



Universitas-XXI, Revista de Ciencias Sociales y Humanas

ISSN: 1390-3837

ISSN: 1390-8634

revistauniversitas@ups.edu.ec

Universidad Politécnica Salesiana

Ecuador

Rojas-Sánchez, Camila; Cárcamo-Ulloa, Luis; Vernier, Matthieu

**Cobertura web de la corrupción en Chile: análisis
automatizado con inteligencia artificial y lingüística**

Universitas-XXI, Revista de Ciencias Sociales y Humanas,
núm. 43, 2025, Septiembre-Febrero 2026, pp. 65-93

Universidad Politécnica Salesiana

Cuenca, Ecuador

DOI: <https://doi.org/10.17163/uni.n43.2025.03>

Disponible en: <https://www.redalyc.org/articulo.oa?id=476182348003>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica Redalyc

Red de revistas científicas de Acceso Abierto diamante

Infraestructura abierta no comercial propiedad de la academia

Cobertura web de la corrupción en Chile: análisis automatizado con inteligencia artificial y lingüística

*Web coverage of corruption in Chile: automated analysis
with artificial intelligence and linguistics*

Camila Rojas-Sánchez

camila.rojas@ufrontera.cl

<https://orcid.org/0009-0009-3677-2268>

Universidad de La Frontera, Temuco, Chile

<https://ror.org/04v0snf24>

Luis Cárcamo-Ulloa

lcarcamo@uach.cl

<https://orcid.org/0000-0003-0633-9606>

Universidad Austral de Chile, Valdivia, Chile

<https://ror.org/029ycp228>

Matthieu Vernier

mvernier@inf.uach.cl

<https://orcid.org/0000-0002-3265-4709>

Universidad Austral de Chile, Valdivia, Chile

<https://ror.org/029ycp228>

Recibido: 12/03/2025 **Revisado:** 30/04/2025 **Aprobado:** 15/06/2025 **Publicado:** 01/09/2025

Cómo citar: Rojas-Sánchez, C., Cárcamo-Ulloa, L. y Vernier, M. (2025). Cobertura web de la corrupción en Chile: análisis automatizado con inteligencia artificial y lingüística. *Universitas XXI*, 43, pp. 65-93, <https://doi.org/10.17163/uni.n43.2025.03>

Resumen

Este estudio busca presentar y validar técnicas de análisis de evaluación de personajes involucrados en casos de corrupción política mediante un clasificador de noticias automatizado que combina inteligencia artificial, lingüística y ciencias sociales y políticas. Para ello, se realiza, con base en estrategias de tratamiento automático del lenguaje, un análisis de tópicos con LDA (Latent Dirichlet Allocation) de un corpus amplio de informaciones extraído de Sophia Search, un clasificador de noticias automatizado, y donde se encuentran 33 699 noticias sobre 07 casos de corrupción considerados escandalosos, provenientes de 64 medios informativos chilenos presentes en la web. Desde este corpus, resulta el análisis de 32 personajes involucrados entre siete hechos considerados como escandalosos políticamente, y que fueron triangulados con 1120 conceptos que ayudaron a ahondar en la evaluación de las y los personajes. Así, los resultados dejan en evidencia diferencias y similitudes en un cruce de un eje entre medios considerados hegemónicos y no hegemónicos, y de otro eje entre medios considerados tradicionales o emergentes, resaltando estrategias propias en cada uno de los tipos de medios. Finalmente se concluye sobre los resultados y se dan futuras proyecciones sobre el uso de herramientas con inteligencia artificial en ciencias sociales y políticas.

Palabras clave

Minería de datos, inteligencia artificial, lingüística, ciencias sociales y políticas, corrupción política, medios informativos, web.

Abstract

This study seeks to present and validate analytical techniques for evaluating figures involved in political corruption cases using an automated news classifier that combines artificial intelligence, linguistics, and social and political sciences. To this end, a topic analysis using LDA (Latent Dirichlet Allocation) is analyzed using automatic language processing strategies. A large corpus of information extracted from Sophia Search, an automated news classifier, contains 33,699 news items about seven corruption cases considered scandalous, from 64 Chilean online news outlets. From this corpus, the analysis of 32 figures involved in seven events considered politically scandalous is obtained. These items were triangulated with 1120 concepts that helped to deepen the evaluation of the figures. Thus, the results reveal differences and similarities across one axis between media considered hegemonic and non-hegemonic, and across another axis between media considered traditional or emerging, highlighting distinct strategies in each media type. Finally, a conclusion is drawn from the results and future projections are made regarding the use of artificial intelligence tools in social and political sciences.

Keywords

Data mining, artificial intelligence, linguistics, social and political sciences, political corruption, news media, web.

Introducción y estado del arte

Desde el retorno a la democracia, Chile ha sido registrado y conocido como uno de los países con mayores índices de transparencia, disputando

el primer lugar latinoamericano con Uruguay (Transparencia Internacional, 2024). Sin embargo, dicho puntaje ha caído significativamente desde 2014, luego de que trascendieran varios casos de corrupción con alto impacto que involucran a figuras e instituciones políticas relevantes del país, poniendo de manifiesto deficiencias estructurales en la prevención y sanción de la corrupción (Wilenman y Feddersen, 2023; Heiss, 2023; Matus, 2020; Nash, 2017; CIPER, 2015a, 2015b).

Si bien, no se puede negar que en las últimas dos décadas Chile ha materializado progresos en lo que dice relación a normas e instituciones ligadas a la transparencia y probidad pública, tales avances, han sido parciales y generalmente impulsados por algún escándalo de corrupción (Nash, 2017; Zalaquett, 2011), lo que ha constituido una baja confianza en instituciones públicas y privadas (Irrarrazabal y Cruz, 2023).

Esta caída en la confianza, en 20 años, no es al azar, ya que, difundidos por los medios de comunicación, los escándalos de corrupción y los privilegios de algunos personajes dentro de la política chilena son evidenciables.

En términos de Thompson (2001), los casos de corrupción son presentados como escándalos políticos y promueven la movilización de la indignación moral y el ejercicio de la denuncia frente a un público (Pereyra, 2012). En este sentido, la corrupción no se puede entender sin tener en cuenta las redes sociales y las plataformas multimediales, pues como señalan diversas investigaciones, el uso de fuentes de información modernas (la web) influye en el interés por la política y específicamente en las temáticas de corrupción (Pastrana, 2019; Nash, 2017), pues, si bien, el mayor consumo de plataformas digitales genera un grado más alto de exposición a contenidos diversos, esto termina por fortalecer convicciones, debido a la capacidad de los algoritmos que ocupan estas plataformas de ofrecer información personalizada y/o ad hoc (Aruguete *et al.*, 2024; Vidal y Olivares, 2020; Waisbord, 2020).

De esta forma, las redes sociales y las aplicaciones tecnológicas van alcanzando la importancia informativa que obtiene la televisión. Así, el WhatsApp, las redes sociales, y la televisión abierta son los medios más frecuentados para informarse por los chilenos y chilenas, según la encuesta CADEM (2025). De esta manera, pareciese que el consumo de medios tradicionales tiene efectos positivos sobre la opinión y participación política, y el uso de redes sociales, en tanto es multidireccional, incentiva otras dimensiones más críticas sobre la participación política (Navia y Ulriksen, 2017). De esta forma, las discusiones políticas en la web estarían reemplazando el uso de los

medios de comunicación tradicionales en los procesos democráticos en América Latina (Rodríguez-Virgili y Fernández, 2017).

En efecto, la digitalización, como señalan Serrano y Salvat (2015), que se inició como un medio de comunicación más, se ha ido ampliando y filtrando por todas las dimensiones de nuestra sociedad, con lo cual ocurre el proceso de migración digital desde la Sociedad de la Información a la actual Sociedad Red, donde la digitalización ha rizomatizado la estructura de la información, dando cuenta de la irrupción de lo digital en la prensa, y donde la Internet es reconocida como un medio particular que se caracteriza por integrar y fusionar en sí a todos los medios hasta ahora conocidos. De hecho, se habla ya de una nueva configuración del sistema de medios en Chile, donde cohabitan las versiones digitales de los medios tradicionales y nuevos medios nacidos para las redes (Vernier *et al.*, 2017), y donde las redes sociales y las aplicaciones tecnológicas van alcanzando la importancia informativa que tenían los medios formales, modificando el sistema informativo chileno (López-Rabadán y Mellado, 2018; Ayala, 2016).

El periodismo está experimentando cambios sustanciales derivados de la evolución tecnológica y de la interacción social que esta evolución posibilita, atravesando una crisis de credibilidad y confianza (Sixto *et al.*, 2021). Por eso, hay quienes plantean que instruir a la ciudadanía consumidora de noticias, desde una perspectiva crítica, es la vía más desafiante para garantizar la calidad de los ecosistemas informativos democráticos (Cárcamo *et al.*, 2021; Cárcamo *et al.*, 2023), frente a este fenómeno donde la ciudadanía parecería vulnerable ante noticias de mala calidad o enmarques poco explícitos (Vernier *et al.*, Scheihing, 2018).

De hecho, en el caso chileno, Mellado y Scherman (2020), realizan un estudio para mapear la diversidad de fuentes y su calidad en las plataformas y medios de noticias en Chile, y según sus resultados, los medios en línea no parecen contribuir a una mayor diversidad en la representación de los personajes sociales, y parecieran tener predilección por informar sobre celebridades y personajes que aparecen en los medios tradicionales, especialmente en televisión. Misma conclusión a la que había llegado Couso (2011), al señalar que, al menos en Chile, y probablemente en otras democracias latinoamericanas también, los nuevos medios de comunicación digital no han transformado radicalmente el escenario de configuración de la agenda pública que producen los medios tradicionales.

