



## Esfera pública digital y Ciberacoso, una propuesta metodológica para su análisis en México

### Cyberbullying, a methodological proposal for its study in Mexico

**Mtra. Carla Irene Ríos Calleja**

Benemérita Universidad Autónoma de Puebla. México.

[carla.rioscalleja@viep.com.mx](mailto:carla.rioscalleja@viep.com.mx)

<https://orcid.org/0000-0001-9656-4238>

**Dra. Araceli Espinosa Márquez**

Benemérita Universidad Autónoma de Puebla. México.

[araceli.espinosa@correo.buap.mx](mailto:araceli.espinosa@correo.buap.mx)

<https://orcid.org/0000-0003-0475-6932>

**Lic. Edwin Garcilazo-Arriaga**

Benemérita Universidad Autónoma de Puebla. México.

[edwin.garcilazoarriaga@viep.com.mx](mailto:edwin.garcilazoarriaga@viep.com.mx)

<https://orcid.org/0000-0002-0199-9587>

Recepción: 10/01/2023 Revisión: 11/01/2023  
Aceptación: 17/01/2023 Publicación: 18/01/2023

#### Resumen

El advenimiento de la pandemia causada por COVID SARS 19 implicó que en México muchas de las actividades diarias se desarrollan en el espacio digital. Las actividades escolares y laborales en su mayoría implicaron que las personas realizaran sus actividades sin una capacitación para el desarrollo de habilidades necesarias en entornos digitales. De ahí que también se incrementaran las conductas delictivas en dichos entornos. En este contexto, el ciberacoso fue una conducta recurrente, sobre todo debido a que muchos niños, niñas y adolescentes incrementaron dramáticamente las horas frente al ordenador. El propósito de este documento es analizar este fenómeno a partir de analizar los datos web provenientes de las búsquedas en Google a partir de un libro de códigos relacionados con el ciberacoso se examina la correlación con los datos de las encuestas del Módulo Sobre Ciberacoso (MOCIBA), del Instituto Nacional de Estadística y Geografía (INEGI) correspondiente a los años 2019, 2020 y 2021. La propuesta involucra esta encuesta recogida a nivel nacional, reconociendo que las encuestas enfrentan nuevos desafíos con el advenimiento de la vida digital al presentarse nuevas estructuras de comunicación de los sujetos que actualmente se auto comunican y *auto informar* en el espacio digital. De esta manera se propone una metodología diseñada expresamente para explorar las coincidencias entre los instrumentos convencionales de medición y datos web no estructurados para medir la incidencia de ciberacoso en México a partir de una metodología híbrida que aprovecha ambas herramientas para entender mejor algunos problemas de la esfera pública digital, como el fenómeno del ciberacoso.

**Palabras clave:** Ciberacoso; Esfera pública digital; Encuestas; Google Trends; Metodología cuantitativa



## Abstract

The advent of the pandemic caused by COVID SARS 19 meant that in Mexico many of the daily activities are carried out in the digital space. School and work activities for the most part meant that people carried out their activities without training to develop the necessary skills in digital environments. Hence, criminal behaviour in such environments also increased. In this context, cyberbullying was a recurrent behaviour, especially as many children and adolescents dramatically increased the number of hours spent in front of the computer. The purpose of this paper is to analyse this phenomenon by analysing web data from Google searches from a codebook related to cyberbullying and examining the correlation with the data from the surveys of the Module on Cyberbullying (MOCIBA) of the National Institute of Statistics and Geography (INEGI) for the years 2019, 2020 and 2021. The proposal involves this survey collected at the national level, recognising that surveys face new challenges with the advent of digital life as new communication structures of the subjects who currently self-communicate and self-report in the digital space are presented. In this way, we propose a methodology specifically designed to explore the overlaps between conventional measurement instruments and unstructured web data to measure the incidence of cyberbullying in Mexico through a hybrid methodology that takes advantage of both tools to better understand some problems of the digital public sphere, such as the phenomenon of cyberbullying.

**Keywords:** Cyberbullying; Digital Public Sphere; Surveys; Google Trends; Quantitative Methodology

## 1. Introducción

El término esfera pública ha sido ampliamente estudiado y definido desde distintos enfoques. En este trabajo, el enfoque utilizado es el sociológico, como lo sugiere Bernard Enjolras y Kari Steen-Johnsen (2017) para no preguntarse cuáles deberían ser las cualidades, elementos y actores ideales de la esfera pública, sino para observar a la opinión pública como un fenómeno social, próximo a una realidad común. Bajo esta visión, los mismos autores ofrecen una definición inicial del concepto como: un espacio que comprende a espacios comunicativos institucionales que facilitan la discusión pública y la formación de opinión pública (2017, p. 101) a partir de un flujo de información, para referirse a que no se puede argumentar la existencia de una sola esfera pública, unificada, sino de una multiplicidad de espacios.

Vale rastrear la evolución del término para identificar que muchos elementos se han transformado, como la estructura de los espacios comunicativos y los elementos instrumentales que acompañan la discusión y la formación de opinión pública, sobre todo lo referente a los canales de *información* que hoy están disponibles.

Los planteamientos de algunos autores como: Lippmann, Dewey, Habermas y Arendt resultan seminales en la construcción y discusión del concepto de esfera pública. Se han tomado como punto de partida conceptual porque cada uno ofrece elementos sustanciales que hoy pueden agregarse para ofrecer una perspectiva actualizada y justificar la existencia de una esfera pública digital.

