

Contenido compensatorio

Un análisis del contenido que desafía el extremismo en línea

Jamie Bartlett
Alex Krasodonski-Jones

Octubre 2015

DEMOS

Open Access. Some rights reserved.

As the publisher of this work, Demos wants to encourage the circulation of our work as widely as possible while retaining the copyright. We therefore have an open access policy which enables anyone to access our content online without charge. Anyone can download, save, perform or distribute this work in any format, including translation, without written permission. This is subject to the terms of the Demos licence found at the back of this publication. Its main conditions are:

- Demos and the author(s) are credited
- This summary and the address www.demos.co.uk are displayed
- The text is not altered and is used in full
- The work is not resold
- A copy of the work or link to its use online is sent to Demos.

You are welcome to ask for permission to use this work for purposes other than those covered by the licence. Demos gratefully acknowledges the work of Creative Commons in inspiring our approach to copyright. To find out more go to www.creativecommons.org



Published by Demos October 2015

© Demos. Some rights reserved.

Unit 1, Lloyds Wharf
2-3 Mill Street
London
SE1 2BD

hello@demos.co.uk
www.demos.co.uk

This document was kindly translated into Spanish by Laura Merino Díez.

This paper was supported by Facebook. The views expressed are those of the authors and do not necessarily reflect those of Facebook. All errors and omissions are the authors' own.

CONTENIDO

Prólogo	4
Metodología	6
Resultados	8
Recomendaciones	14
Formas de abordar el contenido compensatorio en el futuro	15
Notes	18

PRÓLOGO

Casi 1.5 millón de personas emplean Facebook en todo el mundo. Aunque la mayoría de los usuarios utilizan el sitio con fines positivos, algunos emplean la plataforma con malas intenciones. En este sentido, Facebook ha creado las normas de la comunidad, un conjunto de políticas que detalla qué tipo de contenido es apropiado publicar y qué queda prohibido. Por ejemplo, Facebook prohíbe y elimina contenido que incite al odio, que define como "contenido que ataca directamente a individuos en función de su raza, origen étnico, nacionalidad, religión, orientación sexual, sexo, género, identidad de género, discapacidad grave o enfermedades". A pesar de que la incitación al odio está prohibida, en ocasiones ciertos usuarios publican contenido desagradable o perturbador que no infringe las políticas de Facebook.

Para contrarrestar ese tipo de contenido desagradable o extremista, Facebook ha declarado públicamente que el contenido compensatorio (y las herramientas que su plataforma ofrece para promoverlo) desempeñan un papel fundamental en la lucha contra dichas malas prácticas. Facebook opina que esta forma de atajar el problema no es sólo resulta potencialmente más eficaz, sino que además es más probable que tenga éxito a largo plazo.

Contenido compensatorio podría definirse como una respuesta común contra el extremismo o la incitación al odio fomentada por multitud de usuarios. La reacción de los usuarios a este tipo de mensajes extremos a menudo resulta en discrepancia, burla o campañas en contra. Esta forma de luchar contra el extremismo presenta varias ventajas: resulta más rápida, flexible y reactiva, potencialmente originada desde cualquier parte del mundo en cualquier idioma y conserva el principio de respeto por los espacios de debate públicos y accesibles. Sin embargo, el contenido compensatorio puede adoptar tantas formas como el extremismo y, además, es probable que no siempre resulte tan eficaz como podría ser. Algunos tipos de contenido compensatorio podrían incluso resultar contraproducentes.

Facebook confía en la eficacia del contenido compensatorio y ha demostrado un interés creciente en un enfoque riguroso y basado en datos verídicos que permita profundizar en el tema. Por ello, la red social encargó demostraciones con el fin de elaborar una serie de informes de investigación que analizaran en qué medida se crean y comparten diferentes tipos de contenido compensatorio en Facebook.

Este breve informe provisional resume las conclusiones de la fase I en la que se analiza cómo se crea y comparte el contenido que desafía páginas populistas de derecha en toda Europa. Otros informes de esta serie examinan el contenido que desafía la ideología islamista extrema en el Reino Unido entre otros países.

Es muy difícil crear una definición objetiva de una página que incita al odio, y el presente informe no afirma que las páginas incluidas en este estudio o su contenido inciten al odio. En su lugar, el presente estudio se centra en páginas de Facebook populistas y de derecha que han recibido acusaciones frecuentes de crear o compartir un gran volumen de contenido que incita a la violencia o el odio. En adelante, se hará referencia a dichas páginas como páginas "populistas y de derecha" donde se publica una amplia gama de contenidos. A la hora de identificar páginas con contenido compensatorio, se realizó un esfuerzo por identificar de forma explícita las páginas dedicadas en gran medida a compensar o contrarrestar cualquier página o contenido que se considerara como instigadora de la violencia o el odio.