En este sentido, este tipo de tratamiento informativo da cuenta de un sesgo típico de las coberturas mediáticas sensacionalistas: la personalización de la noticia, la cual expresa la tendencia mediática a focalizar su atención en los actores más que en sus acciones, incrementando la atención del público a partir del despliegue de elementos afectivos explícitos dramáticos y controversiales (Zunino y Focás, 2021; Palau y Davesa, 2013). De allí que la corrupción puede ser entendida como un fenómeno de corte interpretativo (Sola-Morales y Rivera, 2017).

Por ello es que esta investigación se plantea como objetivo presentar y validar técnicas de análisis de evaluación de personajes involucrados en casos de corrupción política mediante un clasificador de noticias automatizado que combina inteligencia artificial, lingüística y ciencias sociales. Para ello, en este estudio se analiza, con base en estrategias de tratamiento automático del lenguaje, un análisis de tópicos con LDA (Latent Dirichlet Allocation) de un corpus amplio de informaciones extraído de Sophia Search, un clasificador de noticias automatizado, y donde se encuentran 33 699 noticias sobre 07 casos de corrupción considerados escandalosos, provenientes de medios de todo Chile.

En consecuencia, a ello, se procesan todas las noticias para representarlas como una lista de sustantivos, conceptos claves, y entidades, y se agregan patrones para buscar sus frecuencias.

Además, debido a que se obtienen frecuencias de personajes y conceptos, se realiza un análisis de sentimientos dentro del procesamiento del lenguaje natural (NLP), aplicándolo a evaluaciones morales de las noticias sobre casos de corrupción. De esta manera, se hizo un análisis de sentimientos o también llamado de polaridad, obteniendo diversas puntuaciones, con un valor entre -1 y 1 que viene a indicar de una única vez si el modelo es positivo o negativo. Valores próximos a -1 indican que es muy negativo, próximos a cero indicarían que es neutro y próximos a 1 sería muy positivo.

Luego se genera una triangulación de los datos arrojados, tanto por el análisis de frecuencia y polaridad o sentimientos de los personajes vinculados a los casos de corrupción, como por el análisis de frecuencia conceptual de las noticias relacionadas a los mismos.

Al respecto, hay que tener en consideración que la investigación centrada en la construcción de modelos supervisados de aprendizaje automático para clasificar textos en este idioma todavía es escasa, pero va en aumento (Sánchez *et al.*, 2020), y la mayoría de estudios que clasifican textos en español

tienen su origen en el proyecto TASS (Taller de Análisis de Sentimientos en la Sociedad Española para el Procesado del Lenguaje Natural - SEPLN), que tiene como finalidad promover la investigación y el desarrollo de nuevos algoritmos y técnicas, y que se aplica al análisis de opiniones en español (García Cumbrreras *et al.*, 2016).

De hecho, hay investigaciones que aplican análisis de sentimientos para indagar en las evaluaciones morales y posibles sesgos que podrían estar afectando en los textos de las noticias en la web, tarea desafiante, debido a las dificultades de procesar algunos fenómenos lingüísticos subyacentes como la negación, la ironía, el sarcasmo, la ambigüedad del lenguaje natural (Xu *et al.*, 2024, Sutriawan *et al.*, 2024), y más ampliamente, la subjetividad de las opiniones (García-Vega *et al.*, 2020), que exigen contexto al análisis. De hecho, la lingüística presta atención a la innovación tecnológica en tanto tiene repercusión en la lengua hispanohablante y sus usos (Lovón y Parini, 2024).

Estos nuevos enfoques que aúnan métodos computacionales y big data se empiezan a abordar en la investigación de la comunicación en las ciencias sociales y políticas (Arcila-Calderón *et al.*, 2016), y abren nuevas perspectivas para su aplicación en otros sectores como el de la comunicación política.

Esta investigación culmina con las conclusiones generales del estudio, señalando la experiencia del uso de una herramienta que combina inteligencia artificial con las disciplinas académicas de la lingüística y las ciencias sociales y políticas, e indicando proyecciones de futuros estudios que seguirán permitiendo ahondar en el análisis del pluralismo mediático en Chile.

Material y métodos

Debido al gran volumen de información publicada en línea, la carga recae en los enfoques automatizados para identificar estos marcos dentro del texto. Morstatter *et al.* (2018) prueban un modelo de aprendizaje automatizado (*machine learning*), que permite reconocer patrones e identificar sesgos en los encuadres periodísticos. Asimismo, Saez-Trumper *et al.* (2013) examinan los sesgos en fuentes de noticias en línea y las comunidades que las rodean. Para ello, presentan métodos no supervisados que consideran tres tipos de sesgos: sesgo de selección o “control”, sesgo de cobertura y sesgo de declaración, caracterizando cada uno de ellos a través de una serie de métricas, identificando que dichos sesgos se incrementan en plataformas vinculadas

a redes sociales. También, hay investigaciones que señalan posibles sesgos discriminatorios en los algoritmos que permiten estos análisis (O’Neil, 2017; Martínez *et al.*, 2020).

Este tipo de enfoques metodológicos tiene la dificultad de advertir fenómenos lingüísticos como la ironía, el sarcasmo, y la ambigüedad inherente al lenguaje natural (Xu *et al.*, 2024, Sutriawan *et al.*, 2024), por ello, los enormes volúmenes de textos requieren de herramientas capaces de procesar automáticamente estos mensajes sin perder fiabilidad. Para ello, se pueden utilizar procesos de aprendizaje automático supervisado (*machine learning*), lo que permite tener en cuenta el contexto, mostrando patrones, pero siempre hay que darles chance a las personas investigadoras para afinar el análisis (Baviera, 2017), pues los métodos automáticos de análisis no pueden prescindir de una exigente codificación manual si se desea utilizarlos con fiabilidad en la investigación.

A continuación, se presenta el diseño metodológico de esta investigación, haciendo énfasis en la búsqueda de noticias mediante el software Sophia (Vernier y Ruiz, 2023) y el análisis mediante el lenguaje de programación de código abierto debido al alto volumen de datos (Oliva, 2017), y la triangulación con datos codificados por el equipo de investigación (Otero, 2020).

Diseño General de Investigación

El presente estudio es de carácter mixto, con una metodología cuantitativa y una cualitativa, ya que se busca acceder y rastrear grandes flujos de noticias de distintos medios con presencia en la web sobre casos de corrupción política, y para ello se seguirá un modelo de análisis basado en minería de datos en la web, específicamente bajo un conjunto de medios de comunicación que el software Sophia indexa. Asimismo, se busca acceder a procesos interpretativos sobre el tratamiento del tema de la corrupción política en distintos medios de comunicación de Chile, ya sean a nivel nacional, como a nivel local.

Para explorar el tratamiento noticioso de distintos temas en todas sus etapas comunicacionales, se necesita de la combinación de técnicas cualitativas y cuantitativas (Amadeo, 2008). Así, el análisis de contenido se ofrece como una técnica adecuada y válida para la recolección y el procesamiento de grandes cantidades de datos y la consecuente identificación de los tópicos presentes en las noticias (Sádaba *et al.*, 2012; Koziner, 2015).

Diseño muestral y estrategia de muestreo

El diseño muestral se acogió a una lógica comprensiva que dependía de los avances que se iban alcanzando en el entendimiento del problema estudiado durante el progreso mismo de la investigación. Así, la búsqueda se realizó bajo la plataforma Sophia, herramienta de recopilación de grandes volúmenes de datos de noticias de prensa, lo que permite explorar fenómenos sociales a partir del uso de algoritmos de *machine learning* y de lingüística computacional, desarrollando así metodologías computacionales de análisis del lenguaje humano aplicadas a las ciencias sociales y políticas, utilizando inteligencia artificial y algoritmos.

Sophia es un software de recopilación automática e indexación de documentos web que elabora métodos computacionales e indicadores construidos a partir de la recopilación de grandes volúmenes de datos de noticias de prensa en varios países del mundo, utilizando algoritmos de *machine learning* y de lingüística computacional y desarrollando metodologías computacionales de análisis del lenguaje humano aplicadas a las ciencias sociales.

El conjunto de medios seguidos es un catastro heterogéneo en sus orientaciones políticas, geográficamente distribuidos por todo el territorio nacional y con orígenes mediales diversos, y el periodo de análisis tiene relación con la creación de hitos de los escándalos de corrupción, por tanto, se analizan todos aquellos textos relacionados con los casos de corrupción considerados escandalosos entre 2014 y 2022.

Cobra relevancia entender cómo funciona la plataforma Sophia. Sophia Search es un motor de búsqueda que indexa millones de noticias de prensa publicadas por los medios de prensa (Verniery Ruiz, 2023). A partir de una lista de palabras claves, un rango de fecha y una lista de medios de prensa, devuelve las noticias más relevantes, indicando en particular su fecha de publicación, el medio de prensa que la publicó, el título y contenido de la noticia y su URL.

Plan de análisis

Se almacenaron noticias emitidas sobre distintos casos de corrupción considerados escandalosos. Para ello, en el buscador de Sophia se hizo la siguiente búsqueda, según siete casos de importancia mediática, filtrando por fechas entre 31 de diciembre del 2013 y el 01 de enero del 2023.