Por un lado, Walter Lippmann, parte de tres supuestos para construir el concepto de esfera pública, el primero es que el sujeto coexiste en un ambiente con características estructurales propias y complejas. El segundo, que el sujeto es un ser con capacidades cognitivas limitadas y, por lo tanto, requiere de atajos mentales para reducir dicha complejidad y poder actuar en dicho ambiente. Y, el tercer elemento, se refiere al concepto de estereotipo, que corresponde a un conjunto de generalidades o atajos mentales que son asociados con algo conocido (Livio, 2013). En este sentido, de acuerdo con Lippmann podemos decir que, “lo que vemos”, depende de esta interpretación que hacemos de acuerdo con nuestros modelos mentales. Por tal motivo, siguiendo a Lippmann, la presencia de *expertos* está más que justificada, pues ellos se



encargarán de reducir la complejidad, filtrar y despejar los estereotipos para asesorar - a los dirigentes políticos- en la toma de decisiones con sentido más racional.

Por otro lado, John Dewey, quien criticó el planteamiento de Lippmann, reconoce que muchos elementos son expuestos con claridad, pero señala la figura de los expertos. Si bien está de acuerdo en que ayudan a reducir la complejidad del entorno para hacerla asequible al entendimiento de otras personas, no está de acuerdo en que los únicos que reciban esta claridad sean los dirigentes políticos. En su caso, reconoce que esta actividad es -o debería- ser desempeñada por los medios de comunicación (Livio, 2013). Sin embargo, atisba dos desenlaces posibles en relación con la función de estos expertos. El primero corresponde al riesgo de que estos expertos puedan estar sirviendo a los intereses de los dirigentes políticos -y la clase dominante- para buscar legitimidad y, el segundo, que se conviertan en paladines del interés general aproximándose al público mismo ofreciéndole su conocimiento para fomentar la toma de decisiones más racionales.

Otro argumento de Dewey, y más importante aún, es la existencia de “públicos”, entendidos como el conjunto de personas que resultan agraviadas, directa o indirectamente, por la acción social en general y, éste -el público-, será más efectivo en la medida que se encuentre mejor *informado* (Calhoun, 2017) y, en sociedades complejas, existirán tantos públicos como problemas existan y, en esa misma medida, habrá expertos dispuestos a ofrecer *información*.

Por otro lado, Habermas construye su argumento a partir de la pregunta: ¿en qué medida los actores pueden formarse opiniones que guíen la acción pública sobre la base de un discurso racional? (Calhoun, 2017, p. 26). Él reconoce el uso de la razón para comunicar argumentos en un *espacio* de deliberación enfocado a una noción de *bien público* (Jaramillo, 2010). Pero surge un problema para la eficacia de este planteamiento, que la razón depende de dos elementos, el primero, del acceso al discurso, es decir, a una tribuna y, el segundo, a la apertura (citar).

Podemos entender a la apertura que apunta Habermas como el acceso a *información* de carácter público, por lo menos la mínima, para formar un criterio y poder deliberar. En este sentido, el papel de la generación de información recae en los expertos -para Lippmann- o de los medios de comunicación -para Dewey-. Y es precisamente en los medios de comunicación que Habermas veía una amenaza para la “calidad de la información” si estos caían en el control de los estados o de los mercados para convertirse en plataformas de publicidad (Calhoun, 2017).

Continuando con las contribuciones del concepto de esfera pública, Hannah Arendt escribió sobre éste de forma indirecta, puesto que nunca utilizó el término como concepto teórico. Sin embargo, ella conceptualizó al espacio público a partir de dos elementos, el primero, como un espacio de aparición, es decir, donde existe la posibilidad de ser visto y oído por otros y, el segundo, un lugar común para todos, diferenciado de lo privado (Calhoun, 2017). Contrario a la visión habermasiana de una esfera pública homogénea, Arendt la entiende como el conjunto de pequeños espacios locales, donde la persona buscará el lugar para ser visto y oído por otros y compartir sus puntos de vista. En esta lógica, el individuo también será recíproco en esta interacción con otros y tendrá acceso a múltiples opiniones individuales, y esto último también constituirá uno de sus canales de acceso a otro tipo de *información* que apoyará la formación de sus propias opiniones. Hoy, el Internet y las redes sociales permiten unirse a un flujo de información, en pequeños espacios, que apoyará la formación de la opinión propia (Alexey, 2018).

A modo de recapitulación y para destacar los elementos conceptuales de este trabajo, todos los autores presentados reconocen a un individuo socialmente relevante, que interactúa con un entorno y es capaz de construir una interpretación de la realidad. Y,

esta realidad, es producto de una interpretación y de los esquemas mentales del individuo, los cuales están sujetos a la *información* que reciban del entorno. Vale decir que hoy el entorno no se limita a un espacio físico ni supone que la interacción sea cara-cara, sino un espacio mediado por la tecnología que ha permitido una interactividad virtual (Vázquez, 2013).

Hoy, podría decirse, existe una esfera pública ampliada y el Internet como medio de comunicación masivo y la ubicuidad de su acceso configuran otro espacio (Castrelo, 2018) al cual accede el individuo de forma autónoma. Sin ser objeto de este trabajo discutir el origen y los actores que participan en la creación de *información*, no se puede negar que ésta es un elemento sustancial para la participación del individuo en la esfera pública y en su formación de opinión. Sin embargo, a diferencia del entorno descrito décadas atrás, hoy el individuo tiene acceso autónomo a esta *información* y por múltiples vías.

Ante el avance de la tecnología, se sostiene que ésta ha creado nuevos mecanismos de autocomunicación y *autoinformación* que amplía el rango de fuentes configuradoras de la esfera pública (Pareja Sánchez y Echeverría, 2014, p. 1). Ante la existencia y accesibilidad a estos mecanismos de autocomunicación y *autoinformación* que moldean el espacio público, se plantea: ¿de qué forma se pueden medir estos flujos de *información*, mediada por la tecnología e impulsada por el individuo, para conocer y estudiar más elementos de la esfera pública?

