



**WORLD WIDE WEB
FOUNDATION**

LA MANO INVISIBLE:

el News Feed de Facebook
y nuestra dieta informativa

Abril 2018

www.webfoundation.org

CONTENTS

Prólogo	3
Resumen ejecutivo	4
01 La 'dieta informativa' de Facebook: el experimento Argentino	6
02 Metodología: Qué hicimos	8
03 Análisis: Hallazgos	12
04 Conclusiones y Recomendaciones	16
05 Referencias	18



La Web Foundation se fundó en el año 1009 por Sir Tim Berners-Lee, el inventor de la World Wide Web. Nuestra misión es establecer a la Web abierta como un bien público y un derecho básico.

Reconocimientos

Escrito por Renata Ávila, Juan Ortiz Freuler y Craig Fagan. Claudio Agosti y el equipo de Facebook Tracking Exposed proporcionaron la herramienta y el soporte técnico. Con agradecimiento especial a Martín Becerra, Marius Dragomir, Scott Hale, Guillermo Mastrini, Samuel Maynard, Paula Szewach, Ben Yuan, Martin Becerra y R'lyeh Hacklab por sus aportaciones. Estudio financiado por Open Society Foundations.

Derechos de autor 2018, World Wide Web Foundation, CC BY 4.0



PRÓLOGO

En marzo de 2017, nuestro fundador e inventor de la World Wide Web, Sir Tim Berners-Lee, expresó su preocupación ante una ola creciente de pérdida de control sobre nuestros datos personales y el papel que los algoritmos y la inteligencia artificial están desempeñando en la sociedad. Durante el último año, la Web Foundation se embarcó en estudios para proporcionar marcos de referencia y ejemplos para pensar cómo el uso de los datos personales, algoritmos e inteligencia artificial crean riesgos y oportunidades para los países de ingreso bajo y mediano.¹

Una de las preocupaciones que ha surgido de este estudio es la manera en la que los algoritmos pueden crear un alto grado de opacidad en torno a información y actividades que son cruciales para nuestras sociedades, como el acceso a servicios y el debate público. Dada la misión de la Web Foundation de cerrar la brecha digital, nos preocupa la dinámica de poder entre quienes construyen algoritmos y quienes se ven afectados por ellos. Por eso intentamos poner el foco sobre el impacto de los algoritmos en nuestras sociedades así como en las experiencias en línea de las personas.

"Debemos trabajar para revertir esta tendencia y garantizar que se devuelva el control a las personas que utilizan estas plataformas, y que sus derechos en línea se protejan y respeten."

Este informe se centra en uno de los algoritmos con el que los usuarios de internet interactúan cotidianamente: cómo las redes sociales filtran y distribuyen la información a sus usuarios. Para comprender mejor este proceso, realizamos un experimento controlado utilizando seis perfiles de Facebook idénticos, que seguían las mismas fuentes de noticias en esta red social. Los resultados del experimento muestran cómo se ha delegado en los algoritmos una función crucial de filtrado que define las dietas informativas de los usuarios de Facebook y lo que se podría llamar mejor la "plaza pública" digital. Si bien el estudio se ha enfocado en Facebook, hay preocupaciones que similares dinámicas rijan otras plataformas como YouTube y Twitter—que configuran y estructuran la dieta informativa de los usuarios sin su conocimiento o control, apoyándose en algoritmos similares para tareas de filtrado y priorización.

Debemos trabajar para revertir esta tendencia y garantizar que se devuelva el control a las personas que utilizan estas plataformas, y que sus derechos en línea se protejan y respeten. Consideramos que este informe como un primer paso en esta dirección: echar luz sobre esta dinámica, para que el público en general comprenda mejor cómo funcionan los algoritmos de la plataforma y la mecánica detrás del proceso. Al hacerlo, esperamos que este documento ayude a iniciar una conversación más amplia sobre el papel de los usuarios de internet, los gobiernos y las plataformas para así todos definir los valores a aplicar al diseño, aprovechamiento y responsabilidad de los algoritmos junto a las personas y comunidades afectadas.

¹ WF ha publicado tres documentos técnicos sobre estos temas. Para más información, visite: <https://webfoundation.org/research/white-paper-series-opportunities-and-risks-in-emerging-technologies/>.

RESUMEN EJECUTIVO

Facebook, que con 2 mil millones de usuarios mensuales activos se alza como la red social más grande del mundo, ha desarrollado una amplia gama de tecnologías que, de forma inherente, configuran su relación con los usuarios. Esto incluye definir la información a la que los usuarios de Facebook estarán o no expuestos a través de su plataforma. Este estudio se apoya en esfuerzos de investigación anteriores para mostrar cómo cada usuario de Facebook tiene una "dieta informativa" hecha a medida.² Los experimentos incluidos en este informe se han realizado para iniciar una conversación acerca de qué significa esta dieta informativa, y el papel de los algoritmos de Facebook en su definición. A pesar de los cambios anunciados por Facebook recientemente,³ la función News Feed continúa permitiendo únicamente un control limitado por parte del usuario sobre las funciones de personalización limitadas que los usuarios pueden controlar.⁴

Los resultados de los experimentos incluidos en este informe muestran la existencia de variaciones sustanciales en la información que los usuarios reciben a través de la plataforma, incluso cuando los usuarios tienen perfiles con exactamente las mismas características. El estudio muestra:

- **Gran brecha entre el número de historias que fueron publicadas y las que fueron vistas por nuestros perfiles:** Los perfiles seguían las mismas páginas, y de todas las noticias publicadas por éstas, únicamente se les mostró un promedio de una de cada seis publicaciones.⁵ Una brecha tan amplia entre lo que se publica y lo que efectivamente ve un usuario permite dar una idea del margen de discrecionalidad con el que cuenta para el ejercicio de la curaduría de contenido - es decir, la selección, distribución y difusión de contenidos por parte del algoritmo.⁶
- **Falta total de exposición a ciertas noticias:** La decisión de depender en algoritmos para la curaduría del contenido que se distribuye a los usuarios significa que publicaciones sociales y políticas cruciales que se encuentren fuera de los parámetros de distribución pueden no mostrarse a algunos usuarios. Por ejemplo, durante el período observado, ninguna de las historias sobre femicidios publicadas por las páginas que nuestros perfiles seguían aparecieron en el News Feed de nuestros usuarios.

2 Kincaid, Jason. 2010. "EdgeRank: The Secret Sauce That Makes Facebook's News Feed Tick." <http://techcrunch.com/2010/04/22/facebook-edgerank/>; Taylor, Dan. 2011. "Todo lo que necesita saber sobre el EdgeRank de Facebook". <http://thenextweb.com/socialmedia/2011/05/09/everything-you-need-to-know-about-facebooks-edgerank/>; Napoli, P. (2014). Automated Media: An Institutional Theory Perspective on Algorithmic Media Production and Consumption. *Communication Theory*, 24(3), 340-360.; Devito, M. (2017). From Editors to Algorithms: A values-based approach to understanding story selection in the Facebook news feed. *Digital Journalism*, 5(6), 753-773.; Pariser, E. (2011). El filtro burbuja: Cómo la web decide lo que leemos y lo que pensamos. Penguin.

3 Facebook está lanzando un sistema de publicidad más transparente, que ya se está experimentando en Canadá. Ver: <https://www.propublica.org/article/facebook-experiment-ad-transparency-toronto-canada>.

