

COMISIÓN TÉCNICA DE PRÁCTICAS DE BUENA GOBERNANZA (CTPBG) DE LA
ORGANIZACIÓN LATINOAMERICANA Y DEL CARIBE DE ENTIDADES FISCALIZADORAS
SUPERIORES (OLACEFS)

Concurso Regional sobre Buena Gobernanza, edición 2022

“Participación ciudadana en prevención a la corrupción”

Línea de investigación, Fiscalización y percepción ciudadana sobre corrupción

Título del trabajo:

**¿La percepción ciudadana sobre corrupción en Twitter contribuye a la labor
fiscalizadora?**

Carlos Enrique Cardoso Vargas

María de Jesús Lizbeth Rodríguez Luna

Auditoría Superior de la Federación

México

Julio, 2022

Correo electrónico: cecardoso@asf.gob.mx
Número telefónico: +52 5514 208797

Índice

Constancias	3
Resumen ejecutivo	5
Introducción	6
Objetivos de la Investigación	7
Revisión de Literatura	8
Desarrollo de la Línea de Investigación	10
<i>Recopilación y Depuración de la Información de Twitter.....</i>	<i>11</i>
<i>Análisis Cualitativo.....</i>	<i>14</i>
<i>Análisis Cuantitativo</i>	<i>16</i>
Conclusiones	19
Bibliografía y e-grafía	21
Anexo	23

Tablas y Figuras

Tabla 1. <i>Tuits contenidos en la base de datos por ubicación geográfica y principales palabras relacionadas con el tema de corrupción.....</i>	<i>13</i>
Figura 1. <i>Vinculación Corrupción-Fiscalización en la percepción de usuarios de Twitter....</i>	<i>18</i>
Tabla 2. <i>Temas más relevantes en la percepción asociada a la "corrupción" de usuarios de Twitter.....</i>	<i>23</i>
Tabla 3. <i>Vinculación de la palabra "corrupción" con las tres principales palabras de fiscalización en la percepción de usuarios de Twitter por entidad federativa (cifras en porcentaje).....</i>	<i>25</i>

Constancias



UNIDAD GENERAL DE ADMINISTRACIÓN
DIRECCIÓN GENERAL DE RECURSOS HUMANOS
DIRECCIÓN DE ADMINISTRACIÓN DE PERSONAL

UGA/DGRH/DAP/SSP"1"/155/2022

"2022, Año de Ricardo Flores Magón"

**COMISIÓN TÉCNICA DE PRÁCTICAS DE BUENA GOBERNANZA
DE LA ORGANIZACIÓN LATINOAMERICANA Y DEL CARIBE
DE ENTIDADES FISCALIZADORAS SUPERIORES (OLACEFS)
P r e s e n t e**

Por este conducto se hace constar que el **C. Carlos Enrique Cardoso Vargas** presta sus servicios en esta Auditoría Superior de la Federación desde el **01 de octubre de 2018**, actualmente como personal de **mando superior** ocupando el puesto de **Director/a General**, en la plaza identificada con la clave presupuestal número **10391-560-CF51082-0043**, en la **Dirección General de Auditoría Forense del Gasto Federalizado**, de conformidad con el expediente personal que obra en los archivos de la Dirección de Administración de Personal.

Asimismo, se indica que el correo electrónico institucional del C. Cardoso Vargas es **cecardoso@asf.gob.mx**.

Por lo anterior, se expide la presente a solicitud del interesado con motivo del **Concurso Regional sobre Buena Gobernanza con el tema "PARTICIPACIÓN CIUDADANA EN PREVENCIÓN A LA CORRUPCIÓN"**, a los veintidós días del mes de junio de dos mil veintidós, en la Ciudad de México.

ATENTAMENTE

ING. CARLOS LÓPEZ ALBARRÁN
Director de Administración de Personal



MRH

**COMISIÓN TÉCNICA DE PRÁCTICAS DE BUENA GOBERNANZA
DE LA ORGANIZACIÓN LATINOAMERICANA Y DEL CARIBE
DE ENTIDADES FISCALIZADORAS SUPERIORES (OLACEFS)
P r e s e n t e**

Por este conducto, y de conformidad con la información que obra en los archivos de la Dirección de Administración de Personal, se hace constar lo siguiente:

INFORMACIÓN	DESCRIPCIÓN
Nombre y apellido del participante	María de Jesús Lizbeth Rodríguez Luna
Cargo y área	Prestadora de servicios profesionales por honorarios mediante contrato número SP.DGRH-1166/2022, en la Dirección General de Auditoría Forense del Gasto Federalizado
Correo electrónico	mjlrodriguez@asf.gob.mx

Por lo anterior, se expide la presente a solicitud de la interesada con motivo del **Concurso Regional sobre Buena Gobernanza con el tema “PARTICIPACIÓN CIUDADANA EN PREVENCIÓN A LA CORRUPCIÓN”**, a los veintidós días del mes de junio de dos mil veintidós, en la Ciudad de México.

ATENTAMENTE

ING. CARLOS LÓPEZ ALBARRÁN
Director de Administración de Personal



MIRN

Resumen ejecutivo

El dinámico intercambio comercial, cultural y de información alrededor del mundo a finales del siglo pasado propició un impulso sin precedentes a la participación ciudadana. Las tecnologías digitales evolucionaron rápidamente en lo que va del siglo XXI y, ahora, usuarios de distintas latitudes pueden intercambiar información prácticamente sin restricciones, en cualquier formato y sobre cualquier tema. Estas características, y otras específicas de Twitter, hacen de esta red social el canal tecnológico más accesible, independiente y apto para la denuncia de actos irregulares de servidores públicos por parte de los ciudadanos.

En este contexto, consideramos que Twitter es una fuente para recabar información sobre la percepción ciudadana respecto de la corrupción y una herramienta con el potencial de aportar elementos para llevar a cabo acciones de fiscalización e investigaciones; por lo que, aquí presentamos una investigación que explora si la información relacionada con corrupción que se difunde por medio de Twitter es útil para el desarrollo de la fiscalización que realizan los Entes de Fiscalización Superior. Con tal propósito, este trabajo tiene los objetivos de (i) observar la relación que existe entre la percepción ciudadana sobre la corrupción, expresada en las redes sociales, y las actividades que son materia de la fiscalización, y (ii) discernir cómo contribuye la participación ciudadana a través de las redes sociales a la fiscalización.