Tradicionalmente, para analizar problemas del espacio público, se emplean las encuestas como un instrumento para estudiar, conocer y describir los intereses y opiniones de un grupo de personas que representa, estadísticamente, a una población para realizar inferencias y aproximaciones sobre la esfera pública en general. No se puede negar que estos instrumentos son útiles por su diseño, pero inclusive aún, enfrentan algunos problemas como la cobertura, la baja disposición de las personas a responder y otros errores derivados de la naturaleza de aplicación (Groves y Lyberg, 2010) y queda la pregunta abierta de si las encuestas tendrán aún un papel dominante ante los cambios de hábitos de comunicación y de flujo de información (Reveillac et al., 2022).

Antes de continuar, es pertinente definir conceptualmente el ciberacoso y ponerlo en dimensión para en caso de México. Éste se entiende como el daño intencional y repetido infligido a través de dispositivos electrónicos (Patchin & Hinduja, 2015, p. 2). Y se puede argumentar que corresponde un problema de la esfera pública digital dado que, en 2021, en México, el 21.7% de la población de 12 años o más que uso internet, sufrió una situación de ciberacoso (INEGI, 2022a). De acuerdo con el U-Report de Unicef, en México, el 93% de los encuestados entiende el significado del término ciberacoso y, de esta misma muestra, el 47% declaró haber sido víctima de ciberacoso de los cuales, el segmento de edad con más incidencia (83%) fue de los 31 a los 34 años de edad, siendo el tipo sexual el más común, con 44% (U Report México, 2020).

Dado que las fuentes de información que serán empleadas parten de la condición de la disponibilidad y el acceso a Internet, es preciso poner en contexto la realidad de la población la penetración del Internet en la población mexicana y sus hábitos de consumo. De acuerdo con el INEGI, en 2021 el 75.6% de la población de seis años o más tiene acceso a internet (INEGI, 2022b). De acuerdo con la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares, en 2020, el 60.6% de los hogares en México contó con acceso a internet y el 96.0% de las personas utilizó un smartphone como equipo más frecuente para conectarse a internet y las tres principales actividades fueron: comunicarse (93.8%), buscar información (91%) y acceder a redes sociales (89%) (INEGI, 2020).

En específico, sobre el uso y el acceso al internet móvil, en México, la penetración de la red 3G y 4G es del 87.26% (García Zaballos et al., 2021) lo que supone poco más de 97 millones de usuarios que potencialmente tienen acceso a internet y, de los cuales, 88.5 millones lo emplean para buscar información. De acuerdo con la composición del mercado de teléfonos celulares, el 90% de los dispositivos funciona sobre un sistema operativo Android, el cual es propiedad de Google, lo que podría suponer un acceso casi inmediato a su buscador, ante cualquier indicio del usuario para conseguir información (Insights, 2019). El teléfono celular se ha convertido en el principal dispositivo para acceder a internet, inclusive, dos veces más que la computadora, cuyo uso ha disminuido un 7.8% del 2017 al 2021 (INEGI, 2022b). En México, el sitio más visitado es Google.com, con 2.6 billones de visitas mensuales, de las cuales 108 millones son únicas y consultan en promedio 8 páginas por visita (Similarweb, 2022).

Si bien, la penetración de la telefonía en México (88.5%) es relativamente menor frente a otros países de América Latina (García Zaballos et al., 2021) resulta alentador la cifra de crecimiento de usuarios de internet en zonas rurales que pasó de 39.2% en 2017 a 50.4% en 2020 (INEGI, 2020).

En este sentido, ante la ampliación de la esfera pública a un espacio donde es posible la autocomunicación y *autoinformación*, ¿serán equivalentes los resultados de las encuestas tradicionales y los datos provenientes de los hábitos de uso de la web? En específico se plantea la pregunta: en México, para 2019, 2020 y 2021, ¿existe alguna relación estadística entre los resultados obtenidos en las encuestas del Módulo Sobre Ciberacoso (MOCIBA) y los patrones de búsqueda registrados por usuarios de internet? Se apunta como respuesta provisional la ausencia de diferencia relativa entre los resultados obtenidos en las encuestas y los patrones de búsqueda en internet.

Para comprobar la suposición planteada, se emplearon dos fuentes de información que fueron comparadas mutuamente, para un mismo segmento de tiempo y población. La primera, corresponde al método habitual de encuesta estructurada y se emplearon las encuestas del Módulo Sobre Ciberacoso (MOCIBA), del Instituto Nacional de Estadística y Geografía (INEGI), en específico, el número de personas que reportaron haber sido víctimas de ciber acoso. Por otro lado, se tomó un *proxy* en los macrodatos de Google Trends, sobre las búsquedas de ciertos términos relacionados al ciberacoso para el periodo 2019, 2020 y 2021.

## 2. Marco teórico - conceptual

Se propone una metodología híbrida como planteamiento para entender de forma más amplia a la esfera pública y sus componentes, sobre todo en la época actual donde la tecnología de la información se ha democratizado y donde las personas tienen acceso a diferentes arenas o esferas de las que pueden obtener y aportar información de forma más autónoma.

Las tecnologías de la comunicación impactan socialmente las instituciones de comunicación y los valores al estructurar el tiempo y el espacio (Enjolras & Steen-Johnsen, 2017, p. 102). Argumentando la existencia de una esfera pública digitalizada, se puede decir que constituye un espacio social donde se desarrollan nuevas formas de relaciones de poder y de comunicación, bajo una influencia digital que no se basa en recursos tradicionales sino en el capital social de quien comunica y en su habilidad para hacerlo (Enjolras & Steen-Johnsen, 2017, p. 111). Hoy es posible obtener la magnitud y el flujo de información a partir de las búsquedas hechas por los usuarios de internet en los buscadores, obtenidas de forma pública y organizados como macrodatos.