4 La opción de inhabilitar el News Feed autenticado ("Historias principales") se restablece cada vez que el usuario sale de la aplicación. La opción de priorizar ciertas páginas básicamente permite a los usuarios ubicar el contenido de una persona o página específica "primero" en el News Feed. No está claro qué sucede cuando varias páginas o perfiles están marcados para aparecer primero en el News Feed, o cómo funcionan el News Feed en modo predeterminado, o para las publicaciones publicadas por agentes que no estén marcados como una prioridad.

5 Esta cifra equivale al 18 % de todas las publicaciones publicadas durante el período, pero varía entre las Páginas. Ver <https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecuratorofcontent-Full-English/SeenNotseenBars>.

6 Este porcentaje es sensible a la relación entre el tipo y el número de páginas seguidas (por ejemplo, cuánto contenido se publica) por una parte, y la cantidad de tiempo que se pasa navegando por la línea de tiempo de Facebook por la otra (demanda). Las propias estimaciones de Facebook sobre cómo los usuarios experimentan esta brecha son similares. <https://www.facebook.com/business/news/Organic-Reach-on-Facebook> Nuestro estudio proporciona una ventana para que los usuarios vean cómo sucede esto y se muestran ejemplos de qué tipo de publicaciones se pueden ver y excluir.

- **Diferentes niveles de exposición a diferentes artículos:** Incluso usuarios con el mismo perfil y que siguen los mismos sitios de noticias pueden no estar expuestos al mismo conjunto de historias. Es como si dos personas compraran el mismo periódico impreso, pero descubrieran que se han suprimido diferentes conjuntos de historias de sus ejemplares. El algoritmo coloca a cada usuario en una versión separada e individualizada de lo que debería ser una plaza abierta y pública de información.

Estos hallazgos se basan en pruebas realizadas en Argentina entre los meses de octubre y diciembre de 2017 utilizando una serie de perfiles de Facebook y herramientas automatizadas para estandarizar, controlar y supervisar su comportamiento en línea.⁷ El estudio presenta un enfoque metodológico que podría ayudar a expandir el estudio sobre la dinámica de los algoritmos utilizados por las redes sociales. Los estudios académicos y artículos de prensa que analizan el papel de Facebook en la distribución de noticias generalmente se han centrado en los contextos de la Unión Europea y los EE.UU. Los resultados a nivel de país para Argentina tienen el potencial de ser extrapolados a los usuarios de Facebook en otros países dada la manera estandarizada en que la plataforma distribuye información entre usuarios (ver "metodología").

Recomendaciones

Estos hallazgos sugieren que se produce un grado de curaduría de contenido en la plataforma que parece no estar relacionada con los perfiles e intereses de los usuarios. The Web Foundation cree que los usuarios deben tener un mayor control sobre su News Feed⁸ y sugiere las siguientes medidas para ayudar a Facebook a abordar las brechas encontradas por el estudio:

- **Explorar el uso de paneles de transparencia en tiempo real,** que permitan a los usuarios contrastar las historias a las que fueron expuestos con el universo total de historias publicadas durante un período de tiempo determinado, a fin de aumentar la comprensión pública sobre el funcionamiento del algoritmo.

- **Que la publicación de contenido en orden cronológico** sea la forma predeterminada para su distribución a través de Facebook.⁹
- **Permitir a los usuarios una variedad de opciones para administrar su propia configuración del News Feed,** y así desempeñar un papel más activo en la construcción de sus dietas informativas.¹⁰
- **Responder a las necesidades de los usuarios de Facebook,** generando más espacios para que comenten sobre su experiencia con los algoritmos de la plataforma, y permitiéndoles un rol más activo en las decisiones que se implementarán en la plataforma.
- **Facilitar estudios independientes sobre las externalidades** y los efectos sistémicos de la curaduría algorítmica del contenido.

El modelo actual de control centralizado sobre las "dietas informativas" requiere debate y revisión. Facebook se encuentra estudiando la reestructuración de su News Feed, situación que ofrece una ventana de oportunidad para tener este debate.¹¹ Sin embargo, el anuncio no ha ofrecido detalles sobre cómo los usuarios podrán ejercer un mayor control sobre el contenido de su News Feed, o renunciar a los cambios que se vienen.

Esperamos que este informe aliente a Facebook a redefinir su papel y capacitar a sus usuarios para que determinen sus propias dietas informativas.

7 Se realizaron dos experimentos: el primero con nueve perfiles de usuario y el otro con seis. La segunda prueba se ajustó para los desafíos que se encontraron en el primer experimento con el fin de minimizar las variables que se estaban evaluando.

8 Las encuestas globales muestran que casi 1 de cada 2 personas no se sienten cómodas facilitando información a empresas en línea a cambio de servicios y productos personalizados. Además, casi el 60 % de las personas les importa que las empresas utilicen la información que se les suministra automáticamente. Ver: <https://www.ipsosglobaltrends.com/wp-content/uploads/2017/04/Slide02-5.jpg> and <https://www.ipsosglobaltrends.com/wp-content/uploads/2017/04/Slide03-5.jpg>.

9 La configuración actual permite a los usuarios optar por inhabilitar la autenticación algorítmica. Sin embargo, en el contexto de las pruebas informales, hemos observado que el reinicio de la aplicación restablece esta preferencia al valor predeterminado de autenticación algorítmica. Ver <https://www.facebook.com/help/community/question/?id=10152476808951132>.

10 Ver por ejemplo gobo.social.

11 "No pensamos en términos de "neutral", continuó Stretch, señalando que Facebook intenta darles a los usuarios una fuente de contenido personalizada. "Pero consideramos que estamos (...) abiertos a todas las ideas sin tener en cuenta el punto de vista o la ideología". Transcripción de Stretch, VP de Facebook, en la audiencia del Senado sobre las interferencias en las elecciones rusas de 2016 <https://www.recode.net/2017/10/31/16579072/live-update-facebook-google-twitter-testimony-russia-congress-hearing>

01

LA 'DIETA INFORMATIVA' DE FACEBOOK: EL EXPERIMENTO ARGENTINO

Con más de dos mil millones de usuarios mensuales activos en todo el mundo, Facebook se ha convertido en la red social líder a nivel mundial. Esta posición significa que Facebook posee una enorme cantidad de datos personales de sus usuarios, en los que se apoya para refinar los algoritmos para la que personalizan la distribución de contenido a través de la plataforma. Estos algoritmos se utilizan para determinar qué información se muestra a los usuarios y cuándo. Desde publicidades sobre lo último en zapatos de invierno hasta alertarnos sobre una reciente foto de una amiga con sus perros.¹²

Facebook proporciona a los usuarios una experiencia personalizada, al filtrar las montañas de información que circulan dentro de su plataforma. Este modelo de filtrado tiene por objetivo aumentar la cantidad de horas que pasan los usuarios dentro de la plataforma.¹³

Dicha personalización mediante el filtrado de contenidos tiene un impacto particularmente relevante cuando se trata de la distribución de noticias. Es como comprar un periódico o una revista en la calle y descubrir que el vendedor ha recortado los artículos que él o ella ha decidido que no serían de tu interés. Estas dietas informativas personalizadas, fuera del ámbito de control de los usuarios, generan el riesgo de exacerbar las divisiones entre las comunidades que de manera creciente carecen de una base común de noticias y, por lo tanto, de una realidad compartida.