Creemos que es importante que se preserve el principio de libertad de expresión en Internet, ya que la red debería ser un lugar donde todo usuario sienta que puede expresarse libremente y decir lo que piensa. Por tanto, creemos que el debate, el desacuerdo y el desafío de ideas son casi siempre preferibles a la censura y la retirada de contenidos, incluidos los contenidos extremos o radicales sea cual sea su origen. Sin embargo, también creemos que este principio debe cimentarse en una base empírica para ayudarnos a comprender mejor fenómenos de este tipo y cómo responder. Ese es precisamente el objetivo de este proyecto de investigación.

METODOLOGÍA

Para la realización de este proyecto se recopilaron las publicaciones públicas y los datos de interacción de las páginas públicas de Facebook de 150 sitios de populistas de derechas y de sitios con contenido compensatorio de Reino Unido, Francia, Italia y Hungría a través de la interfaz de programación de aplicaciones (API) pública de Facebook. Para ello, se empleó 'R', un software de código abierto que permite a los investigadores acceder a los datos de carácter público de la API.

En ningún momento se intentó recopilar información personal sobre usuarios ni identificarlos como individuos. Si en algún momento quedaron registrados el nombre de un usuario o su ID de forma accidental, dichos datos se eliminaron automáticamente. En ningún momento se recopiló información sobre grupos ni páginas individuales ni páginas privadas o secretas. Durante todo el estudio se recopiló información sobre páginas de acceso público y visible para todos los usuarios. Con el fin de garantizar la protección de la privacidad individual, no se han citado ni publicado ninguna publicación específica que pudiera dar lugar a la identificación de usuarios concretos.

La selección de las páginas se realizó a través de investigadores de Demos, y por lo tanto, no se deben considerar como una muestra representativa de las páginas pertinentes. Esta forma de selección no es una forma óptima de identificar páginas relevantes, y es necesario señalar que este es un proyecto piloto. En el futuro, otros estudios afinarán la metodología.

Durante un periodo de 2 meses (del 1 de octubre al 1 de diciembre de 2014) se recopilaron 27.886 mensajes publicados en 150 páginas de Reino Unido, Francia, Italia y Hungría, aunque la mayoría del análisis se centró en Reino Unido, Francia e Italia. En este contexto, "mensajes" hace referencia a las actualizaciones que el administrador publicó en la página. Además de los mensajes, también se recopilaron todas las interacciones asociadas, es decir, todos los comentarios, 'me gusta', u otras acciones relacionadas. Los datos sobre interacciones asociadas pueden resultar útiles para estimar el alcance de los mensajes publicados, ya que cada vez que un usuario interactúa con un mensaje, este aparece en la línea del tiempo de sus amigos (dependiendo de la configuración de privacidad). El número total de interacciones asociadas asciende a 8,4 millones.

Una vez recopilados, los datos se sometieron a una serie de análisis que incluyen calcular la media de interacciones por mensaje de acuerdo con los resultados automatizados de la API, calcular el formato de los tipos de datos más populares a través de los datos automatizados de la API, calcular el tipo y el estilo de los tipos de contenido más populares a través de análisis manual humano, calcular los tipos de contenido que ocurren en diferentes páginas a través de análisis manual, calcular el tipo de red de dichas páginas a través del análisis de red automatizado, calcular de la

forma en que diferentes tipos de contenidos se comparten en las páginas en comparación con los propios Titulares de los usuarios que utilizan el análisis automatizado.

Es importante hacer hincapié en que en numerosas ocasiones este estudio emplea metodologías experimentales, puesto que en la actualidad no existen mejores prácticas o estándares concisos para la recopilación y el análisis de datos de esta naturaleza. Dado el carácter exploratorio del presente estudio, los resultados presentados deben interpretarse con cautela.

RESULTADOS

Datos generales sobre el tamaño y la escala de la incitación al odio y el contenido compensatorio

En total, en los cuatro países se identificaron 124 páginas populistas de derecha, en comparación con 26 páginas de contenido compensatorio (por razones explicadas con anterioridad). Asimismo, se publicaron muchos más mensajes en las páginas populistas de derecha (25.522) que en las páginas de contenido compensatorio (2.364). Como era de esperar, también hubo muchas más interacciones en páginas populistas de derecha (un total de 7,8 millones) que en páginas de contenido compensatorio (546.000).

Debido al método de recopilación de datos, la información objeto de estudio incluye muchas más páginas populistas de derecha que de contenido compensatorio. Esto no significa necesariamente que haya más páginas populistas, pero sí es probable que las páginas de contenido compensatorio sean más difíciles de identificar a primera vista. Por lo tanto, con el fin de presentar una imagen más precisa, se calculó la media de comentarios y de interacciones por mensaje en cada país y para cada tipo de página. El resultado demostró que, en general, en Reino Unido hay menos páginas de contenido compensatorio y con una actividad más limitada, pero se comparten y comentan más que las populistas.