Por macrodatos se entiende a aquellos datos que, por su origen y dimensión, resultan complejos y difíciles de procesar por métodos convencionales. Aunque no es un término



nuevo, una de las fuentes generadoras de este tipo de datos ha sido el Internet, a partir de los hábitos de uso que reportan sus usuarios. Dentro de esta categoría de datos, se encuentran los Datos de las Redes Sociales (DRS) también conocido como SMD (por sus siglas en inglés *Social Media Data*) que corresponden a un tipo de datos que provienen específicamente de las redes sociales (Reveilhac et al., 2022).

Pero cabe mencionar una diferencia sutil entre datos provenientes de las redes sociales y los datos provenientes de las búsquedas que los individuos hacen en internet a través de un buscador, en este caso Google. Aunque su estructura es similar, su génesis no es equivalente. Por un lado, en su atributo técnico, los DRS se originan a partir del registro del comportamiento de un usuario o perfil dentro de una red social y, desde su atributo social, dichos datos reflejan un comportamiento que es definido por la estructura de la red además de la interacción con otros usuarios, es decir, este comportamiento está sujeto a una interacción.

Por otro lado, los *datos web* que provienen del buscador se originan por usuarios que realizan una búsqueda de forma privada, es decir, sin miedo ni pudor los usuarios tienen la posibilidad de desnudar sus intereses y emociones, lo que permite suponer que el dato en sí mismo aporta datos fidedignos de un fenómeno. Hay que resaltar en este punto, que el propósito del presente trabajo no es cuestionarse el origen de esta motivación ni tampoco las repercusiones en el individuo por la información que es presentada por dicho buscador y, al respecto, cabe mencionar que tampoco es la finalidad del presente cuestionarse el origen de esta información. Tan solo trata de capturar el interés de un individuo, con acceso a internet, sobre un tema particular.

Para poner en contexto los instrumentos de análisis que se emplearon hay que mencionar que las encuestas del MOCIBA están construidas a partir de una muestra de representación nacional, para una población de 12 y más años que usan internet en cualquier dispositivo y es residente de una vivienda particular, siendo esta la unidad de muestreo. La encuesta del 2019 se levantó del 01 de julio al 23 de agosto, para el 2020 se levantó del 05 de octubre al 27 de noviembre y para el 2021 se levantó del 2 de agosto al 30 de septiembre.

Dada la masividad y la cobertura de usuarios del buscador de Google, se puede argumentar su uso como una fuente válida de datos y existen algunos trabajos que los han empleado para pronosticar movimientos en el mercado de valores (Fong, 2017), para pronosticar tasas de desempleo en Estados Unidos (Nagao et al., 2019) y para predecir la participación electoral en elecciones presidenciales (Prado-Román et al., 2021).

Si bien existe un debate sobre la utilidad y validez de los datos no estructurados de búsquedas de internet frente a instrumentos convencionales para estudiar elementos de la opinión pública, como las encuestas (Schober et al., 2016), existen también argumentos que reconocen algunas ventajas de emplear ambas fuentes, como el trabajo de Reveilhac M., Steinmetz S. y Morselli D. (2022) que encuentra que este tipo de datos puede ofrecer un contexto más amplio para interpretar la opinión pública y mejorar la calidad y utilidad de una encuesta, además de ofrecer ventajas sobre las encuestas, la primera, que ofrece acceso a temas que son difíciles de encuestar como, por ejemplo, la violencia o el racismo y, la segunda, que ofrece la posibilidad de obtener fases iterativas de análisis a lo largo del tiempo.

Respecto a la primera ventaja, algunos estudios han reconocido la poca probabilidad a que algún tema sufra censura social, puesto que los usuarios de buscadores en línea, usualmente acceden en privacidad, lo que permite acceder a temas tabú (Kreuter et al., 2019 en Stephens-Davidowitz, 2014). Adicional a esto, las personas se comportan inusualmente comunicativas en Google, al realizar búsquedas (Stephens-Davidowitz,

2019) lo que permite una riqueza de información aún mayor, no mediada ni acotada por otros factores.

Aprovechando estas ventajas Stephens-Davidowit (2014) encuentra, a partir de datos de Google Trends, que en las elecciones presidenciales de Estados Unidos donde participo Barack Obama, 2008 y 2012, la tasa de búsquedas con carga racial, en un área determinada, son un predictor sólido del apoyo que recibió dicho candidato. Encuentra que esta animosidad racial le costó cuatro puntos porcentuales del voto popular nacional en ambas elecciones. Uno de los aportes más relevantes de este trabajo fue que puso en discusión el impacto del prejuicio racial en las elecciones de ese país, cuando en las encuestas no se había notado.

En este orden de ideas y retomando la pregunta que guía este trabajo se plantea lo siguiente: en México, para 2019, 2020 y 2021, ¿existe alguna relación estadística entre las personas que declararon ser víctimas de ciberacoso en la encuesta del MOCIBA y los patrones de búsqueda de términos relacionados al ciberacoso?

### 3. Metodología y resultados

Para probar el planteamiento se ha tomado como punto de partida los resultados presentados en el Módulo sobre Ciberacoso (MOCIBA) del INEGI, que tiene como propósito identificar las medidas de seguridad que la población implementa para proteger cuentas y equipo, conocer el número de personas que fueron víctimas de ciberacoso en los últimos 12 meses, el tipo de acoso, la percepción sobre la importancia de proteger su información y conocer qué medida la población protege sus datos.