Caso Corpesca: ("Corpesca" OR "Jaime Orpís" OR "Marta Isasi" OR "Ximena Chong" OR "Angelini" OR "Ley General de Pesca" OR "Ley de Pesca") AND ("inhabilitación" OR "asesorías" OR "soborno" OR "cohecho" OR "corrupción" OR "caso" OR "infringir" OR "multa" OR "investigación" OR "multas" OR "acusados" OR "acusado" OR "infracciones" OR "infracción" OR "sanción" OR "sanciones" OR "sanciones")

Caso Cascadas: "cascadas" OR "sqm" OR "Soquimich" OR "Sociedad Química y Minera de Chile" OR "Calichera" OR "Oro blanco" OR "Superintendencia de Valores y Seguros" OR "Ponce Lerou" AND ("corrupción" OR "caso" OR "infringir" OR "multa" OR "investigación" OR "multas" OR "acusados" OR "acusado" OR "infracciones" OR "infracción" OR "sanción" OR "sanciones" OR "corte suprema")

Caso Penta: ("Penta" OR "Pentagate" OR "FUT" OR "Fondo de Utilidades Tributables" OR "Carlos Alberto Délano" OR "Carlos Délano" OR "Délano" OR "Carlos Eugenio Lavín" OR "Carlos Lavín" OR "Hugo Bravo" OR "Jovino Novoa") AND ("corrupción" OR "caso" OR "infringir" OR "multa" OR "investigación" OR "multas" OR "acusados" OR "acusado" OR "infracciones" OR "infracción" OR "sanción" OR "sanciones" OR "soborno" OR "cohecho")

Caso Caval: ("Caval" OR "Sebastián Dávalos" OR "Dávalos" OR "Machali" OR "Natalia Compagnon" OR "Compagnon" OR "Nueragate" OR "Sociocultural de la Presidencia") AND ("hijo" OR "Michelle Bachelet" OR "crédito" OR "Banco de Chile" OR "Banco Chile" OR "madre" OR "predios" OR "plano regulador" OR "suelo" OR "uso urbano" OR "tráfico de influencias" OR "corrupción" OR "caso" OR "infringir" OR "multa" OR "investigación" OR "multas" OR "acusados" OR "acusado" OR "acusadas" OR "acusada" OR "infracciones" OR "infracción" OR "sanción" OR "sanciones")

Caso Milicogate: ("Milicogate" OR "Ley reservada del cobre" OR "ejército" OR "Corte Marcial" OR "militar" OR "Comando de Apoyo a la Fuerza" OR "defensa" OR "Milico" OR "Milicos" OR "Juan Miguel Fuente-Alba" OR "Juan Miguel Fuente Alba" OR "Juan Miguel Alba" OR "Héctor Ureta" OR "Clovis Montero" OR "Francisco Huinchahue") AND ("Codelco" OR "Cobre" OR "datos borrados" OR "uniformado" OR "uniformados" OR "corrupción" OR "caso" OR "infringir" OR "multa" OR "investigación" OR "multas" OR "acusados" OR "acusado" OR "infracciones" OR "infracción" OR "sanción" OR "sanciones" OR "soborno" OR "cohecho")

Caso Pacogate: ("Pacogate" OR "Paco" OR "Carabineros" OR "fuerza policial" OR "Eduardo Gordon" OR "Flavio Echeverría" OR "Felipe Ávila") AND ("documentos faltantes" OR "documentos destruidos" OR "corrupción" OR "caso" OR "infringir" OR "multa" OR "investigación" OR "multas" OR "acusados" OR "acusado" OR "infracciones" OR "infracción" OR "sanción" OR "sanciones" OR "soborno" OR "cohecho")

Caso Rodrigo Rojas Vade: ("Rodrigo Rojas Vade" OR "Rojas Vade" OR "pelado Vade") AND ("cáncer" OR "convención" OR "constituyente" OR "mentira" OR "enferno" OR "donaciones" OR "tratamiento" OR "exconvencional" OR "pelo" OR "pelado" OR "cejas" OR "estallido social" OR "leucemia" OR "desnudos" OR "protestas" OR "salud" OR "vicepresidencia" OR "sífilis" OR "corrupción" OR "caso" OR "infringir" OR "multa" OR "investigación" OR "multas" OR "acusados" OR "acusado" OR "infracciones" OR "infracción" OR "sanción" OR "sanciones")

Luego se unieron estos siete outputs de Sophia Research en un documento CSV, para ser procesado y analizado mediante lenguaje Python, documento al que se le agregan algunas variables, con el fin de hacer una caracterización general de los medios que hablan de casos de corrupción en Chile.

Luego, para poder identificar y comparar el tratamiento de la cobertura periodística de los escándalos de corrupción chilenos ocurridos entre 2014 y 2022, se procesaron todas las noticias para representarlas como una lista de sustantivos como (NOUN); conceptos claves (PATTERN); entidades como (PER) y (ORG), y se agregaron patrones para buscar sus frecuencias.

Además, debido a que se analizaron frecuencias de personajes y conceptos, se puede analizar evaluaciones morales de las noticias sobre casos de corrupción. De esta manera, se hizo un análisis de sentimientos o también llamado de polaridad, obteniendo diversas puntuaciones, con un valor entre -1 y 1 que viene a indicar de una única vez si el modelo es positivo o negativo. Valores próximos a -1 indican que es muy negativo, próximos a cero indicarían que es neutro y próximos a 1 sería muy positivo.

El análisis de sentimientos es un caso de uso del procesamiento del lenguaje natural (NLP) y se incluye en la categoría de clasificación de texto, que implica clasificar un texto en varios sentimientos, como positivo, negativo o neutral. Por lo tanto, el objetivo final del análisis de sentimientos es descifrar el sentimiento subyacente de un texto. Esto también se conoce como Minería de Opinión.

De esta manera, resulta un análisis de 32 personajes involucrados entre los 07 hechos considerados como escandalosos políticamente, y que fueron triangulados con 1120 conceptos que ayudaron a ahondar en la evaluación de los personajes.

Análisis y resultados

El análisis que se presenta a continuación fue elaborado para el estudio de un conjunto de noticias del ámbito nacional sobre hechos de corrupción considerados escandalosos en Chile, entre 2014 y 2022, como lo son: caso Cascadas o también conocido como caso SQM, caso Penta, caso Caval, Milicogate, Pacogate, caso Corpesca, y caso Rodrigo Rojas Vade.

En total, el análisis se aplicó a un conjunto de 33 699 noticias, distribuidas en 64 medios de todo Chile, presentes en la web (tabla 1).

Tabla 1

*Cobertura noticiosa de casos de corrupción chileno
por medio presente en la web entre los años 2014 y 2022*

	Medio	Posts		Medio	Posts
1	emol	4464	33	diariodevaldivia	15
2	horas24	3413	34	lavozdepaillaco	11
3	113	3289	35	noticiaslosrios	10
4	elciudadano	2940	36	clave9	10
5	elllanquihue	2822	37	enlanoticia	9
6	cnnchile	2586	38	enlalinea	8
7	latercera	2562	39	elproa	8
8	elaustral	2154	40	aconcaguadigital	7

9	ahoranoticiasmega	1617	41	eltrabajo	6
10	laestrelladechiloe	1242	42	diariofutro	5
11	biobiochile	1145	43	radioNuevoamanecer	4
12	chvnoticias	992	44	radioestrelladelmar	4
13	ciper	913	45	masnoticia	4
14	eldinamo	630	46	seminariolocal	3
15	elmostrador	601	47	segundos33	3
16	elvacanudo	544	48	prensaDeleuario	3
17	elrepuertero	507	49	elinsular	3
18	lanacion	271	50	elcalbucano	3
19	radiosago	247	51	diariolagoranco	3
20	laquinta	125	52	soypuertomontt	2
21	lasereñaonline	96	53	rioenlinea	2
22	elnavegable	90	54	radioacogida	2
23	laregionhoy	60	55	fresiaahora	2
24	radiopudeto	49	56	diariosanjose	2
25	diariolaguino	35	57	diariolaunion	2
26	elheraldoaustral	30	58	soyosorno	1
27	suractual	29	59	soychiloe	1
28	elquellonino	25	60	elinformador	1
29	ovallehoy	24	61	diarioriobueno	1
30	elhuemul	21	62	diariopaillico	1
31	elperiodista	19	63	diariomafil	1
32	elobservador	19	64	diariolanco	1

Asimismo, para saber qué variables son interesantes de analizar, es que se presenta una panorámica general de la cobertura mediática de los hechos de corrupción en Chile. De esta manera, se puede ver la distribución geográfica de los medios que emitieron al menos una noticia sobre algún escándalo

de corrupción (ilustración 1). En este sentido, es importante señalar que estas 33 699 noticias que están distribuidas en 64 medios están ubicados geográficamente en 39 comunas a lo largo de Chile, y como en la mayoría de la cobertura noticiosa en Chile, esta está concentrada en una mayor proporción en la Región Metropolitana, y en esta ocasión, con un poco más del 70 % de las emisiones de noticias en dicha Región, e inclusive, únicamente en la comuna de Santiago se concentra casi el 25 % de todas las noticias emitidas en el periodo de tiempo que concentra los años 2014 y 2022.