Los hallazgos que identificamos muestran que la percepción ciudadana sobre la corrupción está asociada a los temas de fiscalización y que la participación ciudadana a través de Twitter puede contribuir a esta actividad pública, aportando información susceptible de ser utilizada como una brújula o termómetro para orientar *ex ante* o alertar sobre potenciales temas de revisión; pues, como muestra este estudio, se detectan temas tanto a nivel país, como otros con un carácter muy local que pueden ser de interés de diversos cuerpos auditores. Por estos motivos, consideramos que la información difundida a través de Twitter podría agregar un valor estratégico adicional a la labor fiscalizadora en sus diferentes etapas, comenzando por la planeación.

Introducción

Los cambios económicos profundos que experimentó la sociedad global a finales del siglo pasado, resultado de las crisis económicas y financieras alrededor del mundo, así como las políticas de apertura comercial adoptadas y la modificación del estilo de gobierno de las democracias occidentales significaron un impulso decisivo a la participación ciudadana en la vida pública. El intercambio comercial, cultural y de información propició la vinculación entre actores diferentes y alejados, generando dinámicas de asociación e interdependencia entre sectores económicos y sociales que se apoyaron en las herramientas de la nueva era tecnológica, comenzando por el Internet y el correo electrónico. Las tecnologías evolucionaron rápidamente, pasando desde aplicaciones de mensajería instantánea hasta redes sociales donde usuarios de distintas latitudes ya pueden intercambiar información prácticamente sin restricciones, en cualquier formato y sobre cualquier tema; sin embargo, se especializan de acuerdo con el formato de la información que se comparte, los temas, intereses e, incluso, el tono de la conversación, intercambio o debate.

En este sentido, con más de 200 millones de usuarios activos en el segundo trimestre de 2021 alrededor de la Tierra, 13.9 millones ubicados en México —novenos país con más usuarios en el mundo— (Dixon, 2022), Twitter se ha colocado como un espacio propicio para el debate político, la confrontación, la expresión de ideas y demandas sociales, incluyendo el activismo. También, permite la comunicación directa entre sociedad y gobierno, y ya no sólo con fines electorales o de legitimación sino que también la Suprema Corte de Justicia de la Nación (SCJN) en México, por medio de su Segunda Sala, reconoció que las cuentas de Twitter pertenecientes a los funcionarios públicos son de interés general para la sociedad, en virtud de que “la información contenida y difundida a través de la cuenta [...] tiene relevancia pública o interés general en cuanto puede contribuir al debate en una sociedad democrática al proyectar las actividades y expresiones que realiza[n]” (Ángeles, 2020) los actores públicos.

Estas características, sumadas a la rapidez con la que se difunde la información, a la relativa anonimidad que proporciona a las personas una vez que el mensaje se dispersa, así como a la independencia que provee a los usuarios en el intercambio de información, al permitirle sustraerse de controles oficialistas, de intereses e, incluso, de valores o principios dominantes, hacen de esta red social el canal tecnológico más accesible, independiente y apto para la denuncia de actos irregulares de servidores públicos por parte de los ciudadanos. Al mismo tiempo, desde el enfoque académico, Twitter es una fuente excepcional para recabar información sobre la percepción ciudadana respecto de la corrupción y, para las agencias públicas encargadas de combatir la corrupción, es una herramienta con el potencial de aportar elementos para llevar a cabo acciones de fiscalización e investigaciones.

Con el propósito de reflexionar sobre la relación entre la percepción ciudadana respecto de la corrupción, expresada en las redes sociales, y las actividades que son materia de las revisiones que llevan a cabo las Entidades de Fiscalización Superior (EFS), así como sobre la forma en que la participación ciudadana contribuye a la fiscalización, este documento presenta un análisis exploratorio del contenido de mensajes de Twitter (denominados tuits) que se estructura de la siguiente manera: inicia con los objetivos de investigación; a continuación, hace una breve revisión de la literatura sobre el tema; posteriormente, desarrolla la línea de investigación que proponemos, mientras que, la última sección presentan la conclusiones, limitaciones y posibles cursos de investigación futura.

Objetivos de la Investigación

En presente manuscrito busca indagar respecto al papel de la participación ciudadana en el combate a la corrupción y su posible incidencia en el proceso de fiscalización, tomando como insumo principal la percepción que los usuarios manifiestan en sus mensajes en Twitter. De acuerdo con nuestro conocimiento, es la primera vez que se investigan de manera conjunta

dichos temas utilizando grandes volúmenes de información (*big data*) de una de las principales redes sociales en la actualidad.

El ejercicio que se presenta en este documento consiste en un análisis exploratorio que busca entender si la información relacionada con corrupción que se difunde por medio de Twitter es un insumo útil para el desarrollo de la fiscalización que realizan las EFS. Para tal efecto, esta investigación tiene los siguientes objetivos:

- a) Observar la relación que existe entre la percepción ciudadana sobre la corrupción, expresada en las redes sociales, y las actividades que son materia de la fiscalización.
- b) Discernir cómo contribuye la participación ciudadana a través de las redes sociales a la fiscalización.

Revisión de Literatura

La participación ciudadana, entendida en un sentido amplio, como la manifestación de demandas por parte de la ciudadanía hacia sus gobiernos con la intención de incidir en la toma de decisiones públicas (Díaz, 2017), ha venido encontrando cada vez nuevos canales de exposición y se ha vuelto más autónoma, en buena parte, debido al acelerado desarrollo de las tecnologías de la información. Sin embargo, la participación social ha transitado por diferentes etapas y aún no se encuentra consolidada como elemento central en los procesos de decisión; si bien, a partir de la última década del siglo pasado, los gobiernos empezaron a incluir de manera frecuente acciones de participación ciudadana en sus políticas públicas.

Respecto de la participación ciudadana en la lucha contra la corrupción, Warren (2004) señala que ésta se traduce en actos que, por su propia naturaleza, van en contra de los principios democráticos, en tanto que consisten en una apropiación de lo público, en una afectación al interés general para obtener un beneficio privado y, en consecuencia, son opuestos a los procesos participativos. Por este motivo, el involucramiento de la ciudadanía frente a la corrupción se manifiesta, la mayoría de las veces, en forma de reivindicaciones,

reclamos o denuncias más que como propuestas de acción o de solución ante problemas públicos.