Los algoritmos de personalización están alterando radicalmente la experiencia de consumir noticias en línea. Como usuarios, sólo estamos al tanto del contenido que presenta el algoritmo para nuestro News Feed: los artículos que quedaron dentro de una revista recortada. Los usuarios desconocen en gran medida qué noticias se omitieron de su News Feed y el porqué.¹⁴ Mientras Facebook permite a los usuarios desactivar el algoritmo que gobierna el News Feed, o priorizar el contenido de determinadas páginas, estas opciones son limitadas y difíciles de usar, ya que no proporcionan suficientes detalles para que los usuarios tomen decisiones informadas sobre sus dietas informativas.¹⁵

12 Facebook estima que "De las más de 1.500 historias que una persona puede ver cada vez que inicia sesión en Facebook, el News Feed muestra aproximadamente 300. Para elegir qué historias mostrar, News Feed clasifica cada historia posible (de más a menos importante) analizando miles de factores relativos a cada persona". Ver: <https://www.facebook.com/business/news/Organic-Reach-on-Facebook>.

13 Para estudios relacionados sobre los algoritmos, ver: <http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/>

14 Powers, E. (2017). My News Feed is Filtered?: Awareness of news personalization among college students. Digital Journalism, 5(10), 1315-1335

15 La opción de inhabilitar el News Feed que filtra la información ("Historias principales") se restablece cada vez que el usuario sale de la aplicación. La opción de priorizar ciertas páginas básicamente permite a los usuarios ubicar el contenido de una persona o página específica "primero" en el News Feed. No hay indicios sobre cómo actuará el algoritmo cuando varias páginas o perfiles están marcados para aparecer primero en el News Feed, o cómo funcionan el News Feed en modo predeterminado, o para las publicaciones publicadas por agentes que no estén marcados como una prioridad.

Se estima que, en los Estados Unidos, el 45 % de las personas tiene a Facebook como su fuente principal de noticias en línea. En Argentina, el foco de este estudio, 65% de la población recurre a Facebook como fuente principal de noticias. Por lo tanto, las decisiones sobre la información a la que se expondrá a los usuarios puede tener efectos profundos en nuestra política, economía y cultura. Estudios recientes muestran que el News Feed de Facebook tiene la capacidad de impactar en la cantidad de personas que deciden ir a votar en una elección, e incluso puede tener un pequeño–pero estadísticamente significativo– efecto en el estado de ánimo de los usuarios. Los efectos de Facebook van cada vez más allá de lo que a priori esperaríamos de

una plataforma social, impactando dinámicas de poder, la política y el bienestar de los ciudadanos.

Este informe espera proporcionar a los investigadores un método experimental para obtener más información sobre cómo funciona el algoritmo del News Feed. También tiene como objetivo brindarles a los usuarios una ventana para observar los proceso que determinan sus dietas informativas. Con ese objetivo, hemos preparado un conjunto de paneles interactivos que acompañan el informe para que los lectores puedan explorar los hallazgos, e indagar sobre qué publicaciones se le ocultaron y cuáles se mostraron en el News Feed de nuestros perfiles.

CONTEXTO: ARGENTINA

- **Mercado:** La tasa de conectividad es de casi el 70 %, incluso en las zonas rurales. Tanto el gobierno anterior como el actual expresaron un compromiso público con la reducción de la brecha digital restante, incluso la existente entre mujeres y hombres. Los costos de Internet son relativamente asequibles: Argentina cumple con el objetivo de asequibilidad "1 por 2" de la Alianza para una Internet Asequible, con 1 GB de datos móviles por un costo equivalente al 1.58 por ciento del ingreso mensual promedio.
- **Protección institucional para el flujo de información:** La libertad de expresión es promedio. Argentina ocupa el puesto 50 de 180 países en el Índice Mundial de Libertad de Prensa. Sin embargo, la libertad de Internet se percibe como ambivalente: hay una preocupación creciente con respecto a la consolidación de los medios de comunicación a medida que las empresas de contenidos, empresas de telecomunicaciones y proveedores de cable se intentan fusionar.¹⁶
- **Crecimiento de las redes sociales:** Las plataformas de redes sociales están asumiendo rápidamente un papel mediador entre los usuarios y el contenido web, incluidos los sitios de noticias. Facebook es el sitio web

más visitado en Argentina,¹⁷ y ha logrado una penetración del 75 % entre los usuarios de Internet de Argentina, de los cuales el 61 % accede a la plataforma al menos una vez al día.

- **Papel de las redes sociales en el activismo:** Las redes sociales, principalmente Facebook y WhatsApp, han sido plataformas clave para el debate público y para la coordinación de acciones colectivas, llevando a la gente de las plataformas a las calles. Este ha sido el caso de las protestas feministas bajo el lema #NiUnaMenos que reclama por políticas efectivas contra los feminicidios y la violencia contra las mujeres–así como las manifestaciones públicas masivas ante la desaparición de un joven activista.
- **Papel de las redes sociales en la política electoral:** El partido político que actualmente está en el poder se apoyó en las redes sociales, especialmente Facebook, para interactuar con los votantes.¹⁸ Facebook ha destacado cómo el actual presidente, Mauricio Macri, llegó a más del 92 % de los usuarios de Facebook en Argentina durante la última campaña electoral.

¹⁶ <https://www.wsj.com/articles/telecom-argentina-cablevision-announce-merger-1498924578>, <https://www.publicknowledge.org/news-blog/blogs/argentinas-dangerous-path-toward-media-and-communications-dominance>; <http://www.ambito.com/907340-preocupante-aprobo-el-enacom-fusion-telecom-cablevision>

¹⁷ Diapositiva 35 <https://www.slideshare.net/wearesocialsg/digital-in-2017-south-america> Ver también: <https://www.digitalhouse.com/noticias/las-redes-sociales-mas-elegidas-los-argentinos-marketing-digital/>.

¹⁸ https://www.facebook.com/business/success/mauricio-macri#u_0_0; Ver también <http://www.lapoliticaonline.com/nota/99814/>.

02

METODOLOGÍA: QUÉ HICIMOS

El proyecto se diseñó para proporcionar a los usuarios e investigadores de Facebook una ventana hacia el proceso de curaduría algorítmica de contenido llevado a cabo por la red social en momentos cruciales para la formación de opinión pública. El proyecto tiene como objetivo mostrar un método y un conjunto de herramientas que pueden ayudar a profundizar el análisis y debate en torno a estos procesos, y destacar una serie de aspectos en los que Facebook podría empoderar al usuario para controlar los contenidos a los que es expuesto.

A partir del mes de octubre de 2017, se realizaron algunos experimentos iniciales para probar nuestras herramientas y métodos.¹⁹ A principios del mes de diciembre llevamos a cabo una nueva iteración, creando seis perfiles de Facebook en máquinas virtuales para los cuales utilizamos servidores situados en Buenos Aires. Los seis perfiles se crearon el mismo día de diciembre. Para reducir el universo de variables que podrían influir en los resultados, se crearon seis perfiles de mujeres de 30 años: Lucía, Julieta, Brenda, Valentina, María y Andrea.²⁰

¹⁹ Antes de las elecciones legislativas de Argentina del 22 de octubre de 2017, se crearon nueve perfiles de Facebook específicamente para este trabajo de investigación para seguir 40 fuentes noticias en línea y redes sociales. La herramienta se utilizó para recopilar datos durante dos semanas (unos días antes del día de las elecciones y unos días después de las elecciones). Para leer todos los detalles técnicos, consulte el documento de metodología, que proporciona una descripción detallada de los pasos y las iteraciones del experimento.

