Linterna Verde

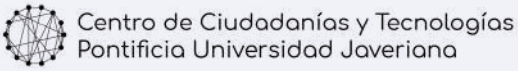
linternaverde.org

La causa de los datos

En busca de una agenda común para América Latina

Marzo, 2025

En alianza con:



Este contenido es posible gracias al apoyo de Luminate y el Gobierno de Canadá. Las opiniones, ideas y contenidos expresados son exclusiva responsabilidad de sus autores y no reflejan necesariamente las posiciones oficiales de las organizaciones mencionadas.

Luminate

Canadá

Contexto y propósito

El 28 y 29 de noviembre de 2024, Linterna Verde organizó en Bogotá un taller con un grupo de integrantes de la sociedad civil y la academia para abordar la pregunta por el acceso a los datos de plataformas de redes sociales para la investigación digital de interés público (ver acá detalles sobre el evento y perfil de los invitados). Este espacio buscó darle continuidad a un par de conversaciones previas que Linterna Verde hizo durante 2023.

Durante el taller se llevó a cabo un panel con tres de los participantes (video disponible aquí). Las reflexiones y conclusiones de cada uno quedaron consignadas en tres textos de la Red de Expertos en Democracia y Tecnología de La Silla Vacía: 'Descolonizando datos y narrativas', 'Ganar lo cualitativo sin perder lo masivo', y 'La causa de combatir el apagón de datos'.

Estos ejercicios permitieron hacer un abordaje del problema desde diversos puntos de vista, pero también sirvieron para constatar la ausencia de un esfuerzo en la región para impulsar una agenda común.

El propósito de este documento es demarcar una hoja de ruta para una iniciativa conjunta. Para el efecto, (i) hacemos una descripción del problema; (ii) enfatizamos la necesidad de una mirada diferencial para América Latina, y (iii) proponemos elementos clave para una estrategia de corto plazo.

Este texto recoge participaciones del taller –a manera de relatoría–, argumentos de otros documentos e ideas de los artículos mencionados. En algunos apartes incluimos citas textuales; en otros, parafraseamos o hacemos planteamientos de manera general. En cualquier caso, aunque este insumo no pretende ser exhaustivo ni representa un consenso formal, refleja una perspectiva plural.

1. El desafío que enfrentamos

La investigación digital desde la sociedad civil latinoamericana constituye un reto cada vez más difícil de sortear. En particular, llevar a cabo monitoreos sobre el ecosistema informativo en redes sociales, páginas web y grupos abiertos de plataformas, se enfrenta a costos incrementales de acceso, falta de infraestructura y el desafío de la retención del talento humano. El lugar de los datos y la dificultad práctica de tener o acceder a servicios de escucha social digital, es un dilema existencial generalizado.

La escucha social digital de interés público enfrenta obstáculos económicos, estructurales y políticos. La sociedad civil depende de la voluntad de las plataformas, al vaivén de las condiciones cambiantes de acceso a las API (como se conocen estas interfaces de conexión entre aplicaciones) y los caprichos de los líderes de Silicon Valley. Por otra parte, las soluciones comerciales –herramientas como Brandwatch o Meltwater– implican inversiones difíciles de costear.

El cierre en agosto de 2024 de Crowdtangle, la herramienta de monitoreo de las redes sociales de Meta, significó un nuevo golpe para la investigación digital de las ONG, la academia y los periodistas. Si bien la empresa abrió una nueva forma de acceso a datos a través de la Librería de Contenidos de Meta, el panorama general es inestable, restringido y precario.

A este proceso de 'apagón' de datos se suman las decisiones de Meta, anunciadas en enero pasado, de terminar el programa de verificación y cambiar algunas políticas sobre moderación de contenidos. Hasta la fecha no se conocen decisiones que impacten de manera directa el monitoreo o el acceso a las API. No obstante, si antes el ambiente ya era adverso para la investigación social, ahora debe navegar también un ambiente político hostil.

En palabras de la Coalición para la Investigación Independiente de Tecnología, *"la sociedad necesita investigación confiable e independiente para aliviar los daños de las tecnologías digitales y avanzar hacia el bien común... Debido a que la investigación independiente sobre tecnologías digitales es una fuerza poderosa para el bien común, también enfrenta una fuerte oposición"*.

Sin datos, el acceso al conocimiento en el entorno digital se vuelve imposible. Esta barrera complica los esfuerzos para entender el ecosistema, los riesgos para grupos subrepresentados y las posibles respuestas regulatorias y de política pública. Con el despliegue de la Inteligencia Artificial, la situación se torna aún más precaria.

2. América Latina: la necesidad de una mirada diferencial

Algunos de los problemas más apremiantes en América Latina se relacionan con las desigualdades significativas en el acceso a la tecnología, la escasez de herramientas en español o portugués y barreras de localización que limitan la capacidad de análisis de datos. Además, hacen falta iniciativas de alfabetización digital, ya que muchos actores en el sector desconocen las oportunidades existentes.