Al respecto, iniciativas para promover la participación civil mediante plataformas de crowdsourcing, como *I Paid a Bribe*, inaugurada en 2016 y con representación en 25 países, han mostrado ser una buena alternativa para hacer frente a la corrupción a pequeña escala, ya que estimulan a las autoridades a iniciar averiguaciones y perseguir los casos denunciados (Kannaiah, 2016). Estas plataformas proporcionan un sitio digital en el que las personas pueden denunciar, de manera anónima, experiencias directas de conductas irregulares de servidores públicos; sin embargo, su alcance es limitado.

Otros estudios se han interesado en observar la adecuación de las redes sociales como medio para evidenciar la corrupción en entornos específicos. Así, mediante el análisis del efecto de publicaciones sobre casos de corrupción en un blog, Enikolopov et al. (2018), encontraron que, en democracias imperfectas, “las redes sociales pueden influir en el mercado y desencadenar mecanismos de rendición de cuentas en grandes empresas controladas por el estado, incluso en un país con competencia política limitada y censura de los medios tradicionales” (p. 153), ya que permiten el intercambio de información de manera independiente. Por su parte, Tang et al. (2019) construyeron un panel de datos para un periodo de cinco años y un conjunto de 62 países, a partir del cual —mediante un análisis econométrico y controlando, entre otras, con una variable proxy del grado de adherencia a las normas sociales y éticas—, encontraron evidencia de que el uso de redes sociales tiene un efecto positivo en el nivel de corrupción percibida de los países medido con el Índice de Control de Corrupción del Banco Mundial.

Además, desde la Primavera Árabe, a las redes sociales, y principalmente a Twitter, les ha sido reconocida cierta capacidad de incidir en la opinión pública. Mediante el estudio de sucesos y factores relacionados con las protestas que llevaron al derrocamiento del gobierno de Hosni Mubarak en Egipto, Acemoglu et al. (2018) llegaron a la conclusión de que el análisis

de mensajes de Twitter tiene el potencial de ayudar a predecir el surgimiento de protestas sociales, las cuales, a su vez, pueden provocar cambios económicos y sociales.

En efecto, Twitter se ha convertido en un espacio de debate político, de confrontación y demanda social, aspectos que están propiciado que se utilice como canal de comunicación directa entre autoridades y ciudadanos, así como entre éstos para compartir información sobre temas de interés. Una ventaja adicional de Twitter es que, a diferencia de los medios de comunicación tradicionales y otras redes sociales —blogs o canales de video—, no está bajo el control de una organización, empresa o grupo en particular, por lo que puede utilizarse como plataforma para la denuncia de conductas irregulares, no sólo de servidores públicos sino también de empresas o actores del sector privado.

En la actualidad, existen escasos estudios sobre el papel de las redes sociales como medios de denuncia y los disponibles no han encontrado evidencia de que conduzca a investigaciones integrales sobre los incidentes reportados. En este sentido, este documento busca abonar sobre este tema, indagando si la información difundida por medio de Twitter puede convertirse en un insumo valioso para la fiscalización, sin dejar de reconocer que su explotación requiere una importante capacidad de procesamiento y análisis.

Desarrollo de la Línea de Investigación

Para los fines de este estudio asumimos que los mensajes recolectados de Twitter son, en sí mismos, apreciaciones o percepciones que los usuarios tienen y auto declaran sobre corrupción en esta red social, sin entrar al análisis de si éstos contienen o no denuncias o reportes de casos concretos de corrupción. Tampoco es relevante la definición de los términos asociados a la corrupción incluidos en el glosario, dado que el objetivo no es confrontar o interpretar el significado que los usuarios de las redes sociales otorgan a los conceptos, sino constatar la existencia de participación por esa vía y observar lo que reportan para, en último término, dirimir si el contenido de los tuits está efectivamente relacionado con la fiscalización.

Los siguientes apartados exponen cómo se llevó a cabo la recopilación y la depuración de los datos; en seguida desarrollamos, primero, el análisis cualitativo y, después, el cuantitativo; en cada apartado presentamos tanto la metodología aplicada, los resultados obtenidos y una breve discusión de éstos.

Recopilación y Depuración de la Información de Twitter

Para la recopilación de datos se utilizó una interfaz de programación de aplicación (API, por sus siglas en inglés) pública de Twitter que permite acceso a la información de los tuits publicados. La extracción se realizó mediante un algoritmo desarrollado en el software estadístico R a través de una cuenta de desarrollador de Twitter que permitió interactuar con la API para recabar información cada tercer día durante el período del 1 de diciembre de 2021 al 31 de marzo de 2022.

Con el objeto de aproximarnos a la percepción de los usuarios de Twitter respecto al tema de corrupción, se obtuvieron los textos de los tuits publicados en castellano cuyo contenido incluyera al menos una palabra o frase de un listado de términos relacionados con el fenómeno de la corrupción¹. El listado comprendió tanto palabras formalmente relacionadas con la corrupción como: “fraude”, “nepotismo”, “soborno”, “peculado”; como expresiones coloquiales comúnmente usadas en México para referirse a dicho fenómeno, por ejemplo: “padrinazgo”, “mordida”, “moche”, “transa” o “póngase la del Puebla”, por mencionar algunas². A partir de la recopilación de información se conformó una base de datos con 4,421,728

¹ La corrupción engloba múltiples prácticas y actos muy disímiles (Arellano-Gault, 2020) que lo vuelve una manifestación multidimensional, por ello, en este estudio se optó por considerar diferentes palabras o frases coloquiales que reflejan el lenguaje usado por la población en México para hablar de la corrupción (Casar, 2018).

² El listado completo se conformó por las siguientes expresiones: abuso de autoridad, abuso de funciones, abuso de poder, acceso condicionado, año de hidalgo, beneficio indebido, carrusel, catafixia, chayote, clientelar, clientelismo, cochupo, cohecho, colusión, ¿cómo nos arreglamos?, compra de votos, conexiones políticas, conflicto de intereses, contratación indebida, corrupción, desacato judicial, desfalco, desvió de recursos, ejercicio indebido de funciones, empresa fantasma, encubrimiento, enriquecimiento ilícito, extorsión, favoritismo, financiamiento clandestino, financiamiento ilegal, fraude, fuero, lavado de dinero, malversación, moche, mochilas pa los cuates, nepotismo, obstrucción de la justicia, padrinazgo, participación ilícita, peculado, póngase guapo, póngase la del puebla, prebendas, prestanombres, privilegiar a empresas, soborno, tráfico de influencias, transa, uso indebido de recursos y utilización indebida de información.

registros, que incluyó el texto los tuits y otras variables asociadas con el mensaje —fecha de creación, cuenta/id de usuario, identificación sobre si el texto fue un tuit o retuit, país y ubicación registrada por el usuario que posteó el mensaje, entre otras—.