Ilustración 1

Distribución geográfica de los medios que cubren hechos de corrupción entre los años 2014 y 2022

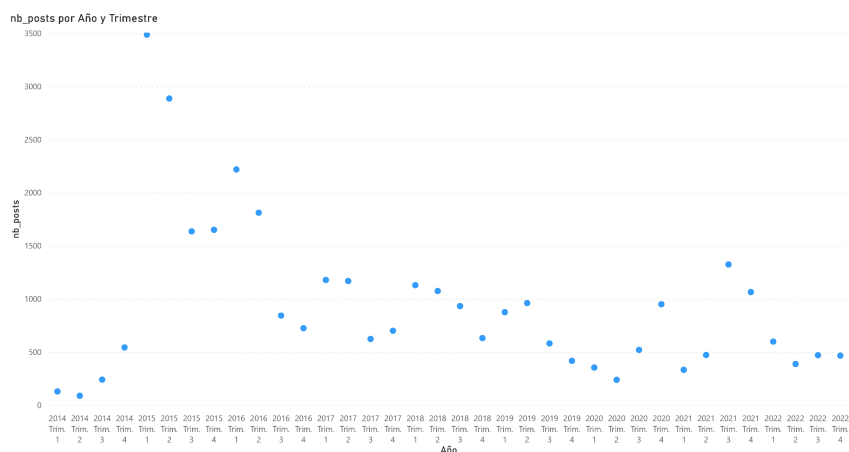


También se puede ver una panorámica a lo largo del tiempo de las emisiones de las noticias de corrupción entre 2014 y 2022 (ilustración 2). Allí se puede observar que el *peak* más alto de post de noticias sobre hechos de corrupción coincide con el comienzo del segundo gobierno de Michelle Bachelet, en el primer trimestre del 2015, donde la temática del financiamiento ilegal emerge a la luz pública, y luego el segundo *peak* coincide con el caso Rodrigo Rojas Vade, en septiembre de 2021, cuando luego de una entrevista a La Tercera, el exconvencional Rodrigo Rojas admitió no padecer de cáncer, pese a que obtuvo donaciones para el supuesto tratamiento.

Llama la atención aquí, que entremedio de estos *peak*, se mantiene de forma relativamente estable la cantidad de post sobre hechos de corrupción en Chile, y eso no quiere decir que no hayan existido otros casos de corrupción.

Ilustración 2

Distribución de la cantidad de noticias sobre casos de corrupción en Chile entre los años 2014 y 2022



Asimismo, para poder generar comparaciones de tipos de medios en la web se pueden buscar características de cada medio, tales como, año de origen y personalidad jurídica, etc., con el fin de poder caracterizar a los medios.

En relación con el año de origen del medio, tal como establece Couso (2011), en Chile, hay un alto grado de digitalización, y desde su instalación en el país, comienzan a surgir “nuevos medios”, donde encontramos (1) me-

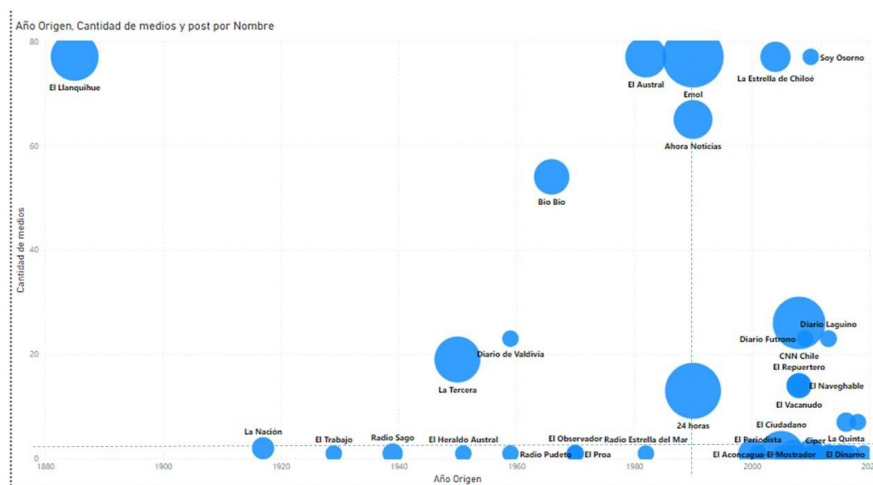
dios tradicionales, incluyendo sus versiones digitales, que son creados antes de 1990 y la llegada profesional de internet en el país; y (2) medios emergentes, incluyendo sus versiones digitales, que son creados antes de 1990 y la llegada profesional de internet en el país.

Referente a la personalidad jurídica de cada medio, con esta información, se puede categorizar a aquellos propietarios que tengan más de un medio como (1) hegemónicos, y aquellos que tengan solo un medio como (2) no hegemónicos.

De esta manera, podemos categorizar los medios que cubren noticias de corrupción, entre medios hegemónicos tradicionales (HT), medios no hegemónicos tradicionales (HNT), medios hegemónicos emergentes (HE) y medios no hegemónicos emergentes (NHE), como se puede ver en la ilustración 3.

Ilustración 3

Distribución de la cobertura de noticias de corrupción, durante los años 2014 a 2022, en medios presentes en la web, según año de creación y según cantidad de medios de su propietario



De esta forma, y como se puede ver también en la tabla 2, se evidencia de que, bajo esta primera caracterización por medios que cubren noticias de corrupción, los medios que postean más noticias sobre hechos noticiosos considerados como escándalos políticos, en forma decreciente, son los he-

gemónicos emergentes (HE), que son un 39,1 % de los medios analizados y cubren un 46,5 % de las noticias bajo sus post; luego están los Hegemónicos Tradicionales (HT), que son solo un 9,4 % de los medios analizados y cubren un 36,3 % de las noticias bajo sus post; después se encuentran los No Hegemónicos Emergentes (NHE), que son un 40,6 % de los medios analizados y cubren un 16,1 % de las noticias bajo sus post; y por último están los No Hegemónicos Tradicionales (NHT), que son un 10,9 % de los medios analizados y cubren solo un 1,1 % de las noticias bajo sus post.

Tabla 2

Porcentajes de medios y su cobertura, según categorización

	Post	Medios	%Post_del_Total	% Medios_del_Total
HE	15 654	25	46,5 %	39,1 %
HT	12 243	6	36,3 %	9,4 %
NHE	5439	26	16,1 %	40,6 %
NHT	363	7	1,1 %	10,9 %
	33 699	64	100 %	100 %

Como se puede ver en la ilustración 3 y la tabla 2, aquellos medios emergentes, es decir que sus orígenes radican desde 1990 en adelante, son el 79,7 %, casi un 80 %, de los medios analizados. Y si se ve desde términos hegemónicos, es decir, si la personalidad jurídica de dichos medios tiene uno o más medios bajo su poder, se percibe que un 82,8 % de los posts que cubren noticias relacionadas a casos de corrupción provienen de medios hegemónicos. Es decir, que pese a la alta emergencia de medios que permite la digitalización, el volumen de emisión en la cobertura sigue estando en manos de medios hegemónicos, sean estos emergentes o tradicionales.

Ahora, si se desea analizar sobre quiénes, de qué, y cómo se habla de los casos de corrupción, además de comparar los distintos tratamientos de los escándalos de corrupción, entonces es conveniente hacer los análisis por caso de corrupción. Para ello, se tuvo en consideración la cobertura por casos, los que, como se puede evidenciar en las tablas 3.1 y 3.2, varían según el caso noticioso.

Tabla 3.1

Porcentaje de cobertura de casos de corrupción por tipo de medio

	%Post_ Corpes ca	%Post_ Casca das	%Post_ Pentag ate	%Post_ Caval	%Post_ Milicog ate	%Post_ Pacog ate	%Post_ RojasV ade	%Post_ del Total
HT	30,31 %	40,90 %	34,25 %	39,54 %	36,75 %	33,56 %	44,68 %	36,30 %
HE	42,90 %	41,76 %	47,39 %	49,67 %	40,47 %	45,52 %	41,88 %	46,50 %
NHT	2,15 %	0,78 %	0,49 %	0,55 %	0,27 %	2,11 %	1,77 %	1,10 %
NHE	24,65 %	16,56 %	17,87 %	10,25 %	22,51 %	18,82 %	11,66 %	16,10 %
	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %	100,00 %

Tabla 3.2

Porcentaje de cobertura de casos de corrupción entre tipo de medio

	%Post_ del Tot al_TM_ Corpes ca	%Post_ del Tot al_TM_ Casca das	%Post_ del_T ot al_TM_ Pentag ate	%Post_ del Tot al_TM_ Caval	%Post_ del Tot al_TM_ Milicog ate	%Post_ del Tot al_TM_ Pacoga te	%Post_ del Tot al_TM_ RojasV ade	Total
HT	14,35 %	27,96 %	19,56 %	19,21 %	4,24 %	10,94 %	3,74 %	100,00 %
HE	16,50 %	23,19 %	21,99 %	19,61 %	3,79 %	12,06 %	2,85 %	100,00 %
NHT	34,30 %	17,94 %	9,50 %	8,97 %	1,06 %	23,22 %	5,01 %	100,00 %
NHE	24,37 %	23,64 %	21,32 %	10,40 %	5,42 %	12,81 %	2,04 %	100,00 %
	22,38 %	23,18 %	18,09 %	14,55 %	3,63 %	14,76 %	3,41 %	100,00 %

A continuación, se presenta una síntesis de los resultados que profundizan el análisis que se hizo de 32 personajes involucrados entre los siete hechos considerados como escandalosos políticamente, y que fueron triangulados con los 1120 conceptos que ayudaron a ahondar en la evaluación de los personajes.