La relevancia de Twitter (ahora X) en el debate público de la región instaló una mirada aislada y consolidó metodologías que se fueron quedando cortas para dar cuenta de las dinámicas en línea. Hoy en día, con una multiplicidad de espacios y audiencias fragmentadas, la limitación del sector se hizo más patente. Las barreras para innovar en los monitoreos incentivan observaciones repetidas que son insuficientes para atender las demandas cambiantes de investigación. Lo anterior sin desconocer buenas prácticas a las que haremos referencia más adelante.

El punto de entrada principal a los datos de redes sociales –el acceso a las API– hace más complejo un proceso de adopción. En Europa los investigadores tienen acceso a la API de TikTok, pero en América Latina no está disponible a pesar de las promesas de la compañía. Mientras Meta ofrece la API restringida, X tiene una versión gratuita básica y paquetes pagos. Linkedin prevé accesos a través de solicitudes directas. La excepción es Youtube, que con una API robusta se ha vuelto “un oasis en medio del desierto de datos”, como lo plantea Carlos Cortés, director de Linterna Verde.

Los avances en la Unión Europea (UE) a través de la Ley de Servicios Digitales (DSA, por su sigla en inglés) no han permeado las políticas de las compañías en América Latina en materia de acceso a datos. Según el artículo 40 de la DSA, mediante un proceso de acreditación, los investigadores pueden acceder a datos de las grandes plataformas. No obstante, el propósito investigativo debe estar alineado con los objetivos de la norma, que incluyen la prevención del riesgo. Para Agustina del Campo, directora del CELE en Argentina, esta provisión termina estrechando el acceso para investigadores. Por el momento, una iniciativa similar se discute en Brasil.

La influencia global de la UE –conocida como el 'efecto Bruselas'– también se enfrenta a regulaciones dispares sobre protección de datos en América Latina. Representantes de TikTok, por ejemplo, lo han planteado como uno de los escollos para abrir la API en la región.

Más allá de eso, promover regulaciones al estilo europeo en nuestra región, bajo el propósito loable de tener estándares y normas de ese nivel, no es necesariamente la mejor alternativa. Como explica Gastón Wright, director de Civic Compass, *"para muchos países en las regiones de la mayoría global, donde la capacidad institucional y los recursos ya están sobrecargados, replicar las regulaciones de la UE puede ser impráctico y contraproducente"*.

Tener una mirada localizada no implica, de ninguna manera, desconectarse del escenario mundial. Al contrario: las iniciativas de fact-checking y periodismo, los análisis de políticas públicas, el monitoreo de las plataformas, el activismo y la estrategia de comunicaciones para la sociedad civil en América Latina requieren de una conversación que parta de una mirada regional y se conecte con la agenda global. Con la excepción de algunos esfuerzos aislados, carecemos de iniciativas coordinadas para tener interlocución con coaliciones de otras latitudes y articular soluciones propias.

Nuestra mirada frente al acceso a datos de las plataformas debe tender puentes con el trabajo sobre moderación de contenidos y transparencia, la protección de datos personales y la Inteligencia Artificial. Por un lado, organizaciones como Linterna Verde, el InternetLab y Signa_Lab Iteso han usado el monitoreo en redes sociales para identificar, entre otras, contenido problemático que puede considerarse incitación al odio, a la violencia o acoso.

Por otro lado, la conexión con el debate de Inteligencia Artificial permite darle vigencia a la causa de los datos y, sobre todo, poner sobre la mesa el impacto para mujeres, grupos racializados y población LGBTIQ+. En palabras de Fernanda Martins, *"las herramientas de Inteligencia Artificial, especialmente el reconocimiento facial, nos desafían a repensar cómo los datos contemporáneos están siendo leídos, interpretados y traducidos en tecnologías que continúan reproduciendo desigualdades históricas"*.

3. La causa en la práctica: una hoja de ruta

Existe una necesidad urgente de articular una iniciativa innovadora y colectiva que permita alentar el trabajo con datos de plataformas en pro del interés público y, a la vez, reenfocar la causa para lograr un acceso más robusto y transparente. Para avanzar en este camino resulta fundamental habilitar espacios de confianza que promuevan intercambios honestos, constructivos y accionables.

A continuación proponemos una serie de oportunidades y acciones en cinco frentes: (i) demanda y uso de datos; (ii) colaboración y alianzas; (iii) innovación y adaptación; (iv) ética y regulación, y (v) enmarcados y narrativas.

a. Demanda y uso datos

- Hacer un diagnóstico de las demandas de datos que tienen las organizaciones, ya sea de plataformas de redes sociales o de otros espacios digitales. Este análisis debe incluir elementos como la escala y tipo de datos que requieren, el trabajo profesional para implementarlo, y los productos desarrollados a partir de estos insumos.

- Identificar otras variables relevantes como el tiempo requerido para proyectos prototipo, costo de opciones comerciales de escucha social y retos generales de las metodologías que emplean.

- Mantener un debate crítico interno frente al tipo de preguntas que pueden responder los datos de redes sociales y la pertinencia de las metodologías desplegadas.

- Con base en estas observaciones, hacer un análisis interno de la viabilidad y necesidad de tener unidades de datos o líneas de trabajo propias en este frente o, en su lugar, contemplar posibles esquemas de apoyo externo.

b. Colaboración y alianzas

- Identificar instancias de colaboración disponibles u oportunidades para forjar nuevas alianzas que permitan canalizar un esfuerzo en torno a esta causa –como hubs de investigadores–. Dicha acción conjunta debe estar acotada en su objetivo y buscar avances de corto plazo.

- Tender puentes con la academia para el desarrollo de buenas prácticas y marcos éticos, al igual que para facilitar el acceso a investigadores afiliados a universidades e institutos de investigación.

- Priorizar el uso compartido de herramientas y metodologías, la sinergia y el fortalecimiento de organizaciones nodo, la búsqueda de apoyos del sector privado y la posible adopción de alternativas no comerciales.

- Sistematizar y compartir las buenas prácticas de organizaciones como InternetLab, en Brasil; Civic House, en Argentina; Signa_Lab ITESO, en México; ColombiaCheck y La Silla Vacía, en Colombia, y otras de alcance regional como el CLIP, DFRLab y Linterna Verde.

- Establecer dentro de las prioridades un proceso de vinculación proactivo con las iniciativas del norte global en la materia, tanto para el acceso a datos en investigaciones clave de interés público (integridad electoral, por ejemplo), como para avanzar en la causa de uso de transparencia y uso de las API.

- Entablar un diálogo con las agencias de cooperación para diseñar indicadores que fomenten el trabajo conjunto y reflejen de mejor manera el derrotero planteado. Estos actores también pueden ser relevantes para apalancar la interlocución con las plataformas.

- Construir estrategias de transparencia y acceso a datos que no se limiten solo a la API, sino que incluyan otras metodologías, tanto cuantitativas como cualitativas, que puedan ser aplicadas por investigadores de la región, como lo explora InternetLab en este artículo.

c. Innovación, adaptación y alfabetización digital

- Pensar y diseñar estrategias de acceso y compilación de datos más allá y a pesar de las API.

- Allanar el camino para la adopción de otras herramientas de análisis, metodologías y trabajo directo con ciudadanías y comunidades. La causa de los datos implica también crear puentes entre lo técnico, lo político y lo social.

- Compartir metodologías propias y analizar alternativas distintas de escucha social con énfasis en la investigación social, como Junkipedia.

- Incorporar un plan de trabajo para el uso de la Inteligencia Artificial y otras tecnologías emergentes. En particular, el sector de verificación o fact-checking cuenta con aprendizajes y buenas prácticas que pueden ayudar a organizaciones pares a enfrentar los constantes cambios en el entorno digital.

- Desarrollar herramientas de procesamiento y post-procesamiento de datos, como los cuadernos de Signa_Lab ITESO, que permiten utilizar y generar diccionarios propios y hacer consultas personalizadas con IA acerca de distintos tipos de discurso.

- Contemplar iniciativas de formación en el uso seguro y ético de datos, el fomento de herramientas descentralizadas y el impacto en comunidades vulnerables.

d. Ética, regulación y seguridad digital

- Incluir mecanismos de protección para investigadores ante ataques digitales, acoso o censura, además de buenas prácticas en privacidad, anonimización y almacenamiento seguro de datos.

- Diseñar estándares de autorregulación y espacios colectivos para abordar dilemas éticos. Contemplar elementos como conflictos de interés y sesgos, riesgos de perfilamiento, vulneración de la privacidad, imputaciones individuales, descontextualización de los datos e impacto en grupos vulnerables.

- Incorporar el conocimiento de otros grupos de la sociedad civil y de la academia en materia de regulación, transparencia y protección de datos. Hacer seguimiento a la experiencia de la DSA e iniciativas similares, como la brasileña, con una mirada crítica y contextual.

- Prever mecanismos de auditoría para evitar sesgos y vulneraciones de derechos, además de analizar marcos regulatorios emergentes en América Latina.

e. Enmarcados y narrativas

- La causa de los datos se moviliza a través de historias. Es importante trabajar en la intersección entre el acceso a datos de redes sociales y la reivindicación y protección de derechos fundamentales y causas de interés público, como la justicia de género y la justicia climática, los derechos de poblaciones vulnerables y las comunidades indígenas.

- Dichos propósitos deben propender por objetivos claros y definidos, articulados a través de activaciones y campañas. En ese esfuerzo, las organizaciones deben buscar roles definidos y voces influyentes en la discusión pública.

Este documento busca ser un punto de partida para fortalecer una agenda común en América Latina que permita promover el acceso equitativo y transparente a los datos de plataformas digitales en beneficio del interés público.

Agradecemos a todas las personas e instituciones que han contribuido a esta discusión y a quienes continúan impulsando soluciones en este campo. Creemos que solo a través del diálogo y la cooperación podremos hacer causa conjunta.

Si desea conocer más sobre esta iniciativa, compartir ideas o explorar posibles colaboraciones, no dude en ponerse en contacto con nosotros.

contacto@linternaverde.co
linternaverde.org
lacausadelosdatos.com

Linterna Verde