En la encuesta 2019 y 2020, se les pregunta si sufrieron alguna situación de acoso, ofreciendo 10 opciones y, para 2021, agregan dos opciones más a la encuesta, para tener un total de 12 tipos de situación de acoso para este año (ver tabla 1). Esta ampliación corresponde a un desglose de la situación de *acoso publicación de información personal*, para explorar el tipo y las circunstancias de este tipo de ciberacoso. En cada uno se les pregunta por la identidad de quien perpetró el acoso, el sexo del acosador, la edad aproximada, la razón por la que creen que lo hizo, la reacción o efecto que provocó sobre la víctima y el canal digital por el cual sucedió el acoso.

Tabla 1

**POBLACIÓN DE 12 AÑOS Y MÁS QUE VIVIÓ CIBERACOSO DURANTE LOS ÚLTIMOS DOCE MESES, POR SITUACIONES DE CIBERACOSO EXPERIMENTADAS. MÉXICO, 2019, 2020 Y 2021. Porcentaje del total**

Tipo de situación	2019	2020	2021
Contacto mediante identidades falsas	33.56	35.03	35.69
Mensajes ofensivos	33.47	34.62	33.12
Insinuaciones o propuestas sexuales	29.03	26.67	24.56
Recibir contenido sexual	26.55	24.42	25.48
Provocaciones para reaccionar de forma negativa	23.89	20.70	20.79
Llamadas ofensivas	23.10	21.91	19.33
Rastreo de cuentas o sitios web	17.59	16.46	16.80
Suplantación de identidad	17.23	17.45	18.09
Críticas por apariencia o clase social	13.72	15.58	16.56
Publicación de información personal	12.72	11.98	
Amenazar con publicar información personal, audios o video para extorsionar			7.57
Publicar información personal, fotos o videos			5.23



Publicar o vender imágenes o videos de contenido sexual	3.07
Otra situación	2.04

Elaboración del autor con datos del Módulo sobre Ciberacoso (MOCIBA), 2019, 2020 y 2021. INEGI.

De acuerdo con los resultados del MOCIBA, en 2019, 17,670,579 personas declararon haber sido víctimas de ciberacoso, en al menos una de sus modalidades, en los últimos 12 meses. Para 2020 se registró una disminución del 8.7%, colocándose en 16,132,831 el número de víctimas y, para el 2021, se registró un incremento del 9.4% para colocarse en niveles previos, en 17,655,776 víctimas.

A partir de la metodología propuesta por Stephens-Davidowitz (2014), una vez obtenidos las situaciones de acoso que capturan la encuesta del MOCIBA se construyó una matriz con términos relacionados a estas situaciones y se convirtieron términos de búsquedas en Google Trends, para un año completo, del 01 de enero al 31 de diciembre, para cada una respectivamente (ver tabla 2). Se construyó una matriz amplia de términos para disminuir el ruido y aumentar la posibilidad de que la búsqueda de estos términos corresponda a una víctima.

La matriz de términos de búsqueda se construyó en función de los términos contemplados en las situaciones de acoso de la encuesta mencionada (ver anexo). El término más importante fue: “ley Olimpia”, el cual no se refiere propiamente a una ley sino a conjunto de reformas legislativas que tienen el propósito de reconocer la violencia digital, también conocida como ciber violencia y sancionar delitos que violen la intimidad sexual de las personas (Unidad General de Asuntos Jurídicos, 2022). El antecedente de este marco regulatorio está en un caso emblemático de una joven poblana, Olimpia Melo Cruz, que fue violentada por su expareja sentimental al publicar contenido íntimo (López Téllez et al., 2022). Aunque la petición de reforma fue presentada en 2014, fue hasta 2018 que se aprobaron las reformas en el Código Penal del Estado Libre y Soberano de Puebla y luego esta conducta fue tipificada en otros estados de la república mexicana como un delito. Este concepto se consideró como un concepto de búsqueda dado que engloba un significado relacionado a la vulneración de la privacidad sexual en México. El término “ciberacoso” se incluyó en la búsqueda por corresponder textualmente al objeto de la encuesta del MOCIBA. Por otra parte, se incluyeron otros términos que pretenden simular una búsqueda sencilla por parte de una víctima como: intentar hacer una “denuncia en línea”, contactar a la “policía cibernética”, obtener orientación sobre “hackearon mi cuenta” y obtener información sobre una “extorsión”.

Tabla 2

### MATRIZ DE TÉRMINOS DE BÚSQUEDA EMPLEADOS EN GOOGLE TRENDS. MÉXICO, POR SUBREGIONES<sup>1</sup>, 2019, 2020 Y 2021

#### Término de búsqueda

Ley Olimpia<sup>2</sup>

Ciberacoso

Denuncia en línea

Policía cibernética

Hackearon mi cuenta

Extorsión





Elaboración del autor

<sup>1</sup> Incluye las 32 entidades federativas

<sup>2</sup> Se incluyen las siguientes variantes: "la ley olimpia", "la ley de olimpia" y "ley de olimpia".

De acuerdo con la arquitectura de búsquedas de Google Trends, se pueden consultar las palabras por "término" o por "tema", la diferencia estriba en que en el primero, los resultados se limitan a considerar a la palabra como un término de búsqueda. En el segundo, por "tema", la búsqueda se amplía a un grupo de términos que comparten el concepto en el idioma consultado. Para los fines del presente trabajo se empleó una búsqueda por "término", sin restricciones de exactitud, es decir, que se consultaron las búsquedas que contienen los términos en cualquier forma, pero que se mostraron juntos como concepto. Esto se hizo por dos razones, la primera, capturar cualquier intención de búsqueda que estuviera relacionada con esos términos, en todas las categorías de búsqueda web y, la segunda, porque la relación de conceptos por "tema" que realiza Google no se conoce con exactitud para todos los términos. La forma en que los resultados son presentados es en función de la popularidad de búsqueda relativa que registraron los conceptos para un periodo dado. Por ejemplo, si se hace una consulta de los términos "A" y "B", la búsqueda completa de "A" + "B" representará el 100% mientras que en lo individual se especificará qué porcentaje del total representa cada término. Si sólo se realiza la búsqueda del "A", Google establece un "B" genérico, siendo el mismo para cada consulta, lo que permite comparar todos los términos.