²⁰ Los seis perfiles hipotéticos en Facebook fueron creados como nuevos perfiles en el mismo día. No se incluyeron más datos personales que el nombre (femenino) y la edad (30). Se compraron tarjetas Sim en Buenos Aires para crear los perfiles usando números de teléfono únicos. Cualquier otro dato personal que se debía proporcionar a Facebook se completó de tal forma que todos los perfiles proporcionaban la misma información.

Se configuraron los seis perfiles para que siguieran las mismas 22 Páginas de Facebook: 20 de ellas administradas por portales de noticias, y dos Páginas políticas (una perteneciente a Mauricio Macri, el actual presidente, y otra creada para Cristina Fernández de Kirchner, ex presidenta y figura clave de la oposición). El 15 de diciembre, una semana después de haber iniciado el experimento, polarizamos nuestros perfiles. De los seis perfiles, dos marcaron “Me Gusta” sobre 30 publicaciones de la página de Facebook de Macri, dos marcaron “Me Gusta” sobre 30 publicaciones de la página de Facebook de Kirchner, y dos permanecieron políticamente neutrales, no marcando ningún “Me Gusta”. El propósito de la variación era determinar si esta ligera polarización generaría cambios en las noticias suministradas a cada usuario de Facebook.



La recopilación de datos se llevó a cabo en 11 días, en dos fases: (1) del 8 al 15 de diciembre, durante la cual se recopilaban datos antes de que los perfiles fueran polarizados, y (2) del 16 al 19 de diciembre, durante la cual se recopilaban datos después de la polarización de los perfiles.²¹ Luego se realizó el análisis comparativo de la información a la que fueron expuestos los seis perfiles a través de sus respectivos News Feed. Este período de tiempo fue de particular interés dado las reuniones de la Organización Mundial de Comercio en Buenos Aires, y las protestas en torno a las reformas del sistema jubilatorio en Argentina.

El sistema de caja negra que caracteriza al algoritmo de News Feed crea un conjunto de desafíos que desalientan la investigación su naturaleza e implicancias.²² En este contexto, elegimos emplear un enfoque de auditoría, habitual en estos escenarios.²³ Los datos fueron recopilados mediante Facebook Tracking Exposed, una extensión Web de código abierto instalada a través de Google Chrome.²⁴ La herramienta recopila los componentes HTML completos de cada publicación pública que aparece en el navegador. Configuramos las seis usuarias de modo que accedieran a Facebook diariamente y a las mismas horas. El acceso fue programado con un UserScript,²⁵ que automatizó las visitas diarias a Facebook de las seis usuarias, visitas que implicaban

21 Los perfiles fueron creados el domingo 5 de diciembre de 2017. El experimento se realizó del 8 al 15 de diciembre de 2017.

22 Devito, M. (2017). From Editors to Algorithms: A values-based approach to understanding story selection in the Facebook news feed. *Digital Journalism*, 5(6), 753-773.

23 Ver ejemplo, Chen, L., Mislove, A., & Wilson, C. (2015). Peeking Beneath the Hood of Uber. *Actas de la Conferencia de Medición de Internet*, 495-508; Bertrand, M., & Mullainathan, S. (2004); y, Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The American Economic Review*, 94(4), 991-1013.

24 <https://facebook.tracking.exposed/>.

25 Tecnología UserScript <https://en.wikipedia.org/wiki/Userscript>; Nuestro script de Auto-scroller <https://facebook.tracking.exposed/autoscroll.user.js>.

hacer un scroll (desplazamiento) del News Feed. Cada scroll expuso a cada usuaria a un promedio de 45 impresiones (o lo que los usuarios suelen llamar publicaciones o posts) por conexión.²⁶ Durante la primera semana de recopilación de datos, cada usuaria se conectó y realizó tres scrolls por día. Después de la polarización de las usuarias, se configuró la automatización para que cada usuaria realizara siete scrolls por día.²⁷ Se programaron los scrolls para que se produjeran durante el día, con el objetivo de aumentar la validez externa (que su comportamiento fuera similar al de usuarias reales), y se coordinaron en todos los perfiles para aumentar la validez interna (comparabilidad entre perfiles).²⁸ Las horas fueron las siguientes:

Primera semana: 08:01, 16:01, 23:55
Últimos días: 8:01, 10:01, 12:01, 14:01, 16:01, 18:01, 20:01

Para documentar la infoesfera (universo total de publicaciones disponibles) con que contaba Facebook, y sobre la cual se apoyó para definir las publicaciones que aparecerían en el News Feed de cada usuaria, utilizamos la API pública de Facebook²⁹: Recopilamos las 11.603 notas publicadas por las 22 páginas de Facebook observadas. Como parte del análisis también se creó un conjunto de variables para representar y poder monitorear tendencias en el uso de palabras clave que fueron de particular relevancia en el país durante el periodo analizado.³⁰

26 La variación se explica por el tipo de publicación: las imágenes y los videos suelen ocupar más espacio que las actualizaciones de texto. Los llamamos impresiones, en lugar de publicaciones, porque un usuario puede estar expuesto a la misma publicación más de una vez.

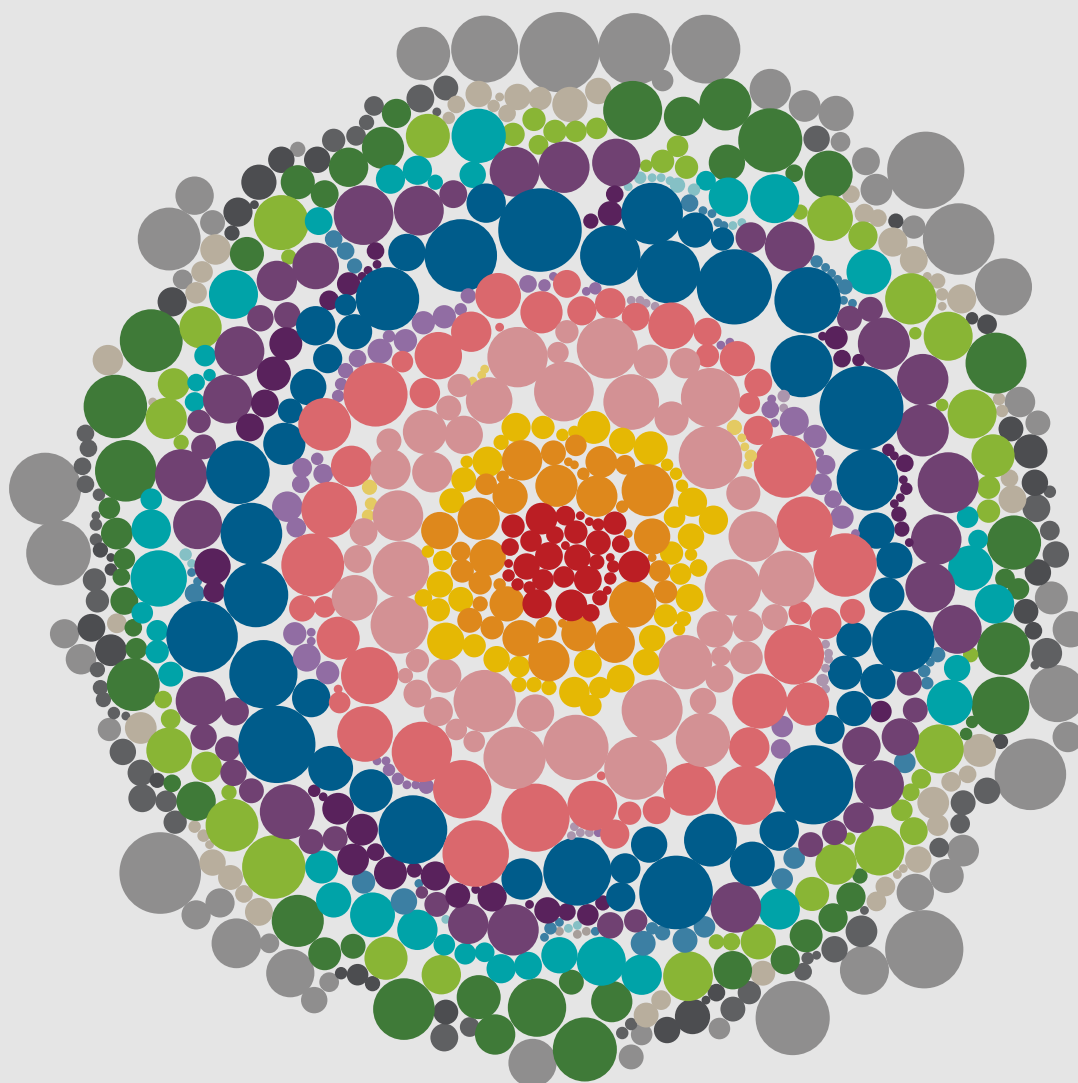
27 Debido a factores fuera de nuestro control, se interrumpió el período de recopilación de datos. Para aumentar el número de puntos de datos para la fase posterior a la polarización, aumentamos el número de desplazamientos diarios, para tener una base de datos más sólida para analizar.