Tabla 1: promedio de interacciones por mensaje en cada país

País	Contenido compensatorio	Contenido populista
Francia	126	285
Italia	37	235
Reino Unido	402	325

Análisis de la estructura de la red y del impacto en la forma de compartir el contenido

Con el fin de comprender mejor la información y las ideas divulgadas en dichas páginas, se seleccionaron las páginas de Reino Unido y se examinó el grado en que las personas que comentaron en una página comentaron en otras. Estos datos se obtuvieron de la API pública de Facebook a través del programa 'R' y se visualizaron en Gephi, una herramienta de análisis de redes de código abierto.

En el análisis de las 92 páginas populistas se registró la participación de un total de 54.495 usuarios únicos con 159.437 mensajes durante el período de estudio. El cálculo se realizó mediante la recopilación de los comentarios y de los códigos de identificación de usuarios participantes. Se estima que 16,2% de los usuarios participan activamente en dos o más páginas y 1,3% lo hace en cuatro o más.

A continuación, se analizaron las 21 páginas de contenido compensatorio. Un total de 116.534 usuarios únicos participó estas páginas con 135.842 interacciones. A pesar de tener más usuarios, estas páginas presentan una menor densidad de red: 12,7% de los usuarios participan en dos o más páginas y 1,15% son activos en cuatro o más.

Esto sugiere que las páginas populistas de Reino Unido tienen una comunidad ligeramente más concentrada: con menos usuarios, pero más activos. Las páginas de contenido compensatorio tienen colaboradores más activos, pero no producen tanto contenido por persona. Sin embargo, este estudio debe repetirse a mayor escala para que las conclusiones sean válidas y significativas.

Tipos de mensajes más eficaces para llegar a un mayor público

Con el fin de determinar qué tipo de mensajes surtieron mayor efectividad a la hora de llegar a un mayor público, examinamos a) el formato de los mensajes y b) el contenido y el tono de los mensajes. Los datos de la API de Facebook permiten a los investigadores determinar qué formato adquieren los mensajes. Se dividieron en vínculo, foto, estado o vídeo. La tabla a continuación muestra qué tipo de formato se empleó con mayor frecuencia en las páginas en cuestión.

Tabla 2 tipos de contenido por país (redondeados)

	País	Vínculo	Foto	Estado	Video
Contenido compensatorio	Francia	64	15	14	6
	Hungría	68	5	26	2
	Italia	29	45	17	9
	Reino Unido	40	34	18	8
Contenido populista	Francia	26	43	17	14
	Hungría	10	83	2	5
	Italia	59	27	9	5
	Reino Unido	57	23	10	10

Como se puede observar, existe una gran disparidad entre países y páginas en la forma más frecuente de compartir contenido. Por ejemplo, en Francia los vínculos son la formato más usado en páginas de contenido compensatorio, mientras que en Italia se prefieren las fotos.

El grado en que un tipo de formato es mejor que otro depende en gran parte de varios factores. Sin embargo, sí es posible extraer conclusiones al calcular el promedio de interacciones por cada tipo de formato. La más importante: las fotos suelen generar más interacciones con los usuarios (ver tabla 3, a continuación).

Tabla 3 tipo de contenido más frecuente por país

	Francia	Hungría	Italia	Reino Unido
Contenido populista	Foto	Vídeo	Foto	Foto
Contenido compensatorio	Foto	Foto	Foto	Foto

Análisis de los mensajes publicados

Con el fin de comprender mejor qué tipo de mensajes y qué tono son más frecuentes, se analizaron 1,000 mensajes de páginas populistas y otros 1,000 de contenido compensatorio. Dichos mensajes fueron seleccionados de forma aleatoria, pero equitativa entre Francia, Italia y Reino Unido. Los mensajes se dividieron en dos categorías: contenido y tono del mensaje. Las categorías fueron seleccionadas por los investigadores y se describen con ejemplos a continuación. La popularidad de los mensajes se determinó en función del número de interacciones que recibieron.

El resultado determinó que el contenido más usado en los mensajes de las páginas populistas corresponde con los comentarios (2.524 interacciones de media)¹ que recibieron alrededor del doble de interacciones que otros contenidos.

El tipo de contenido más usado en las páginas de contenido compensatorio fueron las preguntas (10.934 interacciones de media, aunque esta cifra queda distorsionada por unos pocos mensajes que recibieron un gran número de interacciones). El segundo contenido más usado en las páginas de contenido compensatorio fue el comentario, aunque en Francia los "ataques" resultaron ser la categoría más popular.²

El tono más frecuente de los mensajes en las páginas populistas fue festivo. El mejor ejemplo son los mensajes publicados como homenaje a los caídos durante una guerra o la celebración del orgullo patriótico (6.607 interacciones de media).³ El segundo tono más empleado refleja la manifestación del enfado (4.093 interacciones de media).⁴

El tono más empleado en los mensajes de las páginas de contenido compensatorio en los tres países fue perspicaz o satírico (2.717).⁵

Con el fin de analizar mejor los tipos de mensajes más comunes, se examinaron 100 mensajes en inglés de páginas de contenido compensatorio etiquetados como satíricos, puesto que resultó ser el tipo de mensaje más común.