Es importante aclarar que, al ser una investigación que dice relación con casos de corrupción, entonces las evaluaciones de polaridad o de sentimiento, se inclinan a puntuaciones negativas (García-Vega *et al.*, 2020; Martínez *et al.*, Sustaeta, 2020), debido a que dichos personajes, incluso fiscales o denunciantes de los hechos, están relacionados a conceptos ligados a los delitos mencionados en las noticias, y los que, por su carga moral, darán promedios negativos.

De esta manera, mediante una escala de color de los porcentajes de cobertura, como de la polaridad promedio que puntuaron los personajes, se puede señalar que el personaje con mayor cobertura dentro de los medios, y pese a que es está involucrado en uno de los casos con menor cantidad de noticias, es Rodrigo Rojas Vade. Esto puede responder a que, a diferencia

de los otros casos, a quien se acusa de cometer el delito es a una única persona, lo cual hace que el interés por cubrir el hecho sea únicamente dirigido hacia dicha persona, gatillando la pérdida de credibilidad del primer proceso constitucional.

Algo similar ocurre en el caso Milicogate, en donde el ex comandante Fuente-Alba se encuentra procesado e investigado por determinación de la ministra en visita Romy Rutherford, ambos con altos porcentajes de cobertura y de polaridad promedio negativa. Aunque, los otros acusados por la ministra en visita, a pesar de no tener una alta cobertura, sí tienen altas puntuaciones negativas en la escala de sentimientos, llamando la atención que, si bien son evaluados de forma negativa en la mayoría de los tipos de medios, al ser imputados, no lo están por sobre Rutherford.

Lo mismo sucede con la fiscal Ximena Chong en el caso Corpesca, teniendo siempre una polaridad promedio más negativa que Jaime Orpis, siendo él una de las personas involucradas en el caso y condenado por ello.

Al respecto, en el caso Cascadas, llama la atención la baja cobertura hacia el personaje de Julio Ponce Lerou, principal benefactor de las compraventas de las sociedades cascadas de SQM, cobertura que no alcanza el 7 % en varios de los tipos de medios, con excepción a los medios no hegemónicos emergentes, quienes lo mencionan en casi un 19 % de las noticias vinculadas al caso.

Por el contrario, la cobertura de Natalia Compagnon en el caso Caval es alta, cercana al 30 % y el 40 % según tipo de medio, y con ello, también lo es su alta polaridad negativa en casi todos los tipos de medios.

Con respecto al caso Penta, llama la atención la alta cobertura a la Unión Demócrata Independiente (UDI), la única organización evaluada en esta investigación como personaje, obteniendo la mayor polaridad negativa en casi todos los tipos de medios, a excepción de los hegemónicos emergentes en donde Sebastián Piñera alcanza ese lugar.

En relación con el caso Pacogate, resalta la puntuación negativa en la escala de polaridad del expolicía Bruno Villalobos, incluso por sobre Eduardo Gordon, quien sí estuvo condenado por el delito de malversación de caudales públicos, quedando con arraigo nacional y firma mensual.

En todos los casos, se evidencia la personalización de la noticia, focalizando la atención en los actores más que en sus acciones, con elementos afectivos explícitos (Zunino y Focás, 2021; Palau y Davesa, 2013). Las tablas 4 y 5 resumen estadísticamente todas estas evaluaciones.

Tabla 4

Porcentaje de cobertura y evaluación de personajes de siete casos de corrupción, por tipo de medios

Caso de corrupción	Personaje vinculado al caso	Involucramiento	Porcentaje de mención				Evaluación moral			
			HT	HE	NHT	NHE	HT	HE	NHT	NHE
Corpesca	Ximena Chorra	Directo (fiscal)	34,04 %	24,51 %	9,23 %	18,82 %	-0,0812	-0,0550	0,0077	-0,0589
	Jaime Orpis	Directo (condenado)	2576 %	23,16 %	3,08 %	19,09 %	-0,0501	-0,0431	0,0077	0,0422
	Sebastián Pinera	Indirecto	5,28 %	6,85 %	2,31 %	18,22 %	-0,0169	-0,0154	0,0077	0,0683
Cascadas	Patricio Contesse	Directo (condenado)	19,23 %	12,95 %	4,41 %	13,47 %	-0,0291	-0,0296	-0,0147	-0,0366
	Pablo Lonauera	Directo (acusado)	15,99 %	11,61 %	5,88 %	16,37 %	-0,0285	-0,0260	-0,0294	0,0421
	UDI	Indirecto	7,69 %	6,49 %	4,41 %	16,92 %	-0,0266	-0,0173	-0,0147	-0,0539
	Michelle Bachelet	Indirecto	6,68 %	6,02 %	7,35 %	16,64 %	-0,0154	-0,0156	0,0000	0,0539
	Julio Ponce Lerou	Directo (condenado)	6,85 %	6,24 %	1,47 %	18,44 %	-0,0212	-0,0205	0,0000	0,0698
Pentagate	UDI	Indirecto	17,50 %	14,03 %	19,44 %	25,04 %	-0,0551	-0,0378	0,0000	0,0904
	Carlos Gaiardo	Directo (fiscal)	14,34 %	12,50 %	8,33 %	11,72 %	-0,0388	-0,0303	-0,0556	0,0375
	Jovino Novoa	Directo (condenado)	13,78 %	14,61 %	2,78 %	15,77 %	-0,0368	-0,0352	0,0000	-0,0574
	Carlos Delano	Directo (condenado)	12,66 %	14,73 %	13,89 %	21,36 %	-0,0244	-0,0358	-0,0556	0,0804
	Carlos Lavin	Directo (condenado)	8,75 %	13,66 %	5,56 %	12,94 %	-0,0160	-0,0361	0,0000	-0,0436
	Ivan Moreira	Directo (desaforado)	9,67 %	11,49 %	30,56 %	11,79 %	-0,0248	-0,0312	-0,0278	0,0421
	Sebastian Pinera	Indirecto	9,55 %	8,55 %	19,44 %	29,48 %	-0,0324	-0,0193	-0,0278	-0,1072
Caval	Natalia Compao non	Directo (condenada)	41,05 %	40,19 %	29,41 %	34,85 %	-0,1090	-0,1052	0,1176	0,1068
	Michelle Bachelet	Indirecto	36,13 %	30,31 %	23,53 %	30,93 %	-0,0993	-0,0748	-0,0882	-0,1020
	Sebastián Davalos	Directo (acusado)	27,26 %	34,62 %	23,53 %	34,54 %	-0,0765	-0,0965	0,0000	0,0942
	Mauricio Valero	Directo (condenado)	16,56 %	15,09 %	8,82 %	11,77 %	-0,0382	-0,0405	0,0294	0,0283
	Sergio Moya	Directo (fiscal)	13,10 %	8,42 %	5,88 %	7,69 %	-0,0325	-0,0207	0,0294	-0,0251
Milicogate	Romv Rutherford	Directo (fiscal)	52,58 %	31,16 %	75,00 %	51,81 %	-0,1642	-0,0988	0,0000	0,1717
	Juan Miguel Fuente-Alba	Directo (condenado)	62,18 %	61,81 %	25,00 %	57,83 %	-0,1956	-0,1508	0,0000	-0,1687
	Humberto Oviedo	Directo (acusado)	29,89 %	19,10 %	25,00 %	21,39 %	-0,0941	-0,0653	0,0000	0,0633
	Ricardo Martínez	Directo (acusado)	11,81 %	5,03 %	25,00 %	15,96 %	-0,0277	-0,0184	0,0000	-0,0542
Pacogate	Eduardo Gordon	Directo (condenado)	7,00 %	4,48 %	4,55 %	6,62 %	-0,0207	-0,0111	0,0000	-0,0217
	Bruno Villalobos	Directo (acusado)	18,86 %	17,85 %	6,82 %	12,23 %	-0,0521	-0,0585	-0,0114	0,0395
	Michelle Bachelet	Indirecto	5,57 %	5,85 %	10,23 %	8,41 %	-0,0157	-0,0211	0,0341	-0,0293
	Sebastián Pinera	Indirecto	8,00 %	2,53 %	14,77 %	13,25 %	-0,0271	-0,0068	0,0114	0,0535
Rojas Wade	Rodrigo Roias Wade	Directo (condenado)	74,74 %	74,39 %	78,95 %	76,80 %	-0,2484	-0,1826	-0,1053	-0,2160
	Elisa Loncon	Directo (denunciante)	24,43 %	17,15 %	21,05 %	15,20 %	-0,0605	-0,0178	0,0000	0,0160
	Jaime Bassa	Directo (denunciante)	22,13 %	16,70 %	5,26 %	16,80 %	-0,0689	-0,0245	0,0000	0,0080
	Tomás Ramírez	Directo (abogado)	5,43 %	6,90 %	5,26 %	8,00 %	-0,0146	-0,0111	0,0000	-0,0320