Para comprobar la hipótesis del presente trabajo, se realizó una regresión resuelta por Mínimos Cuadrados Ordinarios (MCO) entre el volumen de búsquedas normalizados de los términos relacionados con el ciberacoso, como variable dependiente y, como variable independiente, el número de personas que declararon ser víctimas de ciberacoso, para cada año respectivamente. Cabe resaltar que ambas series corresponden a resultados por entidad federativa, siendo que en cada regresión se emplearon 32 observaciones respectivamente.

Al agregar millones de búsquedas en Google, los términos relacionados al ciberacoso se relacionan consistentemente con la demografía de aquellos de los que se espera que realicen una búsqueda relacionada. Para 2019, la suma de los términos de búsqueda explica el 54% de la variación en el número de víctimas auto reportadas de ciberacoso y existe una correlación entre ambas del 74%. Para 2020, la suma de los términos de búsqueda explica el 70% de la variación en el número de víctimas auto reportadas de ciberacoso y existe una correlación entre ambas del 84%. Para el 2021, la tendencia se mantiene, la suma de los términos de búsqueda explica el 62% de la variación en el número de víctimas auto reportadas de ciberacoso y existe una correlación entre ambas del 80% (ver tabla 3).

Tabla 3

**RELACIÓN ESTADÍSTICA ENTRE DATOS OBTENIDOS DE GOOGLE TRENDS Y DATOS DEL MOCIBA. MÉXICO, POR ENTIDAD FEDERATIVA, 2019, 2020 Y 2021**

Año	Google Trends	MOCIBA	Coefficiente de correlación	Probabilidad	Est. t	R <sup>2</sup> ajustada
2019	Volumen de búsqueda de	Víctimas auto reportadas por	0.74	1.17e <sup>-06</sup>	6.06	0.54
2020	términos, normalizado, por	entidad federativa	0.84	1.50e <sup>-09</sup>	8.56	0.70



2021	entidad federativa	0.80	5.09e <sup>-08</sup>	7.21	0.62
------	--------------------	------	----------------------	------	------

Elaboración del autor. El estadístico  $t$  y la  $R^2$  proceden de una regresión resuelta por MCO, entre el volumen de búsqueda de términos, normalizado, de los términos relacionados al ciberacoso (ver tabla 2) y el número de víctimas auto reportadas por entidad federativa en el MOCIBA, para cada año respectivamente.

Estas relaciones se mantienen a pesar de las altas señales de ruido que se pueden argumentar, ya que las personas pueden escribir estas palabras en el buscador por razones no relacionadas a ser víctimas de ciberacoso y, de la misma manera, no todas las víctimas de ciberacoso necesariamente buscan estos términos en Google. Sin embargo, al verificar las búsquedas relacionadas a estos términos, se puede argumentar que un gran número de búsquedas están relacionadas a la ocurrencia, propia o cercana, de una situación de ciberacoso, por ejemplo: “*qué es el ciberacoso*”, “*ciberbullyng*”, “*ciberacoso en México*”, “*causas del ciberacoso*”, “*cómo prevenirlo*” y “*qué hacer si hay ciberacoso*”.

Para el 2020 el aumento en esta correlación puede ser atribuida a un incremento en el número promedio de horas de uso diario que registró la población mexicana mayor de 12 años que, del 2019 al 2020, aumentó 18.5%, siendo el segmento de población de 12 a 19 años que más incremento registró, con 32.62% (ver tabla 4).

Tabla 4

**PROMEDIO DE HORAS DE USO DIARIO DE INTERNET DE LA POBLACIÓN DE 12 AÑOS Y MÁS<sup>1</sup>, POR CONDICIÓN DE HABER VIVIDO CIBERACOSO Y GRUPOS DE EDAD SEGÚN DOMINIO**

	Tasa de cambio				
	2019	2020	2021	2019-2020	2020-2021
Promedio México	4.12	4.88	4.99	18.51%	2.17%
De 12 a 19 años	4.32	5.74	5.67	32.62%	-1.13%
De 20 a 29 años	5.07	5.73	6.04	13.18%	5.29%
De 30 a 39 años	4.27	4.92	5.21	15.25%	5.89%
De 40 a 49 años	3.67	4.20	4.20	14.48%	-0.03%
De 50 a 59 años	3.00	3.57	3.65	19.07%	2.40%
De 60 años y más	2.44	2.93	3.06	20.03%	4.20%
Sí vivió ciberacoso	5.10	5.78	5.94	13.37%	2.74%
No vivió ciberacoso <sup>2</sup>	3.81	4.64	4.72	21.85%	1.74%

Elaboración del autor con datos del Módulo sobre Ciberacoso (MOCIBA), 2019, 2020 y 2021.

<sup>1</sup> Incluye sólo a las personas que declaran haber usado Internet diariamente.

<sup>2</sup> Incluye aquellas personas que declaran no saber si vivieron ciberacoso.