En términos de las interacciones, la inmigración, la raza y la religión recibieron de media más de 4.000 interacciones y fueron los tipos de contenido más populares. Por lo general, estas páginas tienden a parodiar el lenguaje extremista empleado en estos temas por las páginas de incitación al odio. El contenido que ganó más fama fue una petición de un regimiento del ejército que pidió que se eliminara una imagen de la página del partido de extrema derecha "Britain First" en la que se ensalzaba al ejército. Una captura de pantalla de la solicitud recibió más de 9.000 veces.

Un análisis del tipo de comentarios y su naturaleza

A diferencia de los mensajes, que sólo puede publicar el administrador, cualquier usuario puede publicar un comentario a un mensaje. Con esto en mente, se analizaron 1.000 comentarios de mensajes en páginas populistas y otros 1.000 de páginas de contenido compensatorio. Tanto los mensajes como los comentarios se seleccionaron al azar y de forma equitativa entre Francia, Italia y Reino Unido. Las categorías fueron seleccionadas por investigadores que midieron la fama de un comentario por el número de "me gusta" que recibió (el único tipo de interacción disponible en un comentario).

En las páginas populistas, 9% de todos los comentarios se clasificaron como contenido compensatorio, es decir, comentarios que se muestran en desacuerdo con el mensaje original o que presentan un mensaje alternativo de carácter más positivo. Las cifras para cada país fueron de 17% en Francia, 7% en Italia, y 5% en el Reino Unido). Si se extrapola esta información, se podría inferir que, sólo en Reino Unido, se escriben cada mes hasta 25.000 comentarios de contenido compensatorio en páginas populistas.

En las páginas de contenido compensatorio, los mensajes de carácter constructivo y en contra del odio o la discriminación constituyeron el tipo de comentario más frecuente (una media de 5,9 "me gusta" por comentario).⁶ El segundo tipo de mensaje más común fue el "debate constructivo" (una media de 5,3 "me gusta" por comentario).⁷ Por el contrario, el contenido compensatorio no constructivo recibió en 3,3 "me gusta" por comentario y la corroboración de datos o hechos, 3,8.⁸ Sin embargo, a pesar de su popularidad, el contenido compensatorio constructivo representó sólo 6% de todo el contenido, en comparación con 20% del contenido compensatorio no constructivo.

Con el fin de comprender mejor el tipo específico de comentarios más comunes, se analizaron 100 comentarios constructivos realizados en inglés en las páginas de contenido compensatorio (las categorías fueron seleccionadas por los investigadores). Esto demuestra que el "debate de temas políticos generales" acapara el mayor número de "me gusta" por mensaje (13,2), seguido de "cuestionamientos /

detalles sobre movimiento políticos" (8,7). Esto sugiere que los comentarios sobre áreas específicas, detalles o información sobre política resultan particularmente efectivos a la hora de propagarse a una amplia audiencia. Sin embargo, este tipo de comentarios supone una proporción mínima del número total de comentarios (6% y 3%, respectivamente).

Análisis sobre dónde y por quién se comparte y debate contenido

A continuación, se examinó la forma en que el contenido se propaga de páginas al propio feed de noticias de los usuarios, y si el lugar donde se publica dicho contenido supone una diferencia en la forma de interactuar. Se analizaron las interacciones de 4.388 mensajes de contenido compensatorio y 75.132 mensajes populistas. También se calculó la proporción de mensajes compartidos desde la página original o desde otro usuario. Este análisis se basa en una muestra de datos mayor que la utilizada para la investigación anterior ya que aunque las páginas son las mismas, el período de recopilación de datos fue más largo. Esta muestra es el único contenido que se recopiló con la asistencia del equipo científico de datos de Facebook y que consistió en datos históricos, agregados y no experimentales.