Tabla 5
Conceptos relacionados a siete casos de corrupción, por tipo de medios

Caso de corrupción	Sustantivos más frecuentes del caso (en orden decreciente)				Grupos nominales más frecuentes del caso (en orden decreciente)			
	HT	HE	NHT	NHE	HT	HE	NHT	NHE
Corpesca	caso; fiscal; senador delitos; ley; años; investigación; empresa; cohecho; millones;	caso; fiscal; ley; senador años; investigación; país; pesca; millones; recursos;	pesca; recursos; caso; año; personas; años;	millones; empresas; años; empresa; ley; caso; proyecto; año; país; investigación;	fiscal nacional; delitos tributarios; fiscal regional; prisión preventiva; juicio oral; millones de pesos; pescadores artesanales; financiamiento irregular; arresto domiciliario; medida cautelar;	pescadores artesanales; delitos tributarios; pesca artesanal; prisión preventiva; fiscal nacional; millones de pesos; ley de pesca; proyecto de ley; fiscal regional; recursos pesqueros;	pescadores artesanales; pesca artesanal; pesca recreativa; recursos hidrobiológicos; pesca ilegal; origen legal; ley de pesca; recursos pesqueros; persecución penal; tulla mínima;	proyecto de ley; pescadores artesanales; millones de pesos; pesca artesanal; gerente general; millones de dólares; grupos económicos; financiamiento ilegal; recursos pesqueros; delitos tributarios;
Cascadas	caso; fiscal; millones; investigación; senador años; personas; delitos; empresa; ministro;	caso; fiscal; investigación; empresa; millones; senador delitos; años; ministro;	años; caso; año; información; sector; horas; personas; comuna; millones; país;	millones; empresa; años; caso; años; año; investigación; sociedad; presidente; país;	fiscal nacional; delitos tributarios; juicio oral; fiscal regional; gerente general; financiamiento irregular; candidato presidencial; millones de pesos; sobsecimiento definitivo; campañas políticas;	delitos tributarios; fiscal nacional; gerente general; campañas políticas; financiamiento irregular; millones de pesos; fiscal regional; candidato presidencial; juicio oral; audiencia de formalización;	año pasado; fiscal regional; delitos tributarios; millones de pesos; nivel nacional; fin de semana; juicio oral; redes sociales; fiscal nacional; financiamiento irregular;	gerente general; delitos tributarios; financiamiento ilegal; millones de dólares; millones de pesos; boletas falsas; campañas políticas; fiscal nacional; año pasado; financiamiento irregular;
Pentagate	caso; fiscal; senador delitos; investigación; millones; diputado; años; ministro; presidente;	investigación; fiscal; senador; delitos; años; ministro; presidente; financiamiento;	caso; días; investigación; año; personas; fiscal; millones; parlamentario; delitos;	millones; caso; empresa; años; investigación; proyecto; sociedad; año; presidente;	fiscal nacional; delitos tributarios; financiamiento irregular; fiscal regional; campañas políticas; arresto domiciliario; general; medidas cautelares; gerente general; acusación constitucional;	delitos tributarios; campañas políticas; fiscal nacional; prisión preventiva; financiamiento irregular; arresto domiciliario; medidas cautelares; fiscal regional; arraigo nacional;	año pasado; delitos tributarios; de queja, prisión preventiva; millones de pesos; financiamiento irregular; campañas políticas; facilitación de boletas; acusación constitucional; fiscal nacional;	millones de dólares; gerente general; delitos tributarios; millones de pesos; financiamiento ilegal; campañas políticas; boletas falsas; paraísos fiscales; aportes reservados; proyecto de

Caval	caso; fiscal; ministro; millones; hijo; empresa; investigación; personas; diputado; país;	caso; fiscal; investigador; millones; empresa; hijo; información; país; ministro;	caso; años; millones; personas; país; año; semana; horas; días; información;	millones; caso; años; empresa; país; fiscal; gobierno; año; investigación; proyecto;	fiscal regional; fiscal nacional; tráfico de influencias; delitos tributarios; juicio oral; millones de pesos; director socioeconómico; semana pasada; secretario general; comisión investigadora;	fiscal regional; millones de pesos; tráfico de influencias; comisión investigadora; delitos tributarios; fiscal nacional; medidas cautelares; arraigo nacional; juicio oral; información privilegiada;	millones de pesos; millones de personas; fin de semana; prisión preventiva; días de presidio; delitos tributarios; embargo de bienes; acción judicial; semana distrital; crecimiento económico;	millones de pesos; tráfico de influencias; fiscal nacional; opinión pública; millones de dólares; clase política; proyecto de ley; delitos tributarios; medios de comunicación;
Milicogate	jefe; general; millones; gastos; ministra; caso; investigación; institución; investigación;	jefe: millones; comandante; caso; ministra; investigación; años; gastos; institución;	general; caso; jefe; dinero; viajes; pesos; información; viaje; ministra;	millones; jefe; general; investigación; ministro; caso; años; comandante; institución; gastos;	gastos reservados; caudales públicos; malversación de caudales; millones de pesos; institución castrense; prisión preventiva; uso de gastos; justicia militar; delito de malversación; comisión investigadora;	gastos reservados; caudales públicos; malversación de caudales; prisión preventiva; institución castrense; comisión investigadora; enriquecimiento ilícito; justicia mistar; delito reiterado;	exceso de equipaje; facturas falsas; empresas de turismo; comisión de servicio; correos electrónicos; monto total; jefe de finanzas; precios de pasajes; valor de factura; millón de pesos;	gastos reservados; millones de pesos; caudales públicos; malversación de caudales; millones de dólares; institución castrense; prisión preventiva; delito reiterado; monto total; fondos públicos;
Pacogate	institución; fraude; caso; fiscal; millones; investigación; director; personas;	institución; fraude; millones; investigación; años; fiscal; general; personas; director	años; personas; caso; año; millones; fraude; institución; horas; días;	millones; años; investigación; personas; año; empresa; general; país; institución;	gastos reservados; millones de pesos; fiscal regional; policía uniformada; caudales públicos; malversación de caudales; prisión preventiva; fiscal nacional; arresto domiciliario; arraigo nacional;	millones de pesos; gastos reservados; prisión preventiva; caudales públicos; malversación de caudales; fiscal regional; asociación ilícita; arresto domiciliario; policía uniformada; fiscal nacional;	millones de pesos; prisión preventiva; año pasado; fin de semana; casos de corrupción; semana pasada; policía uniformada; uicio oral; estallido social; redes sociales;	millones de pesos; gastos reservados; caudales públicos; prisión preventiva; estallido social; millones de dólares; malversación de caudales; año pasado; crimen organizado; redes sociales;
Rojas Vade	convencional; cáncer; mesa; proceso; renuncia; convencionales; enfermedad; órgano; situación;	convencional; cáncer; proceso; caso; mesa; convencionales; renuncia; situación; personas; enfermedad;	cáncer; convencional; situación; renuncia; declaración; enfermedad; delito; caso; órgano; hechos;	convencional; cáncer; renuncia; constituyentes; proceso; septiembre; caso; trabajo; órgano; convencionales;	mesa directiva; proceso constituyente; órgano constituyente; reforma constitucional; redes sociales; convencional constituyente; proyecto de ley; fin de semana; estallido social; millones de pesos;	mesa directiva; proceso constituyente; redes sociales; reforma constitucional; convencional constituyente; pueblos originarios; estallido de patrimonios; vacio normativo; renuncia formal; Tonbocitopenia imunitaria;	convencional constituyente; mesa directiva; proceso constituyente; órgano constituyente; reforma constitucional; estallido de intereses; millones de pesos; declaración de patrimonio; vacio normativo; renuncia formal; Tonbocitopenia imunitaria;	proceso constituyente; redes sociales; órgano constituyente; mesa directiva; reforma constitucional; estallido social; pueblos originarios; declaración de patrimonio; vacio normativo; renuncia formal; Tonbocitopenia imunitaria;

La triangulación de los datos arrojados, tanto por el análisis de frecuencia de mención, como por el de polaridad o sentimientos de los personajes vinculados a los casos considerados como escandalosos políticamente en Chile, y el análisis de frecuencia conceptual de las noticias relacionadas a los mismos, permite identificar tópicos y estrategias predominantes en la cobertura periodística de los escándalos de corrupción chilenos ocurridos entre 2014 y 2022 por parte de medios informativos chilenos presentes en la web, además de comparar los distintos tratamientos de dichos escándalos de corrupción.

Además, este tipo de herramienta al entregar las noticias más relevantes, se pudo realizar un análisis de contenido de las noticias (ver tabla 6). De esta manera, en este estudio se deja en evidencia que los medios hegemónicos, sean tradicionales o emergentes, tienen un tratamiento similar del hecho noticioso, con diferencias solo de énfasis sobre quiénes y de qué hablan. Cabe recordar que este tipo de medios contiene casi el 83 % de todas las emisiones de noticias relacionadas a los casos de corrupción estudiadas en esta investigación. Por tanto, tal como lo señalan Mellado y Scherman (2020) y Couso (2011), los medios en línea no parecieran contribuir a una mayor diversidad en la representación de distintos sectores sociales, e incluso parecieran tener predilección por informar sobre celebridades y personajes que aparecen en los medios tradicionales.

Tabla 6

Ejemplo de noticias sobre distintos personajes involucrados en casos de corrupción

Personaje	Tipo de medio	Nº noticias del caso	Nº noticias personaje	Conteo personaje	Puntaje personaje	Ejemplos
Eduardo Gordon (Pacogate)	HT	1400	98	7,00 %	-0,0207	“Intensa jornada enfrenta este miércoles el fiscal Eugenio Campos, quien encabeza la investigación del mega fraude al interior de Carabineros, pues desde las 10.00 de la mañana comenzó una rueda de toma de declaraciones a testigos en dependencias del OS-7 de Carabineros. De acuerdo a la información que se ha entregado, están citadas al menos seis personas, entre quienes se encuentra la exsubcontralora Dorothy Pérez, el ex contralor Ramiro Mendoza y el ex general director de Carabineros, Eduardo Gordon.” La Tercera, 20 Nov 2018.