La depuración de la base de datos implicó diversos procesos, el primero consistió en eliminar la información duplicada, para lo cual se excluyeron tuits repetidos o que fueron nuevamente posteados mediante un retuit. En segundo lugar, se identificó con el máximo grado de exactitud posible que los tuits fueran publicados por usuarios localizados en México a través del desarrollo de un algoritmo de análisis de texto para determinar si la localización declarada por los usuarios en sus mensajes indicaba una ciudad o entidad federativa mexicana. Los mensajes que hacían referencia a usuarios ubicados fuera de México fueron eliminados y los identificados al interior de este país fueron clasificados por entidad federativa, en tanto que, los tuits en los que los usuarios únicamente manifestaron que su ubicación correspondía al país sin indicar una ciudad o entidad federativa en específico, fueron conservados y se les etiquetó como una localización genérica.

Posteriormente, se llevó a cabo una limpieza del texto de los mensajes consistente en quitar palabras vacías³ (por ejemplo, “a”, “en”, “la”, “entonces”, entre otras) necesarias para la conformación de los mensajes, pero carentes, por sí mismas, de significado clave o relevante para el análisis. También se retiraron los hipervínculos o links a páginas de internet, que, si bien permiten a los lectores de los tuits dar seguimiento o profundizar sobre el contenido del mensaje, no contienen un texto que se pueda analizar. En el mismo, sentido se prescindió de los caracteres referidos a números, símbolos especiales o emoticones.

Finalmente, se creó una variable para identificar la palabra del glosario de términos relacionados con la corrupción que aparece en primer lugar dentro del texto del tuit. La base de

³ En la literatura sobre el tema estas palabras son conocidas como *stopwords*.

datos resultante consistió en 449,230 tuits depurados que fueron publicados por usuarios localizados en México y que aproximan la percepción de los individuos sobre la corrupción.

Tabla 1.

Tuits contenidos en la base de datos por ubicación geográfica y principales palabras relacionadas con el tema de corrupción

Ubicación geográfica	Corrupción	Fraude	Chayote	Extorsión	Fuero	Nepotismo	Moche	Otros corrup.	Total
MEXICO (genérico)	73,799	13,229	12,404	4,340	3,583	3,558	3,240	25,594	139,747
CIUDAD DE MEX.	45,882	8,556	8,104	2,926	2,329	2,569	2,138	16,827	89,331
JALISCO	14,863	2,244	3,622	558	867	638	554	4,494	27,840
NUEVO LEON	12,329	2,099	2,455	568	843	508	849	4,653	24,304
PUEBLA	8,709	1,718	1,307	1,032	503	257	393	4,048	17,967
VERACRUZ	6,389	1,270	1,250	626	351	288	294	2,468	12,936
QUERETARO	5,532	822	913	239	285	248	227	1,733	9,999
GUANAJUATO	5,111	636	804	616	228	279	208	1,314	9,196
ESTADO DE MEX.	4,453	789	1,091	515	272	211	178	1,526	9,035
BAJA CALIFORNIA	4,091	727	932	244	290	161	192	1,370	8,007
SONORA	3,738	650	539	347	218	228	156	1,296	7,172
OAXACA	2,535	656	426	1,588	133	104	104	1,093	6,639
QUINTANA ROO	3,301	663	525	383	164	156	141	1,276	6,609
MORELOS	3,181	501	637	592	211	132	152	1,141	6,547
SINALOA	3,331	635	579	200	160	161	264	1,126	6,456
YUCATAN	3,238	678	649	165	162	132	169	1,156	6,349
COAHUILA	2,927	604	538	261	136	117	195	1,148	5,926
MICHOACAN	2,652	659	471	582	219	64	167	994	5,808
CHIHUAHUA	2,871	680	569	210	205	99	174	880	5,688
TAMAULIPAS	2,590	587	633	282	202	142	192	963	5,591
SAN LUIS POTOSI	1,941	817	338	350	145	72	82	890	4,635
TABASCO	2,160	483	597	196	186	76	114	802	4,614
HIDALGO	2,292	430	357	282	164	123	115	820	4,583
DURANGO	2,577	319	284	270	75	106	97	571	4,299
CHIAPAS	1,679	331	342	130	77	52	79	648	3,338
AGUASCALIENTES	1,625	338	294	102	89	86	82	565	3,181
GUERRERO	1,409	227	337	226	74	77	71	485	2,906
BAJA CALIF. SUR	1,252	339	218	97	49	52	56	428	2,491
ZACATECAS	998	233	118	381	56	42	50	286	2,164
COLIMA	935	283	149	174	53	43	40	332	2,009
TLAXCALA	753	134	60	86	58	57	34	332	1,514
CAMPECHE	610	135	163	58	48	53	24	311	1,402
NAYARIT	438	83	125	29	37	18	22	168	920
Total	230,191	42,555	41,830	18,655	12,472	10,909	10,853	81,738	449,203

Fuente: Elaboración propia con base en los tuits recopilados del 1 de diciembre de 2021 al 31 de marzo de 2022 y de acuerdo con los criterios de depuración descritos en esta sección.

La información contenida en la Tabla 1 muestra que el término más frecuentemente reportado por los usuarios dentro de la base de datos es “Corrupción”, presente en 51 por ciento de los tuits revisados, seguida por “Fraude” y “Chayote”, ambas con 9 por ciento.

Al observar la cantidad de tuits por entidad federativa, es evidente la concentración en la publicación de tuits relacionados con los términos que conforman el glosario, destaca la Ciudad de México (20%), Jalisco (6%) y Nuevo León (5%). Este comportamiento no sorprende puesto que en dichas entidades federativas se ubican las tres grandes áreas metropolitanas del país, donde los usuarios pueden ser propensos a tener una participación más activa en la difusión de hechos vinculados con actos de corrupción, sin dejar de lado el hecho de que las concentraciones poblacionales favorecen la existencia de mejores condiciones de acceso a Internet, información, cultura o educación, lo que influye en un mayor posteo de mensajes.

