

Traducción
Respuestas de la política de la OCDE al coronavirus (COVID-19)
Combatir la desinformación del COVID-19 en las plataformas en línea

La crisis de COVID-19 surge de una "infodemia" y una pandemia

A medida que continúa la pandemia de coronavirus (COVID-19), también lo hace el volumen de información de noticias relacionadas, lo que lleva a lo que se ha llamado una "infodemia". Si bien los temas de información errónea (la difusión de información falsa, independientemente de si hay una intención de engañar) y la desinformación (la difusión deliberada de información falsa o engañosa con la intención de engañar) han sido ampliamente estudiados y discutidos, COVID-19 arroja en gran alivio los desafíos y las consecuencias de esta importante área política.

La desinformación y la información errónea sobre COVID-19 se difunde rápida y ampliamente en Internet, llegando a muchas personas y posiblemente influyendo en ellas. No existe una fuente única, con diferentes actores impulsados por motivos muy diferentes que producen y propagan información falsa y engañosa para promover sus propios objetivos y agendas. Por ejemplo, algunas personas están utilizando plataformas en línea para difundir teorías de conspiración, afirman que COVID-19 es un arma biológica extranjera, una farsa partidista, el producto de la tecnología 5G o parte de un plan mayor para rediseñar la población. Otros están difundiendo rumores de curas supuestamente secretas, como beber cloro diluido, comer plátanos o apagar la electrónica. Otros están utilizando la pandemia COVID-19 para beneficio financiero, vendiendo kits de prueba, máscaras y tratamientos sobre la base de afirmaciones falsas o engañosas sobre sus poderes preventivos o curativos.

No se pueden exagerar los efectos nocivos de la desinformación sobre COVID-19. Los datos de Argentina, Alemania, Corea, España, Reino Unido y Estados Unidos muestran que aproximadamente una de cada tres personas dice haber visto información falsa o engañosa relacionada con COVID-19 en las redes sociales. La investigación también ha demostrado que la desinformación de COVID-19 se difunde significativamente más ampliamente que la información sobre el virus de fuentes autorizadas como la Organización Mundial de la Salud (OMS) y los Centros para el Control y la Prevención de Enfermedades de los Estados Unidos. Al cuestionar las fuentes y los datos oficiales y convencer a las personas para que prueben tratamientos falsos, la difusión de información y desinformación ha llevado a las personas a ingerir curas fatales para el hogar, ignorar las reglas de distanciamiento social y encierro, y no usar máscaras protectoras, lo que socava la efectividad de estrategias de contención.

Sin embargo, los efectos nocivos de la desinformación van más allá de los problemas de salud pública. Por ejemplo, en el Reino Unido, la falsa afirmación de que las ondas de radio emitidas por las torres 5G hacen que las personas sean más vulnerables a COVID-19 ha resultado en más de 30 actos de incendio premeditado y vandalismo contra equipos e instalaciones de telecomunicaciones, así como alrededor de 80 incidentes de acoso contra técnicos de telecomunicaciones. En los Países Bajos, se han registrado 15 actos de incendio premeditado. Del

mismo modo, en Australia, la Unión Europea y los Estados Unidos, la difusión de la desinformación que enmarca a las minorías como la causa de la pandemia ha alimentado la animosidad contra los grupos étnicos, motivando un aumento de la discriminación y los incidentes de violencia en lo que se ha calificado como "coronaracismo".

Comprender cómo se propaga la desinformación de COVID-19 es esencial para elaborar respuestas efectivas

Las personas acceden cada vez más a información, noticias y contenido a través de Internet en sitios de agregadores de noticias, plataformas de redes sociales y plataformas para compartir videos. A diferencia de los medios tradicionales como las noticias de televisión o los periódicos, las plataformas en línea automatizan y personalizan el contenido al aprovechar los datos relacionados con la actividad en línea pasada de los usuarios, las conexiones sociales, la ubicación, etc. tienden a generar más compromiso del usuario. Como resultado, los algoritmos de personalización de contenido pueden exponer repetidamente a las personas al mismo contenido o anuncios similares o incluso sobre la base de la desinformación. Además, algunas plataformas presentan contenido de noticias junto con noticias, anuncios y contenido generado por usuarios. Esto puede dificultar que los usuarios distingan noticias confiables. A menos que se implementen medidas para eliminar, eliminar el énfasis o limitar la circulación de información falsa y engañosa, puede propagarse rápidamente.