Tabla 4 dónde aparecen las interacciones en los mensajes

	Mensajes populistas			Mensajes de contenido compensatorio		
	% de la página original	% de otro usuario	N=	% de la página original	% de otro usuario	N=
Número de "me gusta"	94	6	4,138,425	82	19	56,884
Número de veces que se compartió	97	3	474,558	99	1	13,430
Comentarios	73	27	2,173,678	60	41	16,232

Esto sugiere que, como porcentaje de "me gusta" totales, las páginas de contenido compensatorio reciben más "me gusta" y son compartidas más a menudo. Es decir, los usuarios prefieren interactuar con el feed de otro usuario que con la página original (aunque también les guste la página de origen de la publicación). Esto significa que el contenido tiene la capacidad de llegar más lejos y podría tener un mayor alcance, si los usuarios lo compartieran más ampliamente con sus amigos. Además, creemos firmemente que es muy probable que haya una gran cantidad de comentarios o debates originados en los contenidos compartidos, es decir, en los feeds de noticias de sus amigos. Es probable sea precisamente ahí donde se crea un gran parte de contenido compensatorio, aunque no recopilamos la información necesaria para confirmarlo por razones de privacidad.

Con el mismo método también examinamos la forma en que los diferentes tipos de contenido llegan a los usuarios a los que no les gusta la página originaria de publicación (en otras palabras, ¿puede propagarse este contenido más allá de aquellos usuarios a los que hayan gustado una página). Calculamos este dato averiguando el porcentaje de usuarios a los que les gustaba o comentaban los mensajes, a pesar de no haberles gustado la página donde se publicó el mensaje originariamente.

Tabla 5 ¿hasta qué punto puede llegar la investigación de contenido más allá de los usuarios que apretan al botón de "me gusta"?

	% de "me gusta" por parte de gente a la que no le gusta la página		% de comentarios por parte de gente a la que no le gusta la página	
	<i>Populista o de derechas</i>	<i>Contenido compensatorio</i>	<i>Populista o de derechas</i>	<i>Contenido compensatorio</i>
Enlaces	50	18	66	32
Fotos	52	5	65	25
Vídeos	68	26	75	48
Estado	21	7	41	23

Esto muestra que las páginas populistas son significativamente más eficaces en la publicación de contenido que va más allá de su red de fans de la página. Tanto para las páginas de contenido como para las populistas, los vídeos representan el tipo de contenido más eficaz para publicar y llegar a un mayor público.

RECOMENDACIONES

De acuerdo con las conclusiones anteriores, las páginas de contenido compensatorio no son tan activas como las páginas populistas. En Reino Unido registran cifras positivas en términos de interacciones medias por mensaje, pero no tanto en Francia o Italia. Si quieren llegar a más gente, este tipo de páginas en Francia o en Italia deben producir más contenido. En Reino Unido, tienen más colaboradores, pero interactúan con menor frecuencia que los usuarios de páginas populistas.

Para aquellos que deseen fomentar la difusión de contenido compensatorio, los resultados de esta investigación también sugieren que un cambio en el enfoque sería de gran ayuda. En concreto:

- Para administradores, publicar más fotos y vídeos en proporción con su producción total. También deberían centrarse en el contenido que puede llegar más allá de la red de personas a las que les gusta la página.
- Para aquellos que comentan, más contenido compensación constructivo en lugar de negativo o destructivo, y más comentarios sobre cuestiones políticas específicas.
- Para los colaboradores de páginas de contenido compensatorio, no estaría mal alentar a amigos a compartir más contenido con sus contactos.
- En general, esto sugiere que si los administradores y los usuarios de dichas páginas se mostraran más activos, y cambiaran su contenido ligeramente, este nuevo enfoque podría aumentar dramáticamente el alcance de sus mensajes.

FORMAS DE ABORDAR EL CONTENIDO COMPENSATORIO EN EL FUTURO

El contenido compensatorio es un fenómeno más complejo que lo que pueda parecer. Es importante entenderlo como una acción que va mucho más allá de un simple desacuerdo o una confrontación de un contenido en una página pública. A veces toma forma explícita, como cuando desafía un punto de vista que aparecen en un feed, o incluso cuando se busca contenido que incita al odio activamente para desafiarlo y cuestionarlo. Otras respuestas son menos explícitas: algunos usuarios denuncian la incitación al odio, otros bloquean o silencian al usuario o, simplemente, comparten el desacuerdo con terceros en un mensaje privado. Para otros, consiste en la creación de grupos humorísticos o serios para oponerse a una página o a un usuario. En este tipo de páginas, hemos encontrado discusiones y debates constructivos. En otras ocasiones nos encontramos con contenido compensatorio menos constructivo, como enviar amenazas agresivas y abusivas. Más específicamente se identificaron los siguientes tipos: contenido compensatorio constructivo, contenido compensatorio no constructivo, corroboración de datos o hechos, debate constructivo, oposición satírica y páginas de oposición en un tono serio. Algunos de estos formatos son más valiosos que otros.

Debido a que hay tantos tipos diferentes de contenido compensatorio, resulta difícil determinar lo que constituye un resultado efectivo. No existe un consenso actual sobre valores clave, pero pueden dividirse en tres tipos de métricas (todas ellas medibles con diferentes grados de dificultad).