Análisis Cualitativo

Con el objetivo de explorar el contenido de los mensajes, la estrategia consistió en dividir en palabras del texto de cada uno de los tuits de manera individual respetando el orden consecutivo que tenía, posteriormente se empleó un algoritmo para calcular qué tan frecuente es que una palabra “x” fuera seguida de otra palabra “y”. Este ejercicio conocido en la literatura como obtención de *bigramas*⁴ se realizó inicialmente para la palabra “corrupción” del glosario de términos. Asimismo, para reducir la dimensionalidad de todos los *bigramas* obtenidos, se seleccionaron las veinte expresiones con el mayor número de apariciones. El proceso de identificación consistió en la revisión directa de cada expresión, por lo que se descartaron frases sin sentido o cuyo significado no tuviera relación aparente con el término “corrupción”.

Las expresiones (*bigramas*) más recurrentes dentro de los tuits que contienen el término “corrupción” fueron nombres de funcionarios, exfuncionarios, instituciones, proyectos y programas cuya confiabilidad ha sido cuestionada por la opinión pública. En este sentido, la mayoría de los *bigramas* está asociada con temas difundidos a través de medios de

⁴ En el análisis de textos mediante herramientas de análisis de datos esta técnica es conocida como *tokenización* por n-gramas, consistente en descomponer el texto en unidades del tamaño que especifique el investigador. El ejercicio descrito en este trabajo consistió en construir conjuntos de dos palabras (denominados bigramas), utilizando el paquete *tidytext* en el software estadístico R. Para una aproximación práctica, puede consultarse a Silge y Robinson (2017).

comunicación; sin embargo, también hay términos relevantes en materia de fiscalización y cuyo análisis más profundo podría aportar información valiosa para la actividad auditora, como son “adjudicaciones directas”, “programas sociales”, “medio ambiente” y “servidores públicos”.

Al replicar este ejercicio a otros términos del glosario, se encontraron *bigramas* referidos a expresiones, a un fiscal, el proyecto de un tren, un programa social, entre otras. Asimismo, en cada término identificamos las expresiones más representativas con base en su frecuencia y las que conciernen más a la fiscalización. Así, en la palabra “fraude” la frase más recurrente fue “fraude electoral”, pero “fraude fiscal” y “fraude inmobiliario” también podrían ser indicios de valor para la auditoría. La palabra “abuso de autoridad” se encontró asociada con una alcaldesa y con la policía; mientras que “chayote” se asocia de manera unívoca con periodistas. “Desvío de recursos” estuvo fuertemente vinculada con un exgobernador, pero también con delitos electorales y con empresas fantasma. En “desfalco” se encontró a una servidora pública de alto nivel y a un exgobernador, ambos, con acusaciones de malversación o desvío de recursos y enriquecimiento ilícito.

En el léxico “enriquecimiento ilícito” aparecieron menciones a nombres de servidores o exservidores, representantes y políticos. Con “moche”, se asocian empresas extranjeras, facturas falsas, compañías asociadas, obra pública y agilizar trámites. Mientras que, en “nepotismo”, inesperadamente, se hallaron expresiones referentes al sector salud: medicinas, medicamento para el cáncer y vacunas; lo cual sugiere que los usuarios perciben que, en la compra de medicamentos, podría existir una irregularidad asociada con este tipo de corrupción.

Si asumimos que las expresiones identificadas son manifestaciones de lo que los usuarios de Twitter perciben como corrupción o, más específicamente, como el tipo de corrupción que indican las palabras del glosario, se puede entrever que los mensajes de Twitter contienen percepciones relevantes en materia de fiscalización, debido a que se trata de asuntos susceptibles de ser revisados y cuyo análisis detallado puede proveer elementos para dirigir y realizar con mayor precisión la práctica de las auditorías, incluso, desde la planeación.

Al hacer un análisis más detallado, considerando la dimensión estatal y sólo el léxico de “corrupción”, los resultados arrojan diferencias y similitudes entre las 32 entidades federativas, como se indica en la Tabla 2 del Anexo. De acuerdo con las mayores frecuencias observadas en este ejercicio, los cinco temas más relevantes se refieren a nombres o lugares de interés local; a irregularidades, indicios o factores de riesgo actuales; a dependencias, autoridades o temas de interés local, y a sectores de interés federal, entre los que están salud, educación, seguridad, economía, energía, recursos y obra públicos, y a obras emblemáticas del gobierno federal.

En todas las entidades se identificaron, entre las principales expresiones, nombres de personas, de lugares o de incidentes. La mención de actores políticos puede estar relacionada no sólo con acusaciones de corrupción, sino que también puede indicar propósitos políticos, podrían ser mensajes partidistas afines o de oposición; entre los nombres se encuentran funcionarios y exfuncionarios públicos electos, así como otros servidores y exservidores públicos designados en casi todas las circunscripciones estatales. Al respecto, se observó que, a medida que la frecuencia de las expresiones disminuye, incrementan aquellas cuyo contenido son nombres de personas con relevancia en el contexto subnacional.

Análisis Cuantitativo

Para profundizar más sobre la asociación corrupción-fiscalización en los textos de los tuits, se aplicó un algoritmo que identificó si al interior del mensaje se encontraba alguna palabra de las contempladas en un listado de con términos relativos a la fiscalización, entre las que se incluyeron, “obra pública”, “adquisiciones”, “contrato”, “denuncia” o “licitaciones”. A este listado de palabras⁵ se añadieron, términos específicos relacionados con el tema de la pandemia provocada por el virus SARS-CoV-2, en virtud de que durante este acontecimiento

⁵ El listado final abarcó 192 palabras o frases asociadas con la fiscalización superior, vocablos referidos al sector salud y dependencias gubernamentales.

se realizaron diversas adquisiciones públicas de bienes y contratación de servicios u obra pública que pudieron ser realizadas en condiciones potencialmente irregulares.