Otra consideración importante es el alcance de la influencia ejercida por diferentes fuentes de desinformación. La desinformación de COVID-19 se mueve de arriba hacia abajo, de políticos, celebridades y otras figuras prominentes, así como de abajo hacia arriba, de la gente común. Sin embargo, el impacto de estas dos fuentes de desinformación difiere dramáticamente. La investigación empírica muestra que la desinformación de arriba hacia abajo constituye solo el 20% de todas las afirmaciones engañosas sobre COVID-19, pero genera el 69% del compromiso total de las redes sociales. Por el contrario, si bien la mayoría de la desinformación de COVID-19 en las redes sociales se crea de abajo hacia arriba por los usuarios comunes, la mayoría de estas publicaciones parecen tener una participación significativamente menor. En consecuencia, las figuras públicas influyentes tienen más responsabilidad en los esfuerzos para abordar los reclamos de desinformación y los rumores sobre COVID-19.

Iniciativas de plataformas en línea para abordar la desinformación de COVID-19

A medida que los científicos y los gobiernos continúan trabajando en tratamientos y vacunas, es imperativo que las plataformas en línea, los gobiernos y las organizaciones de salud nacionales e internacionales trabajen juntas para frenar la difusión de información falsa y engañosa sobre COVID-19.

Algunas plataformas han tomado medidas importantes, como dirigir a los usuarios a fuentes oficiales cuando buscan información de COVID-19, prohibir anuncios de máscaras médicas y respiradores, y reforzar sus esfuerzos para detectar y eliminar contenido falso, engañoso y potencialmente dañino relacionado con COVID-19, incluso finalizando las tiendas en línea o

eliminando listados que hacen afirmaciones falsas o engañosas sobre productos que previenen o curan COVID-19. Además, Facebook, Google, LinkedIn, Microsoft, Reddit, Twitter y YouTube han publicado una declaración conjunta sobre su colaboración con las agencias gubernamentales de salud para combatir el fraude y la desinformación sobre COVID-19. En general, existen tres tipos principales de esfuerzos de colaboración entre plataformas y autoridades de salud pública:

Destacando, destacando y priorizando contenido de fuentes autorizadas. Plataformas como Facebook, Instagram, TikTok y Pinterest están redirigiendo a los usuarios a información de la OMS en respuesta a búsquedas de información y hashtags asociados con COVID-19. Del mismo modo, Google lanzó un micrositio COVID-19 de ventanilla única y una "Alerta SOS", que dirige a las personas que buscan "coronavirus" a noticias y otro contenido de la OMS. YouTube presenta videos de agencias de salud pública en su página de inicio y resalta contenido de fuentes autorizadas en respuesta a búsquedas de información sobre COVID-19. Twitter, a su vez, presenta una página de evento COVID-19 con la información más reciente de fuentes confiables además de los plazos de los usuarios. Snapchat también se ha asociado con la OMS para crear filtros y calcomanías que brindan orientación sobre cómo prevenir la propagación del virus.

Cooperación con verificadores de hechos y autoridades sanitarias para señalar y eliminar la desinformación. Facebook coopera con verificadores de hechos de terceros para desacreditar rumores falsos sobre COVID-19, etiquetar ese contenido como falso y notificar a las personas que intentan compartir dicho contenido que ha sido verificado como falso. Facebook se asoció con la International Fact-Checking Network (IFCN) para lanzar un programa de subvenciones de USD 1 millón para aumentar su capacidad y ha estado eliminando contenido marcado por las autoridades de salud pública, incluyendo "reclamos relacionados con curas falsas o métodos de prevención, como beber curas de cloro el coronavirus, o afirmaciones que crean confusión sobre los recursos de salud disponibles".

Del mismo modo, los informes señalan que Google donó USD 6,5 millones a los verificadores de datos que se centran en el coronavirus. A su vez, Twitter amplió la definición de daño en su plataforma para abordar el contenido que va directamente en contra de la orientación de fuentes autorizadas de información de salud pública global y local.

Ofrecer publicidad gratuita a las autoridades. Facebook, Twitter y Google han otorgado créditos publicitarios gratuitos a la OMS y autoridades sanitarias nacionales para ayudarlos a difundir información crítica sobre COVID-19.

No hay una manera fácil de resolver el problema de desinformación de COVID-19 de manera completa o permanente

Las iniciativas anteriores son importantes, ya que ambas facilitan el acceso a información precisa y confiable sobre COVID-19 y posiblemente salvan vidas. Sin embargo, solo pueden llegar tan lejos. Dos desarrollos resaltan que la crisis de desinformación de COVID-19 no tiene una solución fácil:

