respecto tanto en la sociedad como en los círculos académicos, la explicitación de un entendimiento sería un avance en la transparencia sobre qué tipo de contenido puede ser impactado por las medidas de análisis, alertas o sanciones. No abordarlo directamente no disminuye el problema y abre espacio para que otros lo hagan, con decisiones unilaterales y poco claras. Ninguna de las empresas, además, reportó tener una estructura específica para abordar la cuestión de la desinformación, lo que puede dificultar la coordinación de las iniciativas en el ámbito de cada organización.

En cuanto a la moderación de contenidos desinformativos, se observa un aumento de decisiones automatizadas por parte de las plataformas, pero no queda claro, incluso en las explicaciones dadas por las corporaciones para esta investigación, los criterios para la utilización de personas naturales en el análisis y decisión acerca de las medidas de moderación del contenido. El chequeo de contenidos, principalmente por agencias externas, es una práctica presente en la mayoría de plataformas. Cuando ocurren, los chequeos se realizan según las categorizaciones de la plataforma, pauta necesaria para evitar diferentes tratamientos por parte de cada chequeador. Aun así, naturalmente, la complejidad al analizar las "tonalidades de gris" entre uno y otro extremo conlleva riesgos de valoraciones cuestionables, por lo que el chequeo debe contar con mecanismos efectivos de debido proceso para mitigar abusos y errores. En las plataformas donde se comprueba el contenido, se utilizan etiquetas (flagged). Agentes políticos y académicos señalaron la ineficacia de este mecanismo. Los estudios han demostrado que las alertas incluso llaman la atención sobre el contenido. Aun con este "efecto secundario", parece ser una herramienta necesaria.





Durante la mayor parte del período analizado por la investigación, las plataformas se resistieron a eliminar el contenido desinformativo

La moderación también cubre anuncios y contenido promocionado. En tales casos, existen restricciones sobre el contenido que ya se considera desinformativo. La extensión de las medidas de desinformación a la publicidad es fundamental. Esto se debe a que ahí se ubica la dimensión económica de la práctica de producir y difundir noticias falsas como negocio, explotada por las "fábricas de contenidos engañosos" y también por las propias plataformas, que se benefician del cobro de publicaciones pagas. En el caso de anuncios con discursos de políticos, las medidas se topan con las excepciones al discurso de estos agentes adoptadas por algunas compañías.

Durante la mayor parte del período analizado por la investigación, las plataformas se resistieron a eliminar el contenido desinformativo como lo hacen en otras categorías, en base a lo que se define unilateralmente en sus pautas. Pero esta actitud empezó a cambiar en 2020, en el contexto de la nueva pandemia de coronavirus, con la propagación de desinformación sobre la enfermedad y "formas milagrosas" de cura. Los graves riesgos para la salud de la población han presionado a las plataformas para que respondan de manera más rápida y dura. En este escenario, la admisión de situaciones excepcionales de eliminación de contenido en casos de evidente riesgo o grave daño parece una posibilidad razonable, siempre y cuando esté conectado a las normas del debido proceso que permitan la impugnación, la evaluación de los recursos por parte de las personas y la reparación en caso de error en la moderación aplicada, lo que no se garantiza en ninguna de las plataformas analizadas.

La reducción del alcance, una medida más presente en casos de desinformación, termina por asumir casi una condición de exclusión y aparece como una medida de alto impacto en los contenidos, pero que aún no ha demostrado plenamente su eficacia. Los estudios mencionados en la investigación apuntan a una reducción en la velocidad de propagación de la desinformación, pero la limitación del alcance no impide que circule dicho contenido desinformativo. Es necesario profundizar las investigaciones sobre el tema, con énfasis en investigaciones que efectivamente puedan enfocar el funcionamiento de grupos profesionales que utilizan plataformas con sistemas que permitan eludir estas limitaciones.

Parte de las plataformas analizadas aporta más información

en el caso de contenido verificado. Un tipo de acción de este tipo es la visualización de artículos de las propias agencias de verificación con la información obtenida sobre un tema determinado. Otro es el suministro de información "del contexto" o "fuentes confiables". También se puede proporcionar contenido de medios periodísticos y oficiales. En el contexto de la pandemia antes mencionado, se han implementado varios mecanismos para acceder y recomendar información oficial sobre el tema, especialmente de la Organización Mundial de la Salud y las autoridades sanitarias nacionales.