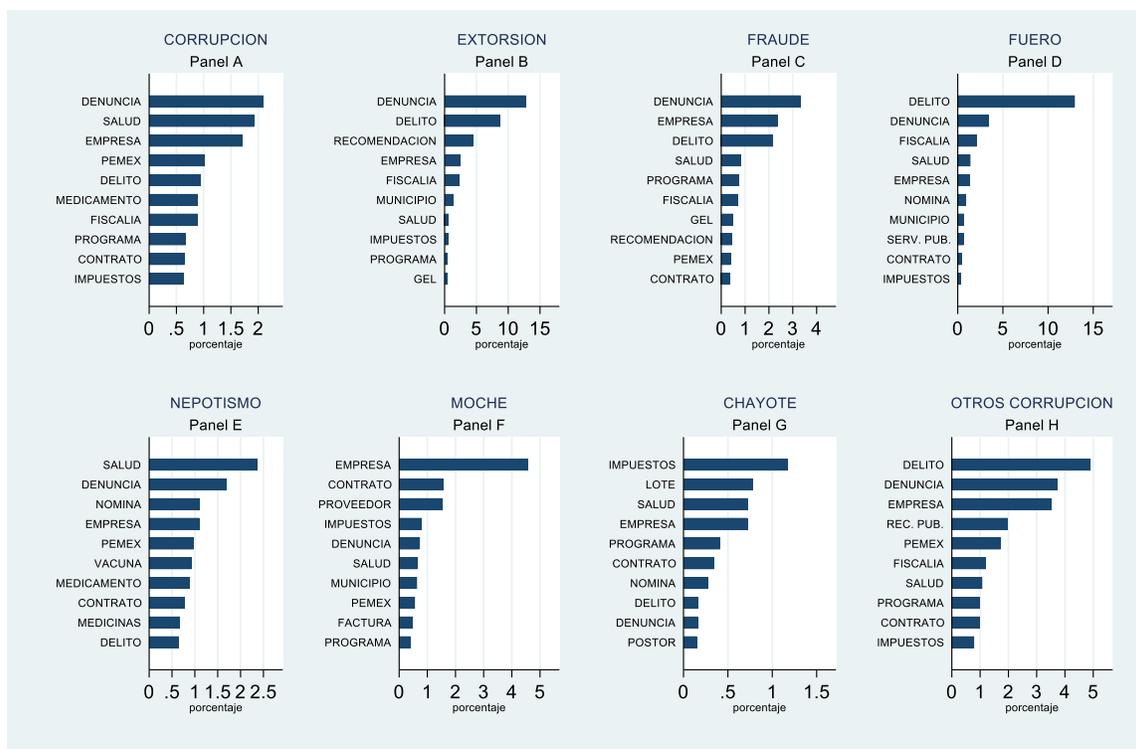
De los 449,203 tuits contenidos en la base de datos, se encontró que 94,550 (21%) de ellos contienen palabras vinculadas a temas de fiscalización, es decir, 1 de cada 5 tuits cuyo mensaje incluye una palabra o frase referente al glosario de corrupción también considera una palabra o frase vinculada con algún tópico relativo al tema de fiscalización.

De esa submuestra de 94 mil tuits, se examinaron los términos de fiscalización *vis à vis* las principales palabras asociadas a corrupción (Figura 1 del Anexo) que aparecen más frecuentemente en los tuits —corrupción, chayote, extorsión, fuero, nepotismo, moche y otros— y se observó que la palabra “salud” aparece en todos ellos; en algunos, su porcentaje de menciones se ubicó en las primeras cinco posiciones, tal es el caso de los tuits que incluyen las palabras corrupción, fraude, fuero, nepotismo y chayote. Este comportamiento parece apuntar a que los usuarios que postearon los tuits analizados perciben cierta asociación entre el sector Salud y diversos actos irregulares; de hecho, considerando solo los datos contenidos en el Panel A de la Figura 1, el número de menciones durante el período de análisis ascendió a 4,810⁶. Asimismo, ese argumento se ve reforzado al encontrar palabras como “medicamento”, “gel”, “vacuna” en mensajes relacionados con términos de “corrupción” (Panel A), “extorsión” (Panel B) y “Nepotismo” (Panel E), de manera respectiva.

Otro aspecto sobresaliente de la Figura 1 es que los usuarios en sus mensajes vinculan todos los términos asociados a corrupción con la palabra “denuncia”, la cual en 6 de los 8 paneles presentados aparece en las primeras dos posiciones. Esto podría apuntar a que los tuiteros estarían percibiendo a Twitter como un canal para manifestar aspectos relacionados con alguna posible irregularidad.

⁶ El método de cálculo empleado para la obtención del resultado expuesto es el siguiente: si el porcentaje de tuits en el que coexisten tanto el término “corrupción” como “denuncia” es de 2.09% (Panel A) y el total de tuits en la base de datos que incluyen la palabra “corrupción” es de 230,191 (Tabla 1), entonces el número de menciones de conjunto de ambas palabras es equivalente a $2.09\% \times 230,191 = 4,810$.

Figura 1.

Vinculación Corrupción-Fiscalización en la percepción de usuarios de Twitter

Fuente: Elaboración propia mediante el uso de la base de datos descrita en este documento

Asimismo, con independencia del término de corrupción que se use, la palabra “empresa” aparece en todos los paneles de la Figura 1 y, en algunos de ellos, en las primeras posiciones, siendo en la palabra “moche” —asociada a alguna cuota ilegal otorgada a un superior o un soborno⁷— donde tiene el mayor peso relativo, es decir, donde los usuarios perciben que ese acto de corrupción puede estar asociado a “empresas” en que el pudo haber de por medio un soborno.

Dada la heterogeneidad en la cantidad de los mensajes a nivel nacional, podría pensarse que el peso que tienen las principales entidades federativas en el total de tuits significa que los resultados antes mencionados solo están recogiendo la percepción de esos lugares y no así

⁷ Definido en Casar (2018), por su parte, la Real Academia de la Lengua Española define “moche” como: dádiva que se otorga o se solicita para recibir (o por haber recibido) algún beneficio valorado en dinero, habitualmente en calidad de expectativa de una evasión institucionalizada.

expresiones de usuarios en otras ubicaciones. Para abonar sobre este asunto, en la Tabla 3 del Anexo⁸, se presenta la vinculación entre palabras de corrupción y las tres principales palabras asociadas a fiscalización al nivel de entidad federativa. Los resultados sugieren que existe una percepción generalizada en el país sobre la vinculación entre “corrupción” y las palabras “salud”, “empresa” y “denuncia” que se manifiesta en la recurrencia de textos que contienen ambos términos. No obstante, también se encontraron palabras que se referían a aspectos más locales, como las menciones a petróleo en Tabasco, medicamento en Aguascalientes y Coahuila, entre otras.

Conclusiones

Mediante los análisis presentados en este documento se encontró que, en México, la corrupción es un tema vigente y de interés generalizado, muestra de ello es la gran cantidad de tuits publicados en sólo cuatro meses, los cuales, como pudo advertirse, captan la percepción ciudadana sobre dicho fenómeno. Estamos conscientes de que esa percepción observada en los mensajes de Twitter puede estar influida, en cierta medida, por información diseminada por éste y por otros canales informativos, debido a que algunas de las expresiones más frecuentes son temas del dominio público. Al mismo tiempo, advertimos que las expresiones con menos menciones llegan a indicar actos o conductas irregulares que son más concretos y específicos a un nivel subnacional.