- **Mayor confianza en la moderación de contenido automatizada.** Como resultado de las medidas de bloqueo y pandemia, las plataformas en línea como Facebook, Google y Twitter se han enfrentado a una escasez de moderadores humanos y, en consecuencia, han aumentado su dependencia de las tecnologías de monitoreo automatizado para marcar y eliminar contenido inapropiado. Sin embargo, estas tecnologías no pueden atrapar cada instancia de desinformación. Además, actuar como moderador de contenido puede tener implicaciones legales para las plataformas bajo la legislación de comunicaciones existente.¹ En consecuencia, para reducir la probabilidad de perder importantes reclamos y rumores engañosos o falsos, tienden a ser programados para errar por precaución. Eso aumenta el riesgo de falsos positivos. Por lo tanto, YouTube ha notado que, como consecuencia de una mayor dependencia de los sistemas de moderación automatizados, "los usuarios y creadores pueden ver mayores eliminaciones de videos, incluidos algunos videos que pueden no violar [sus] políticas". De hecho, ha habido múltiples incidentes reportados de sistemas de monitoreo automatizados que marcan el contenido de COVID-19 de fuentes acreditadas como spam. Por lo tanto, existe el riesgo de que, para suprimir la desinformación de COVID-19, la moderación del contenido sin una supervisión humana adecuada pueda limitar la disponibilidad y la accesibilidad a información confiable sobre COVID-19. Además, sin una adecuada transparencia, responsabilidad y debido proceso, es más probable que la moderación de contenido automatizada tenga un efecto escalofriante en la libertad de expresión, agregando así una capa de complejidad a la infodemia.
- **Prohibición de anuncios que explotan la crisis.** Facebook e Instagram prohibieron los anuncios que sugieren que un producto es una cura garantizada o que evita que las personas contraigan COVID-19, así como anuncios y listados comerciales de máscaras, desinfectantes de manos, toallitas desinfectantes de superficies y kits de prueba COVID-19. Twitter implementó medidas comparables bajo su política de Contenido inapropiado. Del mismo modo, Google y YouTube prohíben cualquier contenido, incluidos los anuncios, que busque capitalizar la pandemia, y sobre esta base han prohibido los anuncios de equipos de protección personal. Estas medidas disuaden a los estafadores que buscan vender productos falsos relacionados con COVID-19 y ayudan a proteger a los consumidores contra el aumento de precios, pero también pueden hacer que sea un poco más difícil para las personas encontrar y comprar productos de higiene en línea. En cualquier caso, la desinformación y la información errónea siguen siendo difíciles de eliminar por completo sin infringir la libertad de expresión y / o los derechos de privacidad de los usuarios. Además, algunas plataformas en línea y productores de contenido aún se benefician de los anuncios que se muestran en el contenido engañoso y falso que escapa a los esfuerzos de moderación de las plataformas. Por ejemplo, el Proyecto de Transparencia Tecnológica encontró varios videos engañosos de coronavirus, curas prometedoras y tratamientos preventivos, en una importante plataforma para compartir videos con anuncios de varios anunciantes, y EUvsDiSiNFO concluyó recientemente que "las principales plataformas continúan monetizando la desinformación y el contenido dañino en la pandemia [...] p.ej al alojar anuncios en línea en páginas que tergiversan a los migrantes como la causa del virus, promueven curas falsas o difunden teorías de conspiración sobre el virus ". Además, después de realizar experimentos en varias plataformas de redes sociales, el DisinfoLab de la UE concluyó que "el proceso de moderación

de [...] contenido conspirador desacreditado ha sido lento e inconsistente en todas las plataformas [...]".

Un primer paso necesario hacia una solución más efectiva a largo plazo es reunir evidencia sistemáticamente. Las actualizaciones periódicas de transparencia de las plataformas en línea sobre la desinformación de COVID-19 que se muestra y se ve, y cómo las plataformas la detectan y moderan, ayudarían a los investigadores, a los encargados de formular políticas y a las propias plataformas a identificar formas de realizar mejoras. La Comisión Europea pidió recientemente a las plataformas que emitan voluntariamente dichos informes de transparencia mensualmente. Si eso se convirtió en un requisito, y particularmente si creció para incluir otras jurisdicciones, la cooperación internacional oportuna y de múltiples partes interesadas y la coordinación en un estándar común de informes podría ser muy valioso para reducir las ineficiencias y los gastos. Además, un enfoque común entre empresas y países podría facilitar las comparaciones globales y multiplataforma. La OCDE es un foro ideal para ese tipo de proyecto y ya está liderando ese esfuerzo con respecto a la transparencia voluntaria que informa sobre el contenido terrorista y extremista violento en línea.

Es probable que la mejora también deba incluir plataformas en línea que asuman un papel más vigilante invirtiendo en más y mejor investigación y verificación de hechos. Esto podría implicar la contratación de más revisores, brindar apoyo adicional a múltiples organizaciones externas de verificación de hechos, continuar desarrollando sistemas de moderación de contenido automatizados y priorizar aún más la calidad y el contenido confiable. En este escenario, el contenido falso y engañoso podría continuar subiéndose, pero la velocidad y la magnitud de su difusión se reducirían considerablemente.