Métricas cuantitativas

En la actualidad, el análisis cuantitativo de contenido procedente de los medios sociales se ha convertido en un lucrativo negocio. Cientos de empresas han construido miles de herramientas que tratan de cuantificar la influencia de medios sociales. Gran parte de este análisis es inadecuado para el análisis académico, pero muchos de los conceptos centrales son aplicables a este estudio:

- La participación es una forma de cuantificar la interacción del usuario. Por ejemplo, podría ser la proporción de usuarios que se inscribieron después de ver la página, o el número de usuarios que compartieron un contenido después de verlo. Esto constituye un instrumento contundente a la par que un indicador útil del alcance potencial. En el presente estudio, hemos sido capaces de cuantificarlo eficazmente.
- El volumen y la exposición ¿Cuántos mensajes se están creando sobre el tema, y el cuántos usuarios únicos lo están debatiendo? ¿Existen picos, donde la densidad del debate es mayor de lo habitual? Responder a estas preguntas es factible, pero requiere acceso a más datos de tráfico en lugar de a datos específicos sobre una página.

- El alcance mide la difusión de una conversación en los medios sociales ¿Cuál es el tamaño de la audiencia? ¿Se limita la incitación al odio a ciertas comunidades aisladas (ya sea por elección propia o a través de los algoritmos de personalización de Facebook)? ¿Con qué frecuencia se extiende a los feeds de otros usuarios fuera de estas comunidades? El alcance de un mensaje o un tema puede constituir una poderosa métrica cuando se combina con la participación, por ejemplo. En el presente estudio, hemos encontrado algunas medidas útiles para fines de investigación.

Métricas de contenido

Las métricas de contenido se centran en un intento por evaluar el tipo y la naturaleza del contenido, en lugar del volumen. A menudo obtener dichos datos requeriría un análisis manual, o bien un sofisticado procesamiento del lenguaje natural basado en análisis. Ambos son posibles, pero con diferentes grados de precisión. Uno de los puntos fuertes de estos enfoques residen en que pueden ser ampliados y aplicados en enormes volúmenes de contenido a gran velocidad.

El análisis de los sentimientos es una técnica que, dado el presente marco analítico, nos permitirá determinar si un fragmento puede clasificarse como contenido de incitación al odio, contenido compensatorio o ninguno de los dos. También se podría evaluar el nivel de extremismo de un contenido mediante el procesamiento del lenguaje empleado. Pensamos que alcanzar algo así es difícil, pero factible. Más específicamente, el y contenido de incitación al odio y el compensatorio podrían clasificarse utilizando el procesamiento del lenguaje natural en varias categorías más pequeñas: la incitación a la línea de acción; debates genuinos, insultos, etc. Pensamos que alcanzar algo así es difícil, pero factible.

Una cuestión pendiente es el grado en que un debate en Facebook evoluciona o se resuelve: ¿supone la publicación de contenido con incitación al odio un elemento que afecta al resto del debate? Pensamos que alcanzar algo así es harto difícil con técnicas automáticas, pero plausible si se realizara de forma manual por análisis humano.

Métricas del mundo real

Este término se refiere a la cuestión del efecto que el contenido de incitación al odio ejerce sobre las actitudes o comportamientos a largo plazo en vivo y sin Internet. Esto es extremadamente difícil de calcular con cualquier grado de precisión - al igual que es difícil determinar si la incitación al odio tiene un efecto sobre las actitudes o el comportamiento. Para obtener indicadores sensibles y defendibles en este sentido sería necesario seguir investigando.

Este estudio exploratorio también ha identificado algunas áreas en las que una cantidad relativamente pequeña de una mayores recursos para investigación podría mejorar nuestra comprensión de la materia.

En primer lugar, es importante crear una mejor forma de medir el volumen total de contenido populista y compensatorio. Resulta extremadamente difícil producir cifras globales fiables en diferentes tipos de contenidos compensatorios sin un método riguroso y científico para determinar objetivamente las páginas identificadas. Un método de identificación especial entraña un riesgo importante de producir datos engañosos. Un enfoque automatizado basado en las páginas más activas y las que más gustaron a los usuarios en diferentes países y categorías proporcionaría una medida mucho más sólida y fiable.

En segundo lugar, hay nuevos tipos de contenidos que pueden ser estudiados con más detalle. Esto incluiría el alcance total de diferentes tipos de contenido (en lugar de sólo las interacciones), el análisis de las secuencias de conversaciones que aparecen a raíz de cada mensaje y evaluar cómo cambia, la existencia de determinados grupos o usuarios que tienen más éxito que otros en la resolución de debates o en la lucha contra la incitación al odio y la violencia y por último, la evaluación de mensajes compartidos en el feed individual en oposición a la propia página

En tercer lugar, sería útil para entender cómo Facebook en calidad de red de usuarios, debería informar a los usuarios de nuestro método, puesto que alberga tanto páginas populistas como de contenido compensatorio. Esto implicaría examinar los diferentes métodos de divulgar contenido. Por ejemplo, mediante el seguimiento de un segmento de contenido específico podemos aplicar el análisis de redes para ver quién lo ha compartido, dónde se originó y hacia dónde se dirige. Esto podría ayudar a responder algunas preguntas fundamentales sobre de cómo este tipo de contenidos se divulgan y qué caminos toman. ¿Tienen diferentes contenidos diferentes formas de viajar? ¿Cómo funciona este aspecto en los diferentes países objeto de estudio?