Independientemente de esto, se pudo observar que los usuarios de esta red social tienen percepciones bien definidas respecto de lo que es corrupción y de los actos relacionados con los diferentes tipos de corrupción. Lo anterior implica que es probable que los reportes y, en su caso, denuncias —realizadas en un lenguaje más simple y expedito— que realizan

⁸ En virtud de su extensión y detalle, una parte amplia del texto y cuadros que comprenden los análisis cualitativos y cuantitativos a nivel entidad federativa quedó fuera de este documento, no obstante, esa información puede ser solicitada a los autores para su consulta.

mediante un tuit, aporten datos que brinden indicios para identificar el acto y al sujeto activo de probables actos irregulares, entre otros. En este sentido, un resultado relevante del presente análisis es que la participación ciudadana y la denuncia relacionadas con corrupción están entre los principales temas mencionados en algunos estados, lo que respalda la tesis de que las redes sociales y Twitter, en particular, son medios adecuados para la lucha contra la corrupción por la vía de la denuncia.

Los hallazgos presentados muestran que, entre los temas asociados a la corrupción que son mencionados con más frecuencia se encuentran varios que están relacionados con la fiscalización que llevan a cabo las EFS. Esto indica que la percepción ciudadana sobre la corrupción y la fiscalización coinciden en un número importante de asuntos públicos. No es casualidad que uno de cada cuatro tuits incluye tanto términos contenidos en el glosario de palabras relativas a corrupción como tópicos referentes a la fiscalización.

Por tanto, este trabajo expone que la participación ciudadana a través de las redes sociales puede contribuir a la fiscalización como una brújula o termómetro que señale indicios de actos irregulares, así como los lugares, dependencias y sujetos que podrían estar relacionados con dichos actos. Esta información, puede llegar a ser estratégica en la labor fiscalizadora en sus diferentes etapas, comenzando por la planeación.

Sin embargo, es relevante subrayar que, dadas las limitaciones de espacio, este trabajo no agotó el análisis, por lo que, en el futuro se deberá analizar con mayor profundidad la naturaleza los tuits, para dilucidar si son denuncias o reportes de actos concretos; así como, observar con mayor detalle los hechos que se denuncian y determinar si se refieren a actos de corrupción o si, en cambio, exponen casos de negligencia o ineficacia de los actores públicos. En este sentido, también será necesario dirimir si los hechos publicados son auténticos o si se trata de manifestaciones dirigidas a afectar o a apoyar intereses políticos o partidistas.

Bibliografía y e-grafía

- Acemoglu, D., Tarek A. H., y Ahmed T. (2018, enero). The Power of the Street: Evidence from Egypt's Arab Spring. *The Review of Financial Studies* 31(1), pp. 1-42, <https://doi.org/10.1093/rfs/hhx086>
- Aguilar, L. (2010). Gobernanza: El nuevo proceso de gobernar. Fundación Friedrich Naumann para la Libertad.
- Alfred, Z. (2014). *Tweeting against corruption: Fighting police bribery through online collective action* [Tesis de Maestría, Universidad de Uppsala]. <https://www.diva-portal.org/smash/get/diva2:747107/fulltext01.pdf>
- Arellano-Gault, D. (2020). Corruption in Latin América, Nueva York, Routledge, 216 pp.
- Casar, A. (2018). Glosario de la corrupción. Revista Nexos, sección especial “Desarmar la corrupción”. México.
- Díaz, A. (2017). Participación ciudadana en la gestión y en las políticas públicas. *Gestión y Política Pública*, 36(2), 341-379.
- Enikolopov, R., Maria, P. y Konstantin Sonin (2018). Social Media and Corruption. *American Economic Journal: Applied Economics*, 10(1), pp. 150-174. <https://www.aeaweb.org/articles?id=10.1257/app.20160089>
- Kannaiah, V. (2016). I Paid a Bribe: Using Technology to Fight Corruption in India. *South Asia @ LSE*. <https://blogs.lse.ac.uk/southasia/2016/05/16/i-paid-a-bribe-using-technology-to-fight-corruption-in-india/>
- Merhi, I. y Ahluwalia, P. (2018). Digital Economy and Corruption Perceptions: A Cross-Country Analysis. *The International Journal of Digital Accounting Research*, 18, pp. 29-4. https://www.researchgate.net/profile/Mohammad-Merhi/publication/323954081_Digital_Economy_and_Corruption_Perceptions_A_Cross-Country_Analysis/links/5ace2fe00f7e9b189657f392/Digital-Economy-and-Corruption-Perceptions-A-Cross-Country-Analysis.pdf

- Mungiu-Pippidi, A. (2013). Controlling Corruption Through Collective Actions. *Journal of Democracy*, 26(1), pp. 101-115. John Hopkins University Press.
<https://www.journalofdemocracy.org/wp-content/uploads/2013/01/Mungiu-Pippidi-24-1.pdf>
- Ángeles, L. (2020). Reseña del Amparo en Revisión 1005/2018. Repositorio de la Suprema Corte de Justicia de la Nación
https://www.scjn.gob.mx/sites/default/files/resenias_argumentativas/documento/2020-10/res-EMMI-1005-18.pdf
- Tang, Z., Chen, L., Zhou, Z., Warkentin, M. y Gillenson, M. (2019). The effects of social media use on control of corruption and moderating role of cultural tightness-looseness. *Government Information Quarterly*, (36)4, pp. XXX.
<https://www.sciencedirect.com/science/article/abs/pii/S0740624X18304994>
- Silge, J. y Robinson, D. (2017). Text Mining with R. <https://www.tidytextmining.com/>
- Warren, A. (2004). What Does Corruption Mean in a Democracy? *American Journal of Political Science*, (48)2, pp. 328-343.

Anexo

Tabla 2.