28 Debido a razones técnicas, hubo algunos segundos o minutos de diferencia entre los desplazamientos. Sin embargo, esta diferencia no debería tener un efecto considerable en la exposición de los perfiles a las publicaciones dada la frecuencia de las nuevas publicaciones en las páginas que se estaban siguiendo.

29 Todos los datos recopilados por los perfiles se pueden encontrar aquí: <https://github.com/tracking-exposed/experiments-data> y el Anexo del informe: <https://docs.google.com/spreadsheets/d/1pLixlZtl6vz3faVxUV3wF8ugZpKQzvZyXUw60u-w/edit#gid=1967898836>. El conjunto de datos sin procesar se puede ordenar por: datos de ubicación, cada vez que cada usuario actualizó el News Feed, la fecha y la hora cada vez que un usuario actualizó el News Feed, la hora de la impresión de cada publicación mostrada por el News Feed a un usuario, el Código HTML de cada publicación pública en el News Feed de los usuarios y metadatos relacionados con las publicaciones.

30 Las palabras clave fueron definidas por los investigadores, basándose en un script automatizado que contó la repetición de ciertas palabras clave a través de los artículos de noticias subyacentes rastreadas. Las palabras clave sin procesar se pueden ver explorando los datos disponibles en Facebook Tracking Exposed GitHub <https://github.com/tracking-exposed/experiments-data> y el Anexo del informe: <https://docs.google.com/spreadsheets/d/1pLixlZtl6vz3faVxUV3wF8ugZpKQzvZyXUw60u-w/edit#gid=1967898836>. El conjunto de datos sin procesar se puede ordenar por: datos de ubicación, cada vez que cada usuario actualizó el News Feed, la fecha y la hora cada vez que un usuario actualizó el News Feed, la hora de la impresión de cada publicación mostrada por el News Feed a un usuario, el Código HTML de cada publicación pública en el News Feed de los usuarios y metadatos relacionados con las publicaciones.

Figura 1 — Infoesfera: las 11.603 notas publicadas por las 22 Páginas de Facebook³¹ (del 8 al 19 de diciembre de 2017)



Nombre de la Página

Agencia Paco Urondo	La Política Online
AmericaTV	Mauricio Macri (Presidente)
C5N	Noticiero Trece
Clarín	Página 12
Cristina CFK (Ex-presidente)	Perfil
Cronica	Radio Mitre
Diario Popular	Radio10
Garganta Poderosa	Telenueve Noticiero
Infobae	Tiempo Argentino
INFOnews	TN Todo Noticias
La Nacion	

El tamaño de cada esfera es proporcional a la cantidad de posts publicados. Cada esfera representa un periodo de tiempo entre dos scrolls (ver Metodología)

³¹ Esto incluye TV Pública Noticias, la que no publicó nada durante el período analizado. <https://www.facebook.com/tvpublicanoticias>. La lista completa de páginas, con sus correspondientes URLs y números de publicaciones por sitio durante el período analizado están incluidos en el anexo.

03

ANÁLISIS: HALLAZGOS

A. Facebook sigue usando curaduría intensa del contenido que vemos

Quizás el hallazgo más revelador del estudio es el gran contraste en el universo de publicaciones a las que ya sea por diseño o defecto-fueron expuestos usuarios con el mismo perfil. El resultado es que incluso los usuarios con el mismo perfil, que siguen los mismos sitios, son sometidos a diferentes dietas informativas.

No es ningún secreto que Facebook desempeña una función de filtrado basado en los perfiles de usuario para clasificar la inmensa cantidad de información que se publica en la plataforma social. Los hallazgos de nuestro estudio sugieren que esta forma de filtrar los contenidos a distribuirse se extiende a usuarios con perfiles aparentemente idénticos. Como muestra la Figura 2, de las 11.603 publicaciones publicadas durante los 11 días observados, Andrea solo vio 1.192 publicaciones, y nuestros seis perfiles combinados estuvieron expuestos únicamente a 2.071 publicaciones, tan solo el 18 % de todas las publicaciones publicadas en las 22 páginas de Facebook. Esto equivale a estar expuesto a alrededor de una de cada seis publicaciones de todos los sitios que se siguen. Pero es importante destacar que Andrea fue sometida a una experiencia muy distinta a la de Lucía: A pesar de seguir las mismas páginas y conectarse a la red social al mismo tiempo, aproximadamente la mitad de los componentes de la dieta informativa de Andrea fue excluida de la línea de tiempo de Lucía, y viceversa.³² Es importante entender los factores que subyacen a la decisión del algoritmo de mostrar u ocultar cada una de las 11.603 notas que publicaron las Páginas de Facebook que seguían nuestros perfiles.. La amplitud de esta brecha entre lo que se ve y lo que se oculta representa el gran margen de discrecionalidad con el que cuenta la plataforma, y que se encuentra fuera del control de los usuarios.

Además, es importante comprender si los usuarios de Facebook están al tanto de estos procesos de curaduría algorítmica de contenido . Un estudio reciente, basado en una encuesta de 147 estudiantes estadounidenses, sugiere que la mayoría de los usuarios desconocen que su News Feed está sujeto a estos procesos de selección automatizada de contenidos a distribuir a los usuarios.³³

32 Esta figura se basa en los intervalos de tiempo 1 a 17, antes de que los perfiles estuvieran polarizados. Ver el gráfico interactivo aquí <https://public.tableau.com/profile/web.foundation#/vizhome/Theinvisiblecurationofcontent-Full-English/PieChartProfiles>

33 "Solo el 24 % sabía que Facebook prioriza ciertas publicaciones y oculta otras de las noticias de los usuarios". (...) 37 % cree que el News Feed incluye todas las publicaciones". Ver Powers, E. (2017). My News Feed is Filtered?: Awareness of news personalization among college students. *Digital Journalism*, 5(10), 1315-1335.