#### 4. Conclusiones

Los resultados presentados de ambas fuentes de datos, por un lado, las encuesta del MOCIBA y, por el otro, una consulta de búsquedas web para los mismos periodos, muestran una alta correlación y, se puede argumentar, que los patrones de búsqueda web explican en buena medida la demografía de las personas que han sido víctimas de ciberacoso. En este sentido, se comprueba la hipótesis planteada que pretendía encontrar una relación estadística entre las personas que declararon ser víctimas de



ciberacoso en la encuesta y los patrones de búsqueda de términos relacionados al ciberacoso, para 2019, 2020 y 2021.

Primero, basta con mencionar que los datos presentados provienen de dos metodologías completamente diferentes. Por un lado, los datos provenientes de la encuesta tienen un arreglo y tienen la ventaja de provenir de una muestra probabilística y ponderada que representa a la población, sin embargo, presenta algunas limitantes en cuanto a la información que recolecta: cobertura temática, tiempo y factores inhibidores.

Los datos provenientes de Google Trends tienen la desventaja de no provenir de una muestra ni permite conocer las características del individuo, tan solo se podría afirmar que provienen de usuarios que realizan una búsqueda con esos términos, sin embargo, cumplen la exigencia de mostrar tendencias o intereses agregados de búsquedas, para múltiples segmentos de tiempo y para distintas áreas. En múltiples disciplinas han encontrado su utilidad para estudiar fenómenos sociales diversos y han mostrado su adherencia con otras fuentes “oficiales”, por lo que no se descarta como una fuente para robustecer un planteamiento (Fong, 2017; Nagao et al., 2019; Prado-Román et al., 2021; Scheitle, 2011; Stephens-Davidowitz, 2013).

El presente trabajo no se cuestiona el origen ni las implicaciones del ciberacoso en México, tan sólo lo considera como un problema que existe en la esfera pública digital para proponer un acercamiento metodológico y rastrear fuentes válidas para su estudio y comprensión. No pretende desestimar ningún instrumento ni metodología, sino por el contrario, proponer metodologías híbridas para entender de mejor forma los fenómenos sociales.

La tecnología, las estructuras de comunicación y los flujos de autoinformación que hoy imperan nos permiten operacionalizar fenómenos sociales desde distintos puntos. Para la investigación social se abre un camino que puede ser explorado para construir metodologías más robustas que incluyan las bondades de distintas fuentes de información que, combinadas, permitan hacer indagaciones más profundas y ofrecer explicaciones más completas. Respecto al ciberacoso, a partir de esta metodología propuesta, se pueden correlacionar matrices de búsqueda que contengan términos cercanos a ciberdelitos específicos y encontrar patrones regionales o temporales que ayuden a entender mejor estos fenómenos. Asimismo, esto abre un camino para la discusión y la investigación para otro tipo de fenómenos y problemas de la esfera pública digital, para comprenderlos en mayor medida.

## 5. Bibliografía

Alexey, S. (2018). Hannah Arendt, Jürgen Habermas, and rethinking the public sphere in the age of social media. *Russian Sociological Review*, 17(4), 88–102. <https://doi.org/10.17323/1728-192X-2018-4-88-102>

Calhoun, C. (2017). Facets of the public sphere: Dewey, Arendt, Habermas. In A. S. D.-G. and A. L. Fredrik Engelstad, Håkon Larsen, Jon Rogstad, Kari Steen-Johnsen, Dominika Polkowska (Ed.), *Institutional Change in the Public Sphere: Views on the Nordic Model* (pp. 23–45). De Gruyter. <https://doi.org/10.1515/9783110546330-003>

Castrelo, V. (2018). La esfera pública habermasiana. Su obsolescencia en tiempos de nuevas plataformas digitales. *InMediaciones de La Comunicación*, 13(1), 71. <https://doi.org/10.18861/ic.2018.13.1.2826>

Enjolras, B., & Steen-Johnsen, K. (2017). The digital transformation of the political public sphere: A sociological perspective. In K. Steen-Johnsen, F. Engelstad, H. Larsen, J. Rogstad, D. Polkowska, A. S. Dauber-Griffin, & A. Levertton (Eds.), *Institutional Change*



*in the Public Sphere: Views on the Nordic Model* (1st ed., Vol. 31, pp. 99–117).  
<https://doi.org/10.1515/9783110546330-006>

Fong, W. M. (2017). Big Data, small Pickings: predicting the stock market with Google Trends. *The Journal of Index InvestIng*, Spring, 75–82.

García Zaballos, A., Huici, H., Puig Gabarró, P., & Iglesias Rodríguez, E. (2021). Cerrando la brecha de conectividad digital: Políticas públicas para el servicio universal en América Latina y el Caribe. In *Cerrando la brecha de conectividad digital: Políticas públicas para el servicio universal en América Latina y el Caribe*. Banco Interamericano de Desarrollo. <https://doi.org/10.18235/0003066>

Groves, R. M., & Lyberg, L. (2010). Total survey error: past, present, and future. *Public Opinion Quarterly*, 74(5), 849–879. <https://doi.org/10.1093/poq/nfq065>

INEGI. (2020). *Comunicado: 094*. INEGI, IFT y Secretaría de Comunicaciones y Transporte.

INEGI. (2022a). *Comunicado de prensa núm. 364/22. Módulo sobre ciberacoso 2021*. INEGI.

INEGI. (2022b). *Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares*. INEGI.

Insights, D. (2019). Estudio: hábitos de los consumidores móviles en México, 2019. In *Deloitte Insights*. Recuperado de: <https://www2.deloitte.com/content/dam/Deloitte/mx/Documents/technology/Global-Mobile-Consumer-Survey.pdf>

Jaramillo, J. (2010). El espacio de los políticos en Habermas, alcances y límites de las nociones de esfera pública y política deliberativa. *Jurid. Manizales (Colombia)*, 7(1), 55–73.