En cuarto lugar, una mayor investigación podría ayudar a determinar lo que podría constituir una medida sensata y sólida de cuantificar la efectividad de un sitio fuera de Internet. Con el fin de proporcionar una visión más práctica y detallada, se podrían solicitar una serie de estudios de casos detallados de las campañas en línea que han tenido un gran impacto en el mundo real.

NOTES

¹ Un comentario es un término "comodín" para cualquier contenido que se crea o se comparte en una página y que añade información sobre una situación sin necesariamente hacer referencia a fuentes externas.

² Un ataque es un tipo de contenido particularmente agresivo dirigido a un usuario o a un grupo de individuos específicos. Un ataque puede ser dirigido a un grupo étnico, político o a una organización.

³ Un contenido festivo es aquel que celebra una página o sus valores.

⁴ Un contenido que manifiesta enfado puede utilizar lenguaje soez y abogar por medidas extremas en relación con el contenido.

⁵ Se refiere a todas las bromas y contenido satírico.

⁶ Cualquier esfuerzo por mantener un debate serio sobre temas específicos relacionados con la incitación a la violencia o al odio (xenofobia, inmigración, etc.)

⁷ Cualquier esfuerzo por mantener un debate serio sobre política, noticias o temas de interés general

⁸ Un contenido compensatorio no constructivo consiste en un texto que desafía la incitación al odio o la violencia de una manera no constructiva (es decir, ataca a un individuo, tiene un carácter ofensivo etc). Corroboración de datos consiste en cuestionar o comprobar una afirmación.

Demos – Licence to Publish

The work (as defined below) is provided under the terms of this licence ('licence'). The work is protected by copyright and/or other applicable law. Any use of the work other than as authorized under this licence is prohibited. By exercising any rights to the work provided here, you accept and agree to be bound by the terms of this licence. Demos grants you the rights contained here in consideration of your acceptance of such terms and conditions.

1 Definitions

- a 'Collective Work' means a work, such as a periodical issue, anthology or encyclopedia, in which the Work in its entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole. A work that constitutes a Collective Work will not be considered a Derivative Work (as defined below) for the purposes of this Licence.
- b 'Derivative Work' means a work based upon the Work or upon the Work and other pre-existing works, such as a musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Work may be recast, transformed, or adapted, except that a work that constitutes a Collective Work or a translation from English into another language will not be considered a Derivative Work for the purpose of this Licence.
- c 'Licensor' means the individual or entity that offers the Work under the terms of this Licence.
- d 'Original Author' means the individual or entity who created the Work.
- e 'Work' means the copyrightable work of authorship offered under the terms of this Licence.
- f 'You' means an individual or entity exercising rights under this Licence who has not previously violated the terms of this Licence with respect to the Work, or who has received express permission from Demos to exercise rights under this Licence despite a previous violation.

2 Fair Use Rights

Nothing in this licence is intended to reduce, limit, or restrict any rights arising from fair use, first sale or other limitations on the exclusive rights of the copyright owner under copyright law or other applicable laws.

3 Licence Grant

Subject to the terms and conditions of this Licence, Licensor hereby grants You a worldwide, royalty-free, non-exclusive, perpetual (for the duration of the applicable copyright) licence to exercise the rights in the Work as stated below:

- a to reproduce the Work, to incorporate the Work into one or more Collective Works, and to reproduce the Work as incorporated in the Collective Works;
- b to distribute copies or phonorecords of, display publicly, perform publicly, and perform publicly by means of a digital audio transmission the Work including as incorporated in Collective Works; The above rights may be exercised in all media and formats whether now known or hereafter devised. The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. All rights not expressly granted by Licensor are hereby reserved.