Humberto Oviedo (Milicogate)	HE	597	114	19,10 %	-0,0653	“La decisión se adopta luego de que, el 1 de agosto pasado fue dejado en libertad el general (r) Humberto Oviedo, quien también es investigado por el mal uso de gastos reservados del Ejército por un monto cercano a \$4.500 millones.” CHV, 21 Ago 2019.
Jaime Orpis	NHT	130	4	3,08 %	0,0077	“Orpis se aterra a la libertad: último recurso ante la Corte aplazaría su ingreso a la cárcel. “Va a ser mi hogar”. Así describió el exsenador UDI Jaime Orpis a la cárcel en la que deberá cumplir la pena efectiva a la cual fue condenado en el caso Corpesca. Radio Saqo, 16 Nov 2021.
UDI	NHE	1448	245	16,92 %	-0,0539	“Aquí hay un problema que atraviesa prácticamente todo el espectro de partidos políticos y, coyunturalmente, ha salido a flote por el lado de la UDI.” El Ciudadano, 09 Oct 2014.

Con respecto al análisis de personajes, llama la atención la táctica de algunos medios, particularmente de los medios no hegemónicos tradicionales de omitir la mención de delitos, ocupando la estrategia de nombrar sustantivos concretos y que no tienen una carga moral negativa, neutralizando así la carga moral del hecho considerado como escandaloso, e incluso nombran adjetivos positivos como “libertad”, por lo que varios de los personajes señalados dentro de estos medios tienen una evaluación neutra o hasta positiva en la escala de polaridad o sentimientos, perpetuando así el *statu quo* del ecosistema informativo, sin una diferencia de perspectiva evidente con los medios hegemónicos, la que, inclusive, no se desea hacer, debido a posibles conexiones de autoridades políticas y públicas con distintos medios aquí analizados.

Asimismo, los medios no hegemónicos emergentes, se diferencian de los medios hegemónicos al dar mayor énfasis al dar mayor cobertura a autoridades nacionales, de todo el espectro político, pero con mayor perseverancia en el sector cercano a la UDI, relacionándolo a palabras negativas y vinculándolos con la dictadura de Pinochet, por lo que sus evaluaciones tienden a ser más negativas. También llama la atención en este tipo de medios la alta cobertura a personajes que no están involucrados de forma directa a los hechos de corrupción, como es el caso de Sebastián Piñera y Michelle Bachelet, ambos mandatarios de Chile desde que comienzan a conocerse los casos de corrupción.

Es importante señalar que en el proceso de análisis fue necesario que el equipo de investigación diera contexto en el caso de algunas ambigüedades que surgieron, pues tal como sugieren Xu *et al.* (2014) y Sutriawan *et al.* (2024), es importante hacer los ajustes correspondientes a la devolución del software ante ambigüedades léxicas. Esto ocurrió en el caso en donde dos personas se nombran exactamente igual, y el software no es capaz de diferenciarlo (como ocurre con gerentes de empresas en donde padres e hijos se llaman igual y ambos son gerentes). También ocurrió cuando los personajes no siempre fueron nombrados de la misma forma en las noticias, por lo que el software asume que, si un apellido es individualizado de su nombre, tomará al apellido como un personaje a parte (ejemplo, Piñera y Sebastián Piñera), o, si en el nombre del personaje hay un acento y hay medios que lo nombran sin acento, entonces también el programa asume que son dos personajes distintos (ejemplo, Loncon y Loncón).

Conclusiones

Este estudio identifica tratamientos y estrategias predominantes en la cobertura periodística de los escándalos de corrupción chilenos ocurridos entre 2014 y 2022 por parte de medios informativos chilenos presentes en la web, además de comparar los distintos tratamientos de dichos escándalos de corrupción.

Es importante señalar también que, con respecto al uso de una herramienta que combina inteligencia artificial, lingüística y ciencias sociales, se presentaron algunas limitaciones. Las observadas en este estudio fueron tres en términos generales.

La primera ocurrió en el caso en donde dos personajes se nombran igual, como ocurre con gerentes de empresas en donde padres e hijos se llaman igual y ambos son gerentes.

La segunda ocurrió cuando los personajes son nombrados de forma distinta, ya sea por su nombre, apellido o sobrenombre, o, si en el nombre del personaje hay un acento o apóstrofe, y hay medios que lo nombran con o sin acento, asumiendo que son personajes distintos.

Ambas limitaciones nombradas resultan de la dificultad de analizar fenómenos lingüísticos de lenguaje natural mediante técnicas de procesamiento automático del lenguaje, como lo es la ambigüedad (Xu, *et al.*, 2024; Sutriawan *et al.*, 2024). En ambos casos ocurren ambigüedades léxicas en donde dos personas que se llaman igual o una persona nombrada de forma distin-

ta, lo que exige de la observación profunda del equipo investigador para dar contexto y poder generar el análisis en profundidad.

La tercera limitación del estudio, que es un poco más específica, resalta ante el hecho de que quienes hacen justicia, como Rurherford y Chong, se carguen de polaridad negativa, lo cual invita a pensar en mayores ajustes necesarios al procesamiento automático del lenguaje, pues en alguna medida, resulta una prueba de las limitaciones que pueden tener estas herramientas. Sin embargo, pese a estas limitantes, el programa ofreció una guía para el análisis de metadatos de forma gratificante.

Ante las limitaciones presentadas en este estudio, se proyectan futuros estudios que guardan relación con el análisis informatizado de grandes volúmenes de datos en la web, con el fin de identificar otros posibles sesgos en los textos multimediales.

El análisis de sentimientos utilizado en este estudio permitió encontrar distintos sesgos, sin embargo, faltó profundizar en otros, como los de género, por ejemplo, al captar que las fiscales mujeres eran peor evaluadas que las personas condenadas, en su mayoría hombres, o también faltó ahondar en los sesgos discriminatorios hacia la figura de Elisa Loncon, a la que vinculaban con los pueblos indígenas, vinculaciones que presuntamente podrían dar una connotación, sea negativa o positiva, pero que afectaría a su evaluación en la escala de polaridad.

Indagar en las evaluaciones morales y posibles sesgos que podrían estar afectando en los textos de las noticias en la web constituye una tarea desafiante, debido a las dificultades de procesar algunos fenómenos lingüísticos subyacentes, como se vio en este estudio con el hecho de poder individualizar nombres de una única forma, o dificultades propias de este tipo de análisis que consisten en calcular o dimensionar la ironía, el sarcasmo y la subjetividad de las opiniones (García-Vega *et al.*, 2020), como también calcular y dimensionar posibles sesgos discriminatorios en los algoritmos que permiten estos análisis (O’Neil, 2017; Martínez *et al.*, 2020), con el fin de generar nuevos algoritmos que posibiliten estudios más equitativos y no discriminatorios.

Agradecimientos

Este trabajo se realizó gracias al financiamiento de la Beca Doctorado Nacional N°21201115 y del Fondecyt Regular N°1241726, ambos de la Agencia Nacional de Investigación y Desarrollo (ANID), Chile.

Referencias bibliográficas

- Amadeo, B. (2008). Framing: Modelo para armar. En M. T. Baquerín (ed.), *Los medios ¿aliados o enemigos del público?* EDUCA.
- Arcila-Calderón, C., Barbosa-Caro, E. y Cabezuelo-Lorenzo, F. (2016). Técnicas Big Data: Análisis de textos a gran escala para la investigación científica y periodística. *El Profesional de La Información*, 25(4), 623-631. <https://doi.org/10.3145/epi.2016.jul.12>
- Aruguete, N., Batista, F., Calvo, E., Guizzo-Altube, M., Scartascini, C. y Ventura, T. (2024). Framing fact-checks as a “confirmation” increases engagement with corrections of misinformation: a four-country study. *Sci Rep* 14, 3201. <https://doi.org/10.1038/s41598-024-53337-0>
- Ayala, T. (2016). Redes sociales, poder y participación ciudadana. *Revista Austral de Ciencias Sociales*, 26, 23-48. <https://acortar.link/1ykPo9>
- Baviera, T. (2017). Técnicas para el análisis de sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength. *Dígitos*, 3, 33-50. <https://doi.org/10.7203/rd.v1i3.74>
- Cárcamo, L., Cárdenas, C., Sáez, D. y Toural, C. (2021). Fake news en Chile y España: ¿Cómo los medios nos hablan de noticias falsas? *Journal of Iberian and Latin American Research*, 1-18. <https://doi.org/10.1080/13260219.2020.1909849>
- Cárcamo-Ulloa, L., Cárdenas-Neira, C. y Scheihing-García, E., Sáez-Trumper, D., Vernier, M. y Blaña-Romero, C. (2023). On politics and pandemic: how do Chilean media talk about disinformation and fake news in their social networks? *Societies*, 13(25), 1-14. doi.org/10.3390/soc13020025
- CIPER. (2015a). *El periodismo que remece a Chile*. Catalonia.
- CIPER. (2015b). *Lo mejor de CIPER 3. El periodismo que remece a Chile*. Editorial Catalonia.
- Cordero, R. y Marín, C. (2006). Los medios masivos y las transformaciones de la esfera pública en Chile. *Documento de trabajo ICSO*, 7(2). <https://acortar.link/1ykPo9>
- Couso, J. (2011). El mercado como obstáculo a la libertad de expresión: la concentración de la prensa escrita en Chile en la era democrática. En B. Sorj, *Democracia y medios de comunicación. Más allá del estado y el mercado* (pp. 109-142). Catálogos.
- García Cumberas, M. Á., Villena Román, J., Martínez-Cámara, E., Díaz Galiano, M. C., Martín-Valdivia, M. T. y Ureña-López, L. A. (2016). Resumen