Figura 2 — Un diagrama de árbol que muestra qué vieron nuestros perfiles en comparación a lo que no vieron.

(El tamaño de cada cuadro es proporcional a la cantidad de publicaciones).



La plaza pública vista por Facebook

La plaza pública vista por nuestros usuarios

Mientras que las plataformas digitales a menudo aseveran que están permitiendo que la plaza pública se amplíe, lo que observamos es que, mediante el uso de algoritmos, están colocando a cada usuario en versiones separadas de la plaza pública, versiones administradas en detalle por sus algoritmos. Estas decisiones de filtrado colocan a Facebook (y otras plataformas) en una posición de control sobre el debate público, lo que trae consigo responsabilidades que no son menores para estas corporaciones.³⁴

³⁴ Koene A. et al. (2015) Ethics of Personalized Information Filtering. En: Tiropanis T., Vakali A., Sartori L., Burnap P. (eds) Internet Science. INSCI 2015. Lecture Notes in Computer Science, vol 9089. Springer, Cham; Taddeo, M., y Floridi, L. (2016). The Debate on the Moral Responsibilities of Online Service Providers. Science and Engineering Ethics, 22(6), 1575-1603.; Koene, A., Vallejos, E. P., Webb, H., Patel, M., Ceppi, S., Jiroka, M., y McAuley, D. (2017). Responsabilidades editoriales derivadas de los algoritmos de personalización.

B. La curaduría algorítmica de contenidos puede ocultar a algunos usuarios noticias sociales y políticas cruciales.

Hay algunas publicaciones que pueden ser relevantes para la sociedad en general, o para un usuario en particular, pero que en función de las variables que fueron priorizadas en el diseño del algoritmo que gobierna el News Feed—pueden terminar sin mostrárselas a dicho usuario de Facebook.³⁵ Cuando una gran parte de la población se apoya en la plataforma para acceder a noticias, esto rápidamente se puede convertir en un grave problema. Dados estos riesgos, Facebook debe diseñar herramientas que permitan una transparencia total con respecto a qué información se visibilizó y cuál se ocultó de los usuarios.

Figura 3 — Proporción de publicaciones que contienen ciertas palabras clave a las que nuestros perfiles estuvieron expuestos ³⁶



Tres de cada cinco argentinos usan Facebook como su fuente de noticias en línea. Si Facebook fuera la única fuente de información fiable de nuestras seis usuarias argentinas, probablemente no se habrían enterado, por ejemplo, de las historias de feminicidio y violación que se publicaron durante el periodo analizado. Ambas cuestiones son de gran relevancia en el debate público y político, lo que ha provocado protestas (tanto en las redes sociales como en la calle) contra el feminicidio y las violaciones.³⁷

³⁵ Para explorar las diferencias en los temas de publicación que aparecen, o que no aparecen, en el News Feed de los perfiles, consulte: <https://public.tableau.com/profile/web.foundation#!/vizhome/LaManoInvisible-TableroAvanzado/DashReactionsShares>.

³⁶ Más información en <https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/PickabooDash>. Versión en español disponible en <https://public.tableau.com/profile/web.foundation#!/vizhome/LaManoInvisible-TableroAvanzado/DashReactionsShares>.

³⁷ Como se esperaba, el algoritmo es particularmente sensible al número de reacciones y comparticiones, que es en sí mismo una decisión editorial. Sin embargo, como el gráfico muestra que este no es el único factor, y en el caso específico bajo escrutinio, algunas de las historias sobre homicidios tuvieron un desempeño peor que las historias omitidas sobre feminicidio y violaciones que no aparecieron. Ver: <https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/DashboardSharesReacts>.

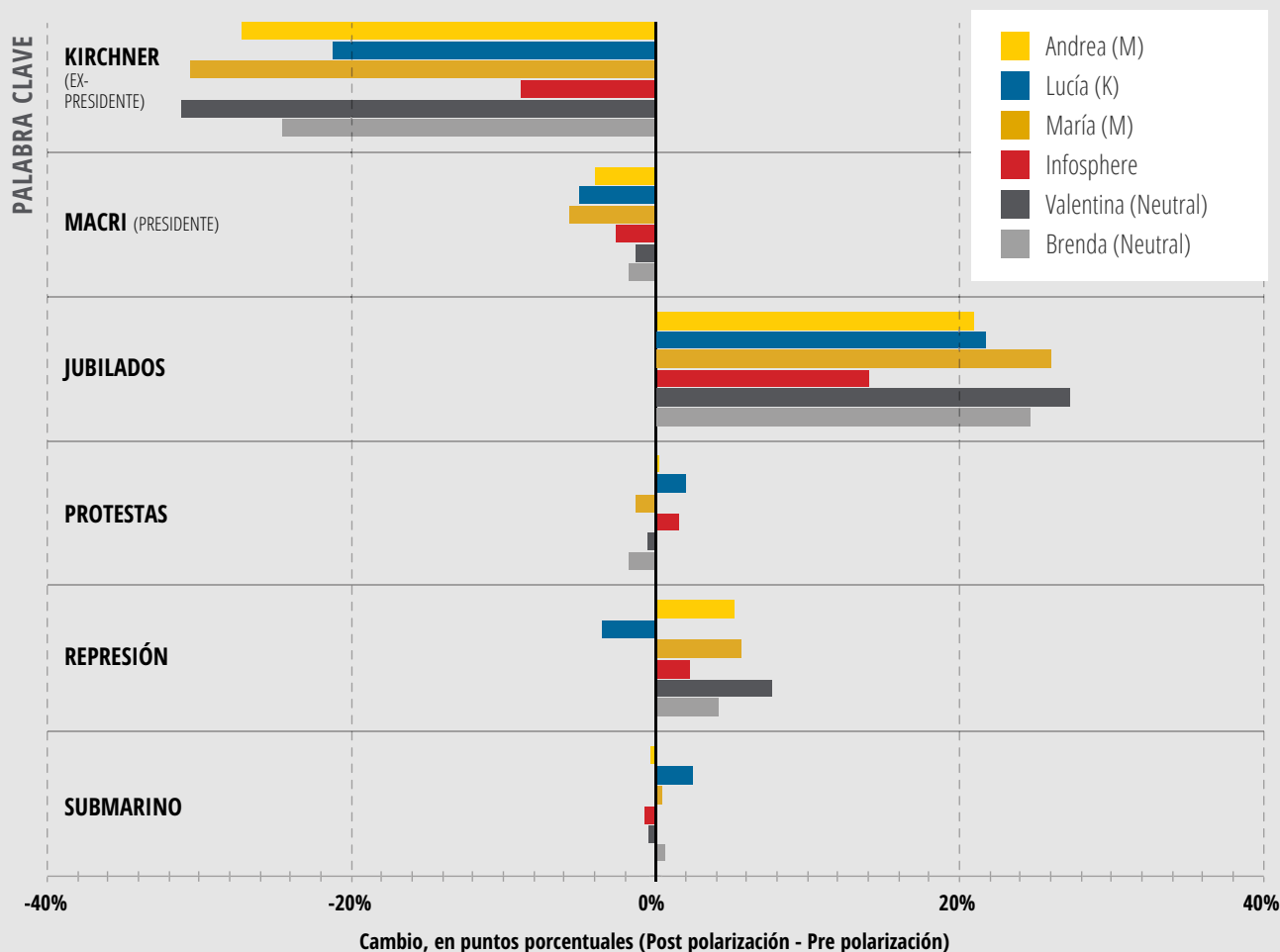
C. La curaduría de contenidos de Facebook significa que usuarios con el mismo perfil que siguen las mismas páginas de noticias no se les expone a las mismas historias publicadas en la plataforma.