4 Restrictions

The licence granted in Section 3 above is expressly made subject to and limited by the following restrictions:

- a You may distribute, publicly display, publicly perform, or publicly digitally perform the Work only under the terms of this Licence, and You must include a copy of, or the Uniform Resource Identifier for, this Licence with every copy or phonorecord of the Work You distribute, publicly display, publicly perform, or publicly digitally perform. You may not offer or impose any terms on the Work that alter or restrict the terms of this Licence or the recipients' exercise of the rights granted hereunder. You may not sublicense the Work. You must keep intact all notices that refer to this Licence and to the disclaimer of warranties. You may not distribute, publicly display, publicly perform, or publicly digitally perform the Work with any technological measures that control access or use of the Work in a manner inconsistent with the terms of this Licence Agreement. The above applies to the Work as incorporated in a Collective Work, but this does not require the Collective Work apart from the Work itself to be made subject to the terms of this Licence. If You create a Collective Work, upon notice from any Licensor You must, to the extent practicable, remove from the Collective Work any reference to such Licensor or the Original Author, as requested.
- b You may not exercise any of the rights granted to You in Section 3 above in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation. The exchange of the Work for other copyrighted works by means of digital filesharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in connection with the exchange of copyrighted works.

C If you distribute, publicly display, publicly perform, or publicly digitally perform the Work or any Collective Works, You must keep intact all copyright notices for the Work and give the Original Author credit reasonable to the medium or means You are utilizing by conveying the name (or pseudonym if applicable) of the Original Author if supplied; the title of the Work if supplied. Such credit may be implemented in any reasonable manner; provided, however, that in the case of a Collective Work, at a minimum such credit will appear where any other comparable authorship credit appears and in a manner at least as prominent as such other comparable authorship credit.

5 Representations, Warranties and Disclaimer

A By offering the Work for public release under this Licence, Licensor represents and warrants that, to the best of Licensor's knowledge after reasonable inquiry:

- i Licensor has secured all rights in the Work necessary to grant the licence rights hereunder and to permit the lawful exercise of the rights granted hereunder without You having any obligation to pay any royalties, compulsory licence fees, residuals or any other payments;
- ii The Work does not infringe the copyright, trademark, publicity rights, common law rights or any other right of any third party or constitute defamation, invasion of privacy or other tortious injury to any third party.

B except as expressly stated in this licence or otherwise agreed in writing or required by applicable law, the work is licensed on an 'as is' basis, without warranties of any kind, either express or implied including, without limitation, any warranties regarding the contents or accuracy of the work.

6 Limitation on Liability

Except to the extent required by applicable law, and except for damages arising from liability to a third party resulting from breach of the warranties in section 5, in no event will licensor be liable to you on any legal theory for any special, incidental, consequential, punitive or exemplary damages arising out of this licence or the use of the work, even if licensor has been advised of the possibility of such damages.

7 Termination

A This Licence and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this Licence. Individuals or entities who have received Collective Works from You under this Licence, however, will not have their licences terminated provided such individuals or entities remain in full compliance with those licences. Sections 1, 2, 5, 6, 7, and 8 will survive any termination of this Licence.

B Subject to the above terms and conditions, the licence granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different licence terms or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this Licence (or any other licence that has been, or is required to be, granted under the terms of this Licence), and this Licence will continue in full force and effect unless terminated as stated above.

8 Miscellaneous

A Each time You distribute or publicly digitally perform the Work or a Collective Work, Demos offers to the recipient a licence to the Work on the same terms and conditions as the licence granted to You under this Licence.

B If any provision of this Licence is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this Licence, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

C No term or provision of this Licence shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.

D This Licence constitutes the entire agreement between the parties with respect to the Work licensed here. There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You. This Licence may not be modified without the mutual written agreement of Demos and You.

Contenido compensatorio -aquél que argumenta, discrepa o presenta un punto de vista opuesto- hace referencia a una estrategia potencialmente importante de hacer frente al contenido extremista u ofensivo en línea. Las principales ventajas del contenido compensatorio: resulta más rápido, flexible y reactivo; puede surgir de cualquier parte del mundo en cualquier idioma y conserva el principio de respeto por los espacios de debate públicos y accesibles. Sin embargo, el contenido compensatorio puede adoptar tantas formas como el extremismo y, además, es probable que no siempre resulte tan eficaz como podría ser. Algunos tipos de contenido compensatorio podrían incluso resultar contraproducentes.

Este breve informe provisional resume las conclusiones de la fase I en la que se analiza cómo se crea y comparte el contenido que desafía páginas populistas de derecha en toda Europa. Se basa en el análisis de la actividad de 150 páginas de Facebook durante un período de dos meses. El objetivo consiste en evaluar cómo los usuarios comparte e interactúan con el contenido producido en dichas páginas. También presenta nuevas metodologías y enfoques para medir cómo el contenido se propaga en línea, cómo podría medirse su impacto y cómo mejorar su eficacia.

Este es el primer informe de una serie de investigaciones que examinan contenido compensatorio o que desafía contenido extremista en línea. Los próximos informes examinarán la ideología islamista en Reino Unido entre otros países.

Jamie Bartlett es Director del Centro para el Análisis de los Medios de Comunicación Social de Demos. Alex Krasodonski-Jones es investigador en el Centro para el Análisis de los Medios de Comunicación Social de Demos.

ISBN 978-1-909037-93-9