Livio, M. (2013). Revisitando el “Debate Lippmann-Dewey” Acerca del valor de la opinión pública en la práctica científica. *II Congreso Nacional de Estudiantes y Graduados En Filosofía: La Filosofía En Su Contemporaneidad 21, 22 y 23 de junio de 2013. Mar Del Plata. Argentina*, 209–214.

López Téllez, D., Fernández- Cuevas, M. P., & Callejas-Téllez, A. (2022). La ley Olimpia: Un punto de inflexión en la regulación normativa penal de la violencia digital y mediática en México. *DIVULGARE Boletín Científico de La Escuela Superior de Actopan*, 9(18), 28–38. <https://doi.org/10.29057/esa.v9i18.9140>

Nagao, S., Takeda, F., & Tanaka, R. (2019). Nowcasting of the U.S. unemployment rate using Google Trends. *Finance Research Letters*, 30(April), 103–109. <https://doi.org/10.1016/j.frl.2019.04.005>

Pareja Sánchez, N., & Echeverría, M. (2014). Public Opinion and New Technologies. Reflections for its Analysis in Mexico. *Revista Mexicana de Opinión Pública*, 51–68.

Patchin, J. W., & Hinduja, S. (2015). Measuring cyberbullying: Implications for research. *Aggression and Violent Behavior*, 23, 69–74. <https://doi.org/10.1016/j.avb.2015.05.013>

Prado-Román, C., Gómez-Martínez, R., & Orden-Cruz, C. (2021). Google Trends as a Predictor of Presidential Elections: The United States Versus Canada. *American Behavioral Scientist*, 65(4), 666–680. <https://doi.org/10.1177/0002764220975067>

Reveilhac, M., Steinmetz, S., & Morselli, D. (2022). A systematic literature review of how and whether social media data can complement traditional survey data to study public opinion. *Multimedia Tools and Applications*, 81(7), 10107–10142. <https://doi.org/10.1007/s11042-022-12101-0>



Scheitle, C. P. (2011). Google's Insights for Search: A Note Evaluating the Use of Search Engine Data in Social Research. *Social Science Quarterly*, 92(1), 285–295. <https://doi.org/10.1111/j.1540-6237.2011.00768.x>

Schober, M. F., Pasek, J., Guggenheim, L., Lampe, C., & Conrad, F. G. (2016). Social Media Analyses for Social Measurement. *Public Opinion Quarterly*, 80(1), 180–211. <https://doi.org/10.1093/poq/nfv048>

Similarweb. (2022). *Mexico's Top Websites Ranking in April 2022*. Recuperado de: <https://www.similarweb.com/top-websites/mexico/>

Stephens-Davidowitz, S. (2013). Who Will Vote? Ask Google. In *Working Paper*.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics*, 118, 26–40. <https://doi.org/10.1016/j.jpubeco.2014.04.010>

Stephens-Davidowitz, S. (2019). *Todo el mundo miente. Lo que Internet y el big data pueden decirnos sobre nosotros mismos*. Paidós México.

Unidad General de Asuntos Jurídicos. (2022). *Ficha técnica Ley Olimpia*. Secretaría de Gobernación. Gobierno de México.

U Report México. (2020). *Ciberseguridad*. Recuperado de: <https://mexico.ureport.in/opinion/4746/>

Vázquez, M. (2013). Primeras aproximaciones a la esfera pública virtual. In M. Fernández & M. David López (Eds.), *Lo público en el umbral. Los espacios y los tiempos, los territorios y los medios* (Primera ed, pp. 138–163). Facultad de Periodismo y Comunicación Social Universidad Nacional de La Plata.

## Anexo

### Definiciones de las situaciones de acoso empleadas en el módulo sobre ciberacoso (MOCIBA)

Situaciones de ciberacoso	Definición
Contacto mediante identidades falsas	Situación donde una persona contacta a otra bajo una identidad falsa para molestarle o dañar.
Mensajes ofensivos	Situación bajo la cual una persona le envía a otros mensajes ofensivos para molestarle o dañar.
Insinuaciones o propuestas sexuales	Situación donde una persona le envía a otra fotos o videos de contenido sexual para molestarle.
Recibir contenido sexual	Situación donde una persona le envía fotos o videos de contenido sexual a otra, para molestarle.
Provocaciones para reaccionar de forma negativa	Situación en la cual una persona provoca a otra, ya sea molestandola o retándola, en línea para que reaccione de forma negativa.
Llamadas ofensivas	Situación bajo la cual una persona le llama a otra para dirigirle improperios.
Rastreo de cuentas o sitios web	Situación bajo la cual una persona vigila los sitios o cuentas en Internet de otra, para causarle molestia o daño.



Suplantación de identidad	Situación en la cual una persona toma la identidad de la víctima para enviar información falsa, insultar o para agredir a otros.
Críticas por apariencia o clase social	Situación bajo la cual una persona critica o se burla de otra, en línea, por su apariencia o clase social.
Publicación de información personal	Situación bajo la cual un informante, falso o verdadero, publica información personal de la víctima, vulnerando la privacidad.
Amenazar con publicar información personal, audios o video para extorsionar	Situación bajo la cual un informante, falso o verdadero, amenaza a una persona con publicar audios o videos de carácter personal, para obtener algo a cambio.
Publicar información personal, fotos o videos	Situación bajo la cual un informante, falso o verdadero, publica fotografías o videos de carácter personal de la víctima, vulnerando la privacidad.
Publicar o vender imágenes o videos de contenido sexual	Situación bajo la cual un informante, real o simulado, ofrece por medios digitales contenido sexual.

Elaboración del autor con datos del Módulo sobre Ciberacoso (MOCIBA), 2019, 2020 y 2021. INEGI.