Desde el punto de vista de la información sobre cómo lidiar con los contenidos desinformativos, se destaca la baja transparencia de las plataformas. Las medidas existentes no se presentan de forma organizada y buena parte del trabajo de esta investigación fue precisamente la búsqueda y organización de información dispersa en noticias de sitios oficiales, sitios de "ayuda" o FAQ de cada empresa - la excepción fue You-Tube, que recién en octubre de 2020 comenzó a dedicarle una página al tema. La falta de reconocimiento de la gravedad del problema y la implicación de las plataformas para enfrentarlo termina creando obstáculos para que la comunicación con los usuarios sobre el tema sea efectiva. Sin embargo, el elemento más problemático en términos de transparencia es quizás la falta de equilibrio de las acciones anunciadas. Una vez más, la información es escasa y dispersa, por lo que la medición de la concreción de las acciones se restringe a las afirmaciones de las empresas, acercándose más a promesas no comprobables. Tampoco se evalúa la efectividad de lo implementado o el debido proceso, mecanismo que podría permitir que los usuarios sean notificados y se defiendan en procesos de moderación de contenidos o en eventuales sanciones a sus publicaciones o cuentas aplicadas por las plataformas.

Concluimos que las acciones aún carecen de organización, solidez, transparencia y evaluación. El problema de la desinformación necesita ser reconocido, comunicado y afrontado de manera efectiva por las plataformas, lo que implica la revisión de la estructura y modelo de negocio de estas empresas, bajo el riesgo de seguir ofreciendo curas que no logran interrumpir una forma de comunicación que hoy se ha convertido en uno de los principales desafíos para las democracias en todo el mundo.

Síntesis de las medidas adoptadas por las plataformas digitales (2018-2020)











Facebook

Instagram

WhatsApp

YouTube

Twitte

ABORDAJE DEL FENÓMENO	Política específica y estructurada de combate a la desinformación					
	Concepto de desinformación adoptado y definido					
	Mención a la desinformación en directrices o políticas			Х	Х	Х
BORDAJI	Medidas correlacionadas que inciden sobre la desinformación	Х	Х	Х	Х	Х
⋖	Estructura institucional específica para el tema					
	Verificación de contenidos desinformativos por equipo propio				Х	X
	Verificación de contenidos desinformativos por terceros	X	X	Х		Х
*	Eliminación de contenidos desinformativos				Х	Х
ONTENIE	Reducción del alcance	Х	Х		Х	Х
N DEL CO	Desmonetización y medidas restrictivas en anuncios y contenidos promovidos	Х			Х	Х
MODERACIÓN DEL CONTENIDO*	Suspensión de cuentas por desinformación				Х	Х
	Recurso para denunciar la desinformación				Х	X
	Notificación de usuarios en caso de sanción por desinformación	Х	X		Х	X
	Recurso a la plataforma en caso de sanción por desinformación				Х	Х
	I .					



Facebook



Instagram



WhatsApp



YouTube



Twitter

		. accoon	stag.a	TTTT CO. TP P	1041450	
	Alertas de contenidos desinformativos	Х	Х			Х
	Disponibilidad de la información sobre verificación de contenido	X	Х	Х		Х
	Disponibilidad de contenido contextual o fuentes oficiales	Х	X	Х	X	Х
	Medidas para la transparencia de anuncios	Х	Х		Х	Х
	Canales de información sobre medidas contra la desinformación					
	Balance de las medidas adoptadas contra la desinformación en los informes de transparencia					
	Iniciativas de educación mediática	X	Х	Х	Х	Х
	Proyectos de apoyo al periodismo profesional	X	Х	Х	X	Χ
	Prohibición de cuentas no auténticas	X	Х		Х	Χ
	Restricción a cuentas y herramientas automatizadas			Х	X	Χ
	Prohibición del discurso de odio	X	Х		X	Χ
	Restricción de materiales manipulados	X	Х		Х	Χ
	Restricción a la circulación de mensajes			Х		
+						