- de TASS 2016. *TASS 2016: Workshop on Sentiment Analysis at SEPLN Proceedings*, 13-21. <https://acortar.link/1ykPo9>
- García-Vega, M., Díaz-Galiano, M., García-Cumbreras, M., Plaza del Arco, F., Montejo-Ráez, A., Jiménez-Zafra, S., Martínez, E., Aguilar, C., Sobrevilla, M. A., Chiruzzo, L. y Moctezuma, D. (2020). Overview of TASS 2020: Introducing Emotion Detection. *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2020) Co-Located with 36th Conference of the Spanish Society for Natural Language Processing (SEPLN 2020)*, 163-170. <https://acortar.link/1ykPo9>
- Heiss, C. (2023). De las calles a las urnas. El estallido social tras el fracaso del proceso constituyente. En Batthyány, K. (ed.), *Constituir en la revuelta. Nuevos desafíos a la democracia*. CLACSO. <https://acortar.link/1ykPo9>
- Irrarázabal, I. y Cruz, F. (2023). Confianza institucional en Chile: un desafío para el desarrollo. *Punto de referencia* (682), 1-34. <https://acortar.link/1ykPo9>
- Koziner, N. (2015). El framing: Un programa de investigación para el estudio de las comunicaciones mediáticas. *Animus*, 14(28), 22-45. <https://doi.org/10.5902/2175497720203>
- Lopez-Rabadan, P. y Mellado, C. (2018). Twitter as a space for interaction in political journalism. Dynamics, consequences and proposal of interactivity scale for social media. *Communication & Society* 32(1), 1-18. <https://doi.org/10.15581/003.32.37810>
- Lovón, M. y Parini, A. (2024). Inteligencia artificial, discurso digital y lenguaje: La lingüística en la era de las nuevas tecnologías. *Lengua y Sociedad*, 23(2), 577-598. <https://doi.org/10.15381/lengsoc.v23i2.29773>
- Martínez, C. D., García, P. D. y Sustaeta, P. N. (2020). Sesgos de género ocultos en los macrodatos y revelados mediante redes neurales: ¿hombre es a mujer como trabajo es a madre? *Reis: Revista Española de Investigaciones Sociológicas*, 172, 41-59. <https://doi.org/10.5477/cis/reis.172.41>
- Matus, A. (2020). The end of the Chilean fantasy. *Dissent*, 67(2), 97-105. <https://dx.doi.org/10.1353/dss.2020.0020>
- Mellado, C. y Scherman, A. (2020). Mapping source diversity across Chilean news platforms and mediums. *Journalism Practice*, 15(7), 974-993. <https://doi.org/10.1080/17512786.2020.1759125>
- Morstatter, F., Wu, L., Yavanoglu, U., Corman, S. y Liu, H. (2018). Identifying Framing Bias in Online News. *ACM Trans. Soc. Comput. I*(2), 1-18. <https://doi.org/10.1145/3204948>

- Nash, C. (2017). Dictadura, transición y corrupción. Algunas lecciones del caso chileno. *Revista Electrónica de Investigación y Asesoría Jurídica-REDIAJ*, 1(8), 353-408. <https://acortar.link/1ykPo9>
- Navia, P. y Ulriksen, C. (2017). Tuiteo, luego voto. El efecto del consumo de medios de comunicación y uso de redes sociales en la participación electoral en Chile en 2009 y 2013. *Cuadernos.Info*, (40), 71-88. <https://dx.doi.org/10.7764/cdi.40.1049>
- Oliva, S. (2017). *Introducción a Pandas y Jupyter Notebook de Python*. <https://bit.ly/4f9gjWA>
- O'Neil, C. (2017). *Armas de Destrucción Matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia*. Capitán Swing.
- Otero, V. (2020). *Political media's bias, in a single chart*. Ad Fontes Media.
- Palau, A. y Davesa, F. (2013). El impacto de la cobertura mediática de la corrupción en la opinión pública española. *Revista Española de Investigaciones Sociológicas*, 144(1), 97-126. <https://doi.org/10.5477/cis/reis.144.97>
- Pastrana, A. (2019). Estudio sobre la corrupción en América Latina. *Revista mexicana de opinión pública*, (27), 13- 40. <https://doi.org/10.22201/fcpys.24484911e.2019.27.68726>
- Pereyra, S. (2012). La política de los escándalos de corrupción desde los años 90. *Desarrollo Económico*, 52(206), 255-284. <https://bit.ly/41FjIPp>
- Rodríguez-Virgili, J., López-Escobar, E. y Tolsá A. (2011). La percepción pública de los políticos, los partidos y la política, y uso de medios de comunicación. *Comunicación y Sociedad*, 24(2), 7-39. <https://bit.ly/4nQbjKe>
- Sádaba, T., Rodríguez, J. y Bartolomé, M. (2012). Propuesta de sistematización de la teoría del framing para el estudio y praxis de la comunicación política. *Observatorio Journal*, 6(2), 109-126. <https://doi.org/10.15847/obsOBS622012540>
- Saez-Trumper, D., Castillo, C. y Lalmas, M. (2013). Social media news communities: Gatekeeping, coverage, and statement bias. In Proceedings of the 22nd ACM International Conference on Information & Knowledge Management. *ACM*, 1679-1684. <https://doi.org/10.1145/2505515.2505623>
- Sánchez-Holgado, P., Martín Merino, M. y Blanco Herrero, D. (2020). Del data-driven al data-feeling: análisis de sentimiento en tiempo real de mensajes en español sobre divulgación científica usando técnicas de aprendizaje automático. *Anuario Electrónico de Estudios en Comunicación Social "Disertaciones"*, 13(1), 35-58. <https://doi.org/10.12804/revistas.urosario.edu.co/disertaciones/a.7691>

- Serrano, V. y Salvat, G. (2015). *Tras el ensayo digital. Una aproximación interdisciplinar a la sociedad de la información*. Ediciones UACH.
- Sixto-García, J., Soengas-Pérez, X., Rodríguez-Vázquez, A.-I., Vázquez-Herrero, J. y López-García, X. (2021). Percepción social del periodismo en España. *AdComunica*, (22), 191-210. <https://doi.org/10.6035/2174-0992.2021.22.11>
- Sola-Morales, S. y Rivera Gallardo, R. (2017). El tratamiento periodístico sobre la corrupción política. Análisis comparado del caso SQM en Chile y del caso de Bárcenas en España. *Estudios sobre el Mensaje Periodístico*, 23(1), 647-662. <https://doi.org/10.5209/ESMP.55619>
- Sutriawan, S., Rustad, S., Shidik, G. F., Pujiono, P. y Muljono, M. (2024). Review of ambiguity problem in text summarization using hybrid ACA and SLR. *Intelligent Systems with Applications*, 22, 200360.
- Vernier, M., Cárcamo Ulloa, L. y Scheihing García, E. (2017). Diagnóstico de la estrategia editorial de medios informativos chilenos en Twitter mediante un clasificador de noticias automatizado. *Revista Austral de Ciencias Sociales*, (30), 183-201. <https://doi.org/10.4206/rev.austral.cienc.soc.2016.n30-09>
- Vernier, M., Cárcamo Ulloa, L. y Scheihing García, E. (2018). Pensamiento crítico de los jóvenes ciudadanos frente a las noticias en Chile. *Comunicar*, 26(54), 101-110. <https://doi.org/10.3916/C54-2018-10>
- Vernier, M. y Ruiz, F. (2023). *Manual de usuario. Sophia search: software de recopilación automática e indexación de documentos web para la investigación en comunicación y territorios*. <https://bit.ly/4nZzaXW>
- Vidal-Sepúlveda, M., y Olivares-Rodríguez, C.. (2021). ¿Quién elige mi dieta informativa? La datificación automatizada de personas en el ambiente digital. *Información, cultura y sociedad*, (44), 143-154. <https://doi.org/10.34096/ics.i44.8627>
- Waisbord, S. (2020). ¿Es válido atribuir la polarización política a la comunicación digital? Sobre burbujas, plataformas y polarización afectiva. *SAAP*, 14(2), 249-279. <https://bit.ly/44YnbBe>
- Wilenman, J. y Feddersen, M. (2023). Contesting authority in the crisis of neoliberalism: The Chilean Spring and the mobilization of human rights frames. En Steven A. Boutcher, Corey S. Shdaimah y Michael W. Yarbrough (eds.), *Research Handbook on Law, Movements and Social Change*. Edward Elgar. <https://doi.org/10.4337/9781789907674.00036>
- Xu, S., Li, Z., Mei, K. y Zhang, Y. (2024). Core: Llm as interpreter for natural language programming, pseudo-code programming, and flow programming of ai agents. *arXiv preprint arXiv:2405.06907*, 2.

- Zalaquett, J. (2011). Conflictos de intereses: normas y conceptos. *Anuario de Derechos Humanos*(7), 179-189. <https://doi.org/10.5354/adh.v0i7.17363>
- Zunino, E. y Focás, B. (2021). La corrupción como problema público en la televisión argentina: interacción entre medios y audiencias. *Anagramas Rumbos y Sