

Declaración de Danielle Coffey
Presidenta y directora ejecutiva News Media Alliance
Comité Judicial del Senado de los Estados Unidos
Subcomité de Privacidad, Tecnología y Derecho

10 de enero de 2024

Supervisión de la IA: el futuro del periodismo

Presidente Blumenthal, miembro de alto rango Hawley y miembros del Subcomité, gracias por invitarme a testificar aquí hoy sobre “Supervisión de la IA: el futuro del periodismo”. Mi nombre es Danielle Coffey y soy presidenta y directora ejecutiva de News/Media Alliance (N/MA), y represento a más de 2.200 editores de noticias, revistas y medios digitales, que van desde los editores de noticias y revistas más importantes hasta periódicos hiperlocales, y desde medios exclusivamente digitales hasta periódicos que han impreso noticias desde antes de la Convención Constitucional.

Mis miembros producen contenido periodístico y creativo de calidad que busca informar, educar y conectar con los lectores y enriquecer su vida diaria. Nuestras publicaciones se adhieren a principios y procesos que respaldan la verificación, la exactitud y la fidelidad a los hechos. Cubrimos desastres naturales, zonas de conflicto, juntas escolares, ayuntamientos, ayuntamientos, entretenimiento y artes, y otros asuntos de interés público para las comunidades locales, nacionales e internacionales. Cumplimos con los estándares y códigos de conducta y brindamos a los lectores voz a través de políticas de corrección para garantizar la precisión en los informes.

Lamentablemente, no puedo decir lo mismo de la mayor parte de lo que se difunde y prolifera en el resto de Internet. Sin salvaguardias adecuadas, no podemos confiar en un conjunto común de hechos que promuevan un discurso público saludable. Sin informes de calidad, no podemos tener un electorado informado y una sociedad funcional. Esto es particularmente preocupante ahora que comenzamos un año electoral en el que los lectores dependerán de la precisión de la información para tomar decisiones importantes. Con un mercado ya inundado de hechos falsificados, la Inteligencia Artificial Generativa (GAI) podría crear un riesgo aún mayor para el ecosistema de la información con imprecisiones y alucinaciones si no se frena con contenido de calidad proveniente de publicaciones de noticias.

El riesgo de que el contenido GAI de baja calidad domine Internet se ve amplificado por la drástica caída económica de las publicaciones de noticias en las últimas dos décadas. Los estudios muestran que más de dos periódicos cierran cada semana, y que un tercio de las publicaciones

de los países cerrarán antes de 2025. Se trata de periódicos locales y comunitarios que cubren escuelas de barrio, partidos de fútbol, aperturas de empresas y constituyen el tejido de nuestra país. Aunque los medios de comunicación locales luchan por sus vidas, la gente está más interesada que nunca en las noticias. Si bien ha habido una disminución en los ingresos del 56 % desde 2014, durante ese mismo período de tiempo, el tráfico a los 46 sitios de noticias principales aumentó un 43 %. ¿A qué se debe este aumento del consumo y la producción de periodismo de calidad y, al mismo tiempo, la pronunciada caída del ingreso financiero? Porque los distribuidores dominantes de contenidos noticiosos (también las empresas de Internet más grandes del mundo) están eliminando los sitios web de las publicaciones y vendiendo porciones para generar participación e información personal para dirigirse a los usuarios con publicidad.

En los últimos años, ha habido innumerables estudios, investigaciones y litigios por parte del Departamento de Justicia en las dos últimas administraciones que han encontrado conductas anticompetitivas por parte de los distribuidores monopolísticos de contenido de noticias. Existe un nexo directo entre esta conducta anticompetitiva y la pronunciada caída de los ingresos financieros devueltos a las publicaciones de noticias. Esto ha llevado a una dramática escasez de ingresos necesarios para financiar los costosos gastos asociados con la recopilación de noticias y la contratación de periodistas.

Estas plataformas dominantes han acaparado el mercado de los datos de los usuarios y la publicidad dirigida a través de sus jardines amurallados que han creado para mantener a los usuarios interesados, en parte a través del contenido de noticias que crean mis miembros. Aunque hay menos usuarios que hacen clic, y cuando los usuarios hacen clic, los monopolistas de las grandes tecnologías obtienen tarifas injustas. Este Comité y las dos últimas Administraciones han estado examinando activamente el poder de monopolio en el ecosistema de tecnología publicitaria que da como resultado que hasta el 70% de los pagos publicitarios vayan a las plataformas dominantes. La extorsión y el arbitraje son indiscutibles. Debido a que más del 70 % de nuestra audiencia de referencia proviene de solo dos plataformas, nos vemos obligados a confiar en ellas y aceptar estos términos. Por lo tanto, no sorprende que los editores no obtengan ingresos a pesar de la demanda masiva de nuestro contenido. Agradecemos profundamente a los senadores Klobuchar y Kennedy por su liderazgo en la Ley de Preservación y Competencia del Periodismo, que aborda este desequilibrio del mercado, y agradecemos a este comité por aprobar esta legislación que es tan vital para la industria de las noticias en julio pasado.

Sin embargo, por más terrible que sea la dinámica actual, GAI solo aumentará este desequilibrio del mercado. Los desarrolladores de GAI rastrean sitios web y utilizan muros de pago para entrenar sus modelos. Un análisis que encargó la N/MA, así como las quejas pendientes ante los tribunales, demuestran la importante incorporación de contenido informativo de calidad en los modelos de entrenamiento de IA. Para colmo de males, los resultados de GAI “resultados” de las consultas de los usuarios a menudo contienen resúmenes, extractos e incluso copias textuales completas de artículos escritos y verificados por periodistas humanos. Estos productos compiten en el mismo mercado, con la misma audiencia y tienen el mismo propósito que los artículos

originales que alimentan los algoritmos en primer lugar. GAI es una exacerbación de un problema existente en el que los ingresos no pueden ser generados por quienes crean el trabajo original, sino que de hecho son desviados de ellos.

A largo plazo, proteger a los creadores y titulares de derechos del uso no autorizado de sus obras para la formación de GAI ayudará tanto a las empresas de tecnología como a los editores de noticias a ofrecer mejores productos y servicios a los consumidores. Los modelos y productos de GAI no serán sostenibles si destruyen el contenido de calidad del que se alimentan. Se han tomado obras expresivas protegidas por derechos de autor sin autorización y sin consentimiento, y se han utilizado repetidamente en la formación, el procesamiento y la exhibición de modelos. Debido a que estos usos van mucho más allá de las barreras establecidas por los tribunales, que ahora sustituyen abiertamente el contenido protegido por derechos de autor y usurpan los mercados de licencias, esto no debe considerarse un uso legítimo según la ley de derechos de autor actual.

Para ser claros, los editores de noticias no se oponen a las tecnologías de IA generativa; queremos ayudar a los desarrolladores a desarrollar su potencial de manera responsable. Los miembros de News Media Alliance están en general dispuestos sentarse a la mesa y discutir soluciones de licencias razonables para facilitar el acceso confiable y actualizado a contenido confiable y autorizado. Una solución constructiva beneficiará a todas las partes interesadas y a la sociedad en general y evitará una incertidumbre prolongada. Algunos desarrolladores de GAI son buenos actores y buscan asociaciones y acuerdos de licencia con publicaciones de noticias, y aplaudimos sus esfuerzos. Sin embargo, queda más trabajo por hacer.

Con motores de búsqueda dominantes que combinan conjuntos de datos de IA con resultados de búsqueda en tiempo real, un proceso conocido como “generación aumentada de recuperación” o “conexión a tierra”, las leyes antimonopolio deberían proteger a los usuarios de estos comportamientos anticompetitivos. La práctica de depender de un medio dominante, en este caso la búsqueda, deja a las publicaciones pocas opciones en cuanto a si permitirían la inclusión y la aquiescencia de este acuerdo comercial. Esto es lo que se conoce como elección de Hobson o, en términos antimonopolio, acuerdos vinculantes. Un acuerdo de este tipo proporcionaría poco o ningún retorno a una publicación en la que hay incluso menos probabilidades de que el usuario haga clic.

A menos que se observen inmediatamente algunas correcciones en el mercado, pedimos al Congreso que intervenga e investigue estas prácticas. También ofrecemos las siguientes sugerencias para que el Congreso garantice que la GAI avance de manera sostenible: transparencia, explicabilidad y trazabilidad. El Congreso debe apoyar la legislación que requiera el mantenimiento de registros y la divulgación de usos de capacitación no autorizados de material que esté protegido por derechos de autor, por medidas técnicas de protección o regido por contratos.

términos que prohíben el scraping para revelar el uso y la ponderación del uso específico de contenido de terceros. Deben desarrollarse medidas sustanciales de transparencia en torno al uso de materiales protegidos por derechos de autor en las tecnologías GAI. El público debería poder saber en qué se entrenaron los modelos de IA y, si así lo desea, realizar las evaluaciones necesarias para seleccionar modelos más fiables o de origen más ético. Y los editores tienen derecho a saber quién copió su contenido y para qué lo utiliza. Los incentivos para evitar la divulgación son demasiado fuertes para apostar por una solución autorreguladora. Las obligaciones deben adaptarse a los raspadores, los desarrolladores y aquellos que configuran modelos fundamentales en aplicaciones personalizadas.

Se debe exigir a los desarrolladores de GAI que proporcionen explicaciones que detallen cómo sus modelos producen resultados y proporcionen enlaces a los materiales citados en los resúmenes. Si bien es valioso la armonización internacional y el tratamiento conjunto de otras preocupaciones relacionadas con los datos, cualquier resultado debe lograr el objetivo central de proporcionar suficiente transparencia en la ingesta y el uso de materiales protegidos por derechos de autor para permitir a los titulares de derechos analizar suficientemente dichos modelos.