Temas más relevantes en la percepción asociada a la "corrupción" de usuarios de Twitter

Entidad	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_10	T_11	T_12	T_13	T_14	T_15	Total
CIUDAD DE MEXICO	0	2	1	3	4	3	2	1	0	2	1	0	0	0	1	20
JALISCO	4	3	1	4	3	1	3	1	0	1	0	0	0	0	0	21
NUEVO LEON	2	3	1	4	3	2	3	1	0	0	1	0	0	0	0	20
PUEBLA	7	0	3	6	2	0	1	1	0	0	0	0	0	0	1	21
VERACRUZ	5	2	2	4	3	1	2	1	0	1	0	0	0	0	0	21
QUERETARO	2	3	1	3	3	2	3	1	0	1	1	0	0	0	0	20
ESTADO DE MEXICO	2	5	3	3	1	1	0	1	1	1	0	1	1	0	0	20
GUANAJUATO	2	4	1	0	3	5	3	0	0	0	2	0	0	0	0	20
BAJA CALIFORNIA	3	3	2	2	4	3	1	1	0	1	0	0	0	0	0	20
YUCATAN	3	0	8	4	2	1	0	1	0	0	0	1	0	0	0	20
SONORA	4	4	0	2	2	4	2	1	0	0	0	1	0	0	0	20
OAXACA	6	2	3	2	1	1	0	1	1	0	0	2	0	1	0	20
SINALOA	3	3	2	2	4	1	4	0	1	0	0	0	0	0	0	20
QUINTANA ROO	7	2	4	1	1	2	1	1	1	0	0	0	0	0	0	20
MORELOS	5	2	4	3	1	1	0	1	1	0	0	2	0	0	0	20
COAHUILA	2	4	0	4	1	4	4	1	0	0	0	0	0	0	0	20
MICHOACAN	3	4	1	2	3	3	2	1	1	0	0	0	0	0	0	20
TAMAULIPAS	5	2	2	2	3	2	2	1	0	1	0	0	0	0	0	20
CHIHUAHUA	3	5	2	3	3	1	1	1	0	0	0	0	1	0	0	20
SAN LUIS POTOSI	4	3	2	4	1	2	1	1	1	1	0	0	0	0	0	20
TABASCO	4	1	2	5	2	3	1	1	1	0	0	0	0	0	0	20
HIDALGO	8	0	2	2	2	2	1	1	2	0	0	0	0	0	0	20
DURANGO	8	0	5	2	2	1	0	1	0	1	0	0	0	0	0	20
BAJA CALIFORNIA SUR	2	5	5	3	2	0	0	1	1	1	0	0	0	0	0	20
CHIAPAS	5	1	5	2	2	0	1	0	1	0	1	0	1	1	0	20
AGUASCALIENTES	7	1	2	0	3	5	1	0	0	0	0	0	0	1	0	20
GUERRERO	3	0	5	4	1	2	1	1	1	1	1	0	0	0	0	20
ZACATECAS	9	4	3	1	1	1	0	0	1	0	0	0	0	0	0	20

COLIMA	3	7	2	2	1	1	1	1	0	1	1	0	0	0	0	20
TLAXCALA	7	2	2	1	1	0	0	1	1	0	0	0	5	0	0	20
CAMPECHE	3	7	0	1	1	2	1	1	1	0	1	1	0	0	1	20
NAYARIT	0	4	7	1	2	3	1	1	0	0	0	1	0	0	0	20
Alusiones al tema	131	88	83	82	68	60	43	27	16	13	9	9	8	3	3	643
Promedio simple estatal	4.1	2.8	2.6	2.6	2.1	1.9	1.3	0.8	0.5	0.4	0.3	0.3	0.3	0.1	0.1	20

Notas: Cada estado cuenta con 20 expresiones más frecuentes, éstas se clasificaron en 15 temas. Número de veces que las expresiones aluden a cada tema por entidad federativa.

T_1: Nombres o lugares de interés estatal.

T_2: Irregularidades, indicios o factores de riesgo actuales

T_3: Dependencia, autoridad o temas de interés local.

T_4: Sectores de interés federal: Salud, educación, seguridad, economía, energía.

T_5: Obras emblemáticas del gobierno federal.

T_6: Nombres o casos federales que no están en primeros lugares en el ámbito nacional.

T_7: Programas sociales o federales en general.

T_8: Servicio público.

T_9: Participación ciudadana y términos relacionados.

T_10: Tipologías de la corrupción y políticas anticorrupción.

T_11: Dependencias y organismos federales.

T_12: Denuncia y términos relacionados.

T_13: Fiscalización.

T_14: Obras públicas locales.

T_15: Casos de corrupción o similares locales pasados.

Tabla 3.

Vinculación de la palabra “corrupción” con las tres principales palabras de fiscalización en la percepción de usuarios de Twitter por entidad federativa (cifras en porcentaje)

Palabra asociada a fiscalización	MEX																																	
	AGS	BC	BCS	CAMP	CHIA	CHIH	CDMX	COA	COL	DGO	EDO.	GTO	GRO	HGO	JAL	MEX (gen)	MICH	MOR	NAY	NL	OAX	PUE	QRO	QROO	SLP	SIN	SON	TAB	TAM	TLAX	VER	YUC	ZAC	
DENUNCIA	2.0	1.6	2.2	2.8	3.8	1.8	1.9		2.5	1.5	3.0	2.1	2.1	1.8		2.2	2.5	4.4	2.3	1.4	3.4	1.9		2.3	2.6	1.6	3.2		2.7	4.6	2.4	3.2	1.6	
SALUD		2.1	1.8		1.1	1.8	1.7	2.2		1.0	2.3	3.3	2.1	1.9	2.4	1.8	2.5	1.6	2.5	2.4		2.3	2.3	1.4	2.0	2.5	2.0	2.1	1.6		2.0	1.9	1.9	
EMPRESA	2.0	1.8		2.6	3.3	1.6	1.7	1.7	1.5		1.3	1.8	1.4	1.7	1.8	1.8	1.3		1.6	1.9		1.4	1.8	1.4		1.3	2.0	2.4	1.7		1.6	2.1		
FISCALIA										1.0								2.5			2.0									4.5			3.1	
DELITO			1.4	1.5																	2.4													
PROGRAMA PEMEX									2.2															1.8										
CONVENIO																												3.8						
MEDICAMENTO	0.8																																	
MUNICIPIO															2.2																			
MEDICINAS																							1.3											
Total	4.8	5.5	5.4	6.9	8.2	5.2	5.3	5.3	6.2	3.5	6.7	7.3	5.5	5.4	6.4	5.8	6.3	8.5	6.4	5.7	7.8	5.6	5.5	5.1	6.3	5.3	7.1	8.3	5.9	11.7	5.9	7.1	6.6	

Fuente: Elaboración propia.