La presencia de Facebook en Argentina y en todo el mundo como principal fuente de noticias para muchos ciudadanos significa que sus algoritmos (y las modificaciones que se le hacen) afectan el flujo de información, el debate público y la sociedad en general.

Los experimentos descubrieron que a los distintos perfiles se les mostraron historias diferentes sobre un grupo específico de temas: protestas, represión y la desaparición de un submarino militar.³⁸ Por ejemplo, Lucía K, a quien le gustaban las publicaciones de una figura clave de la oposición, vio más publicaciones

sobre las protestas y el submarino desaparecido que otras en el estudio. Lucía K también estuvo expuesta a menos publicaciones sobre la represión que tuvo lugar en el contexto de tales protestas, lo que parecería contradictorio con la hipótesis de que las historias se seleccionan según los perfiles políticos. Como resultado, aunque el estudio puede demostrar que la distribución de historias difiere incluso para perfiles similares, no puede confirmar si el algoritmo de Facebook reacciona ante pequeñas señales de interés político divergente (ver Figura 4). Es necesario seguir investigando para probar esta hipótesis.³⁹

Figura 4 — Cambio, en puntos porcentuales, de la proporción de publicaciones que contienen palabras clave seleccionadas, antes y después de la polarización⁴⁰



³⁸ El ARA San Juan, que se convirtió en una tragedia nacional y suscitó críticas públicas del actual gobierno de Macri por su manejo del asunto. Ver: <https://www.theguardian.com/world/2017/nov/30/argentina-calls-off-missing-submarine-rescue-effort>.

³⁹ Para más información sobre estos resultados, consulte el Anexo II. Además: gráfico interactivo que muestra la variación en estos porcentajes a lo largo del tiempo <https://public.tableau.com/profile/web.foundation#/vizhome/Theinvisiblecuratorofcontent-Full-English/Selectedkywdsdash> y la distribución <https://public.tableau.com/profile/web.foundation#/vizhome/Distributions-Invisiblecuratorofcontent-English/AdvancedDist>.

⁴⁰ Para reducir las interferencias, se basó este gráfico en un subconjunto más pequeño de publicaciones, en el que la recopilación de datos fue más estable. Prepolarización: 6 de diciembre a las 9:00 p.m. hasta el 12 de diciembre a las 5:00 p.m. Estos se representan como Intervalos de tiempo 1 a 17 en el conjunto de datos y gráficos interactivos. Post-polarización: Entre el 15 de diciembre a las 5:00 p.m. y el 18 de diciembre a las 7:00 p.m. Estos se representan como Intervalos de tiempo 27 a 47 en el conjunto de datos y gráficos interactivos. Nota: El perfil de Julieta se excluyó de este análisis porque su perfil se bloqueó el 17 de diciembre y no fue posible recopilar datos a las 3:00 p.m., como se esperaba (Intervalo de tiempo 38).

04

CONCLUSIONES Y RECOMENDACIONES

Este estudio documenta la dinámica cambiante entre un usuario de Facebook, las fuentes de noticias que sigue, y las noticias a las que es expuesta a través de la plataforma, como resultado de:

- **La curaduría de noticias:** Esto implica que las dietas informativas, tanto en Facebook como en otras plataformas en línea, son definidos por algoritmos que controlan cómo se informa y qué porción de la esfera pública verán los usuarios, lo que a menudo hace que noticias potencialmente relevantes sean ocultadas de los usuarios.
- **La naturaleza opaca de los algoritmos:** Facebook parece definir las dietas informativas de los usuarios según criterios que son invisibles para los usuarios. Además, el algoritmo es lo suficientemente complejo u opaco como para que incluso usuarios aparentemente idénticos que siguen las mismas fuentes de información reciban diferentes publicaciones en sus líneas de tiempo.

El resultado de estos procesos es que cada persona recibe una dieta informativa muy diferente.

El seguimiento de las páginas de noticias de Facebook que realizaron nuestras seis usuarias era el equivalente de antaño de estar suscrito a seis periódicos y recibirlos diariamente. Antes de la aparición de las redes sociales la decisión de “seguir” a los mismos periódicos habría significado que estas seis usuarias habrían estado expuestas a exactamente las mismas noticias, como dos vecinas que reciben los mismos periódicos. La “curaduría algorítmica de contenidos” estaba limitada a lo que cada lectora decidiera que era interesante, y para lo que tuviera tiempo. Ellas determinaban su dieta informativa.

Nuestro estudio descubrió que los algoritmos de Facebook parecen decidir qué periódicos se reciben, cuándo se reciben y su contenido. Los mismos algoritmos también afectan lo que reciben los “vecinos”, lo que traduce en que ninguna persona encontrará los mismos periódicos o los mismos artículos en la puerta de su casa.

En el caso de Argentina, un país con una intensa participación en las redes sociales, se crea un escenario donde los usuarios de Facebook ven el mundo a través de estos algoritmos.

Estos hallazgos sugieren que:

- La curaduría algorítmica de contenido es persistente y parece depender de factores aparentemente arbitrarios, que los usuarios no controlan ni informan de manera activa.
- Los usuarios necesitan comprender mejor cómo la tecnología funciona y afecta sus dietas informativas, y consiguientemente, su percepción de la realidad. Esto se puede hacer explicando el fenómeno con ejemplos que muestran la dinámica detrás de la curaduría algorítmica del contenido.
- Las plataformas, especialmente las que desempeñan un papel central en la definición de la dieta informativa de un país, deben revelar sus prácticas de curaduría algorítmica de contenido, así como estar abiertas a revisar esos mecanismos con los usuarios, e informarles cómo se estructuran sus dietas informativas.
- Es necesario comprender mejor si las leyes y los reglamentos vigentes son adecuados para abordar los desafíos planteados por los algoritmos que personalizan la distribución de información.

Algunas sugerencias para que Facebook comience a abordar estos desafíos incluyen:

- **Explorar el uso de paneles de transparencia en tiempo real**, que permitan a los usuarios contrastar las historias a las que fueron expuestos con el universo total de historias publicadas durante un período de tiempo determinado, a fin de aumentar la comprensión pública sobre el funcionamiento del algoritmo.
- **Que la publicación de contenido en orden cronológico** sea la forma predeterminada para su distribución a través de Facebook.⁴¹
- **Permitir a los usuarios una variedad de opciones para administrar su propia configuración del News Feed**, y así desempeñar un papel más activo en la construcción de sus dietas informativas.⁴²
- **Responder a las necesidades de los usuarios de Facebook**, generando más espacios para que comenten sobre su experiencia con los algoritmos de la plataforma, y permitiéndoles un rol más activo en las decisiones que se implementarán en la plataforma.
- **Facilitar estudios independientes** sobre las externalidades y los efectos sistémicos de la curaduría algorítmica del contenido.

Entendemos que estas sugerencias son una forma en que Facebook puede aproximarse a su objetivo de “brindar a las personas el poder para construir una comunidad y acercar al mundo”. Son pasos fundamentales para devolver a los usuarios control de sus dietas informativas.

⁴¹ La configuración actual permite a los usuarios optar por inhabilitar la autenticación algorítmica. Sin embargo, en el contexto de las pruebas informales, hemos observado que el reinicio de la aplicación restablece esta preferencia al valor predeterminado de autenticación algorítmica. Ver <https://www.facebook.com/help/community/question/?id=10152476808951132>.

⁴² Ver por ejemplo gobo.social.

05

REFERENCIAS

Anexo I

Lista de páginas de Facebook seguidas

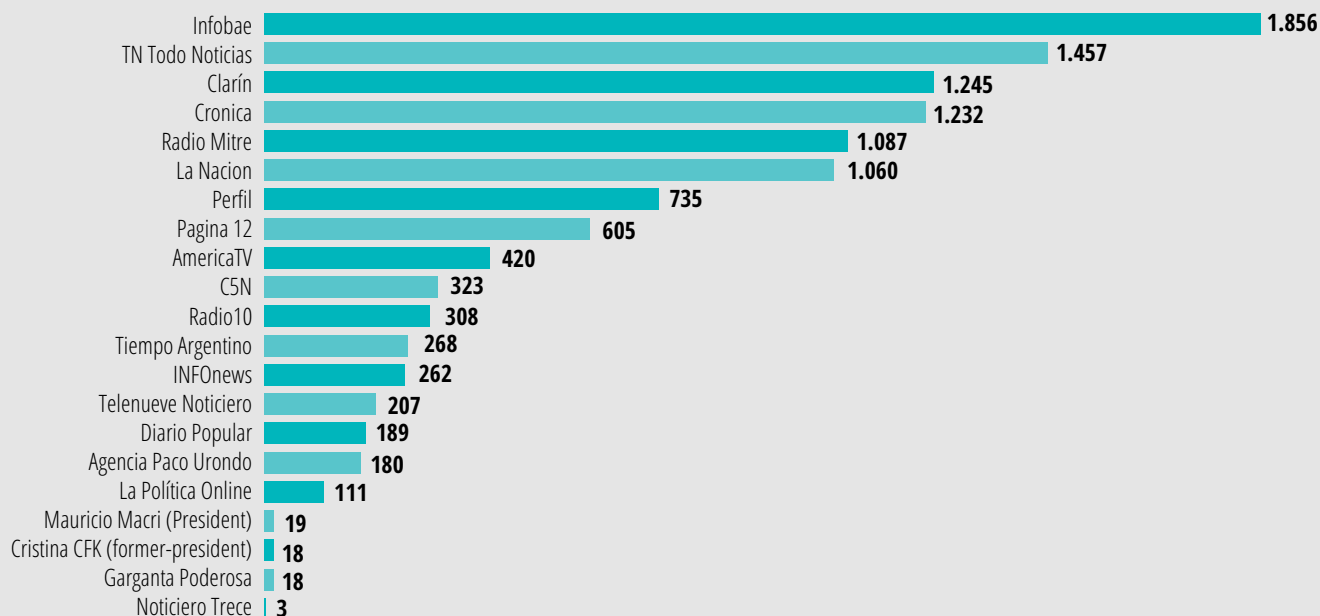
TIPO	NOMBRE	HIPERVÍNCULO DE FB	SEGUIDORES (2 DIC 2017)
Blog	Agencia Paco Urondo	https://www.facebook.com/agenciapacourondo/	120.636
Televisión	América 2	https://www.facebook.com/AmericaTV/	1.435.601
Televisión	C5N	https://www.facebook.com/C5N.Noticias	4.082.608
Televisión	Canal 13	https://www.facebook.com/NoticieroTrece/?ref=br_rs	93.675
Televisión	Canal 7	https://www.facebook.com/tvpublicanoticias/	144.586
Televisión	Canal 9	https://www.facebook.com/telenueveC9/?ref=br_rs	81.314
Periódico	Clarín	https://www.facebook.com/clarincom/	5.684.851
Político	Cristina Fernández de Kirchner	https://www.facebook.com/CFKArgentina/	2.358.232
Periódico	Crónica	https://www.facebook.com/cronicadiario/	1.435.565
Periódico	Diario Popular	https://www.facebook.com/populardiario/	112.030
Blog	Garganta Poderosa	https://www.facebook.com/La-Garganta-Poderosa-213440425391495/?fref=ts	505.251
Periódico	Info News	https://www.facebook.com/INFOnewscom	103.130
Periódico	Infobae	https://www.facebook.com/Infobae/	2.312.842
Periódico	La Nación	https://www.facebook.com/lanacion	3.342.628
Blog	La Política Online	https://www.facebook.com/pages/lapoliticaonline/304176325478	132.228
Político	Mauricio Macri	https://www.facebook.com/mauriciomacri/	4.488.499
Radio	Mitre	https://www.facebook.com/radiomitre/	1.458.622
Periódico	Página 12	https://www.facebook.com/Pagina12ok/	391.303
Periódico	Perfil	https://www.facebook.com/perfilcom	644.724
Radio	Radio 10	https://www.facebook.com/radio10/	398.785
Periódico	Tiempo Argentino	https://www.facebook.com/DiarioTiempoArgentino/	331.903
Televisión	TN	https://www.facebook.com/todonoticias/?ref=br_rs	6.825.397

TOTAL

36.484.410

Total de publicaciones por sitio (durante el período de observación):

Nombre de la Página



Palabras clave seleccionadas para variables ficticias (utilizadas para probar durante la polarización)

NOMBRE DE LA VARIABLE	BOOLEANO	EXTRA (CON CONTROL MANUAL DE LOS RESULTADOS)
Mención CFK	CFK Cristina Fernandez Cristina Kirchner	Cristina
Mención Macri	Mauricio Macri Macri	el presidente Mauricio
Movilización (y vars)	Protesta Movilizació marcha cacerolazo	Ninguna
Jubilados	previsional jubila	Ninguna
Cacerolazo	Cacerola	Ninguna
Represión	Represio reprimi balas de goma gas lacrimógeno	Ninguna
Submarino	ARA San Juan Submarino	Ninguna
Inflación	inflacio	Ninguna

(Anexo II)

Documentación técnica

<http://webfoundation.org/docs/2018/01/Annex-FB Tests OSF.xlsx>

Investigación conjunta

Comparta sus hallazgos y llevemos a cabo un esfuerzo conjunto para comprender mejor el funcionamiento de Facebook. Si está utilizando el conjunto de datos, le recomendamos que comparta este conjunto de datos utilizando el hashtag #LaManoInvisible #InvisibleHand o simplemente comparta su investigación enviándonos un correo electrónico a contact@webfoundation.org.

Encontrará gráficos interactivos en: <https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/PickabooDash> (inglés)

<https://public.tableau.com/profile/web.foundation#!/vizhome/LaManoInvisible-TableroAvanzado/DashReactionsShares> (español)



WORLD WIDE WEB
FOUNDATION

World Wide Web Foundation
1110 Vermont Ave NW, Suite 500
Washington, DC 20005, USA

www.webfoundation.org

Twitter: [@webfoundation](https://twitter.com/webfoundation)