Sin embargo, las plataformas no pueden estar solas en este esfuerzo. Los gobiernos, las organizaciones de salud pública, las organizaciones internacionales, la sociedad civil, las organizaciones de medios y las empresas tecnológicas deben unir fuerzas para combatir la desinformación. Eso incluye trabajar juntos para mejorar las habilidades de alfabetización mediática, digital y de salud de las personas. Si los usuarios pueden identificar con precisión la fuente de lo que leen, quién lo escribió o pagó, y por qué se les muestra la información que ven, es menos probable que sean persuadidos y manipulados por afirmaciones y rumores engañosos.

Recomendaciones clave de política

La difusión de la desinformación de COVID-19 es un tema complejo que garantiza la cooperación, la coordinación y la confianza entre las plataformas en línea, los gobiernos y las organizaciones de salud nacionales e internacionales, así como una combinación de medidas cuidadosamente equilibradas. Aunque ya se están dando varias respuestas, se puede hacer más. Las siguientes consideraciones están destinadas a guiar a los interesados en sus esfuerzos para llegar a soluciones cada vez más efectivas para la desinformación de COVID-19.

Apoye una multiplicidad de organizaciones independientes de verificación de hechos. La desinformación sobre el coronavirus cuestiona las acciones, recomendaciones y competencia de

gobiernos, médicos, científicos y organizaciones de salud nacionales e internacionales. Debido a que daña su credibilidad, la desinformación también obstaculiza su capacidad de desacreditar o refutar información falsa y engañosa, incluida la información que los arroja en una luz negativa en primer lugar. Los verificadores de hechos independientes pueden proporcionar un análisis imparcial de la información mientras ayudan a las plataformas en línea a identificar contenido engañoso y falso; Los gobiernos y las autoridades internacionales pueden ayudar apoyando y confiando en sus análisis para restaurar la confianza pública. La mayor dependencia de Facebook y otras plataformas y la asistencia financiera para los verificadores de datos son ejemplos a seguir no solo por otras compañías tecnológicas, sino también por el sector público. Las plataformas también podrían considerar etiquetar el contenido que ha pasado con éxito las verificaciones de hechos por dos o más organizaciones independientes de verificación de hechos con una marca de confianza. Se ha demostrado que tales marcas de confianza son efectivas para aumentar la confianza de los consumidores en el contexto del comercio electrónico, por lo que también podrían ser una herramienta útil para aumentar la confianza en la información en línea sobre COVID-19.

Asegurar que existan moderadores humanos para complementar las soluciones tecnológicas.

Mientras que los sistemas de monitoreo automatizados son una herramienta importante para detectar y eliminar la desinformación de COVID-19, la moderación de contenido también requiere intervención humana, especialmente cuando se requieren decisiones más matizadas. Este es un problema complicado sin una solución fácil, particularmente durante la pandemia, porque si bien enviar moderadores de contenido de regreso al trabajo demasiado pronto sería un riesgo inaceptable para la salud pública, hacer que trabajar desde casa genere problemas de privacidad y confidencialidad. Sin embargo, las plataformas en línea deben buscar alternativas para tener acceso a la fuerza laboral de moderación de contenido requerida. Para aquellas plataformas con recursos adecuados, puede ser útil contratar (más) moderadores como personal de tiempo completo.

Emitir voluntariamente informes de transparencia sobre la desinformación de COVID-19. Las actualizaciones periódicas de las plataformas en línea sobre la naturaleza y la prevalencia de la desinformación de COVID-19, y las acciones que las plataformas están tomando para contrarrestarla, permitirían enfoques mejores y más basados en evidencia por parte de las partes interesadas del sector público y privado. Si varias jurisdicciones consideran requerir tales informes, se necesitará la cooperación internacional y la coordinación a través de procesos de múltiples partes interesadas para evitar una serie fragmentada de estándares de presentación de informes que innecesariamente aumentarían los costos y complicarían las comparaciones globales entre plataformas.

Mejore las habilidades de alfabetización mediática, digital y sanitaria de los usuarios. Las personas necesitan las habilidades para navegar y dar sentido a lo que ven en línea de manera segura y competente, y para entender por qué se les muestra. Esto incluye saber cómo verificar la precisión y confiabilidad del contenido al que acceden y cómo distinguir las noticias reales de las opiniones o rumores. Para este fin, la colaboración entre plataformas, organizaciones de medios, gobiernos y educadores es crítica, al igual que los esfuerzos para mejorar la

alfabetización en salud. La reciente asociación entre la Unión Europea, la UNESCO y Twitter para promover la alfabetización mediática e informativa en medio de la crisis de desinformación COVID-19 es una iniciativa loable que debería ser replicada por otras plataformas y partes interesadas relevantes.