Infracción de derechos de autor y otros usos nocivos: la copia no autorizada del contenido del editor para entrenar y alimentar sistemas comerciales que produzcan resultados sustitutivos debe reconocerse como una infracción. Los formuladores de políticas deberían presionar a la industria para que reconozca que la copia desenfrenada de contenido de medios expresivos para capacitar a los LLM que luego compiten con ese contenido viola los derechos exclusivos de los editores y usurpa injustamente sus mercados.

News Media Alliance cree que la ley existente establece que esta sistemática y la conducta competitiva es infractora. Y nos sentimos alentados por los signos emergentes de usos autorizados de GAI para otras obras creativas, como música e imágenes. Pero un reconocimiento más amplio de que el contenido de los medios no es gratuito es fundamental para fomentar negociaciones significativas entre los desarrolladores y editores de GAI.

El Congreso también debería considerar una legislación que aclare que los desarrolladores e implementadores de GAI son responsables del diseño de su tecnología y no califican para los puertos seguros establecidos hace décadas para alentar a una industria tecnológica incipiente a alojar contenido sin tener que monitorear y eliminar activamente el material dañino creado íntegramente por terceras personas. Los sistemas GAI no son anfitriones pasivos ni conductos para el contenido creado o utilizado por otros; estos sistemas utilizan y generan contenido activamente por sí mismos. Los sistemas GAI y quienes los desarrollan deben ser responsables y rendir cuentas, como cualquier otro negocio.

Licencias y competencia: Los formuladores de políticas deben fomentar soluciones de licencias basadas en el mercado y respetar las leyes y políticas establecidas que desalientan la regulación

gubernamental de los mercados de licencias como primer recurso. Para algunos mercados, esto puede incluir licencias colectivas voluntarias, como ya lo permite la ley. El gobierno debe impedir que los desarrolladores condicionen o modifiquen la prestación de otros servicios, como publicidad o clasificación de búsqueda, a que (a) un propietario de contenido u operador de sitio ponga a disposición contenido para capacitación o (b) un propietario de contenido u operador de sitio permita el uso de dicho contenido o sitio en búsqueda u otros servicios o imponer términos y condiciones razonables.

Diseño responsable y rendición de cuentas: el Congreso debe garantizar que el desarrollo de la IA generativa esté diseñado para ser responsable, en lugar de mitigar a posteriori. Sería útil dejar explícito que las inmunidades previstas en la Sección 230 de la Ley de Decencia en las Comunicaciones no se aplican al contenido generado por IA. La experiencia de los editores en otros contextos es que los estándares más bajos de responsabilidad reducen los incentivos para que las plataformas negocien el uso de contenido de calidad, aumentan el daño al público y colocan a los editores de medios en una desventaja competitiva. Las primeras tendencias de la GAI revelan que los daños llamados “sobreadaptación” o “no intencionales” son demasiado comunes. Se debe incentivar a los desarrolladores para que incorporen principios de seguridad desde el diseño y mantengan programas para evitar resultados peligrosos del contenido generado por IA, como contenido ilegal y otros daños graves en línea.

Aplicación de la ley y lucha contra la piratería: el Congreso debe garantizar herramientas adecuadas para evitar el scraping y el “blanqueo” no deseado de contenido protegido por derechos de autor. Los web scrapers deben respetar y seguir los términos de uso y cumplir con las señales automáticas que indican que el contenido en línea debe limitarse a usos específicos. Por ejemplo, las herramientas podrían señalar el deseo de los usuarios de bloquear el rastreo para capacitación y al mismo tiempo permitir usos beneficiosos, o de que la capacitación se limite a usos o usuarios particulares. Los sitios piratas conocidos deberían estar prohibidos para fines de entrenamiento de IA, incluso si esos propietarios de sitios permitieran la extracción de datos y los esfuerzos de cumplimiento del Departamento de EE. UU.

El Ministerio de Justicia, la Oficina del Coordinador de Observancia de la Propiedad Intelectual (IPEC) y otras partes del gobierno deben recibir poderes y financiación.

El Cuarto Poder ha desempeñado un valioso papel en este país durante siglos, llamando a los gobiernos y a los líderes cívicos a actuar responsablemente en sus posiciones de poder. Las noticias locales, especialmente, han descubierto e informado sobre eventos en todo el país que mantienen a los lectores informados, educados e involucrados en sus comunidades. Simplemente no podemos permitir que se ignore la prensa libre a expensas de tecnologías nuevas y apasionantes. Ambos pueden existir en armonía y ambos pueden prosperar. Debemos garantizar eso por el futuro de nuestro país y el futuro de nuestra sociedad.

Gracias por la oportunidad de participar en esta discusión. Esperamos trabajar con el Congreso,

la Administración, los Estados y nuestros homólogos de todo el mundo a medida que avanza esta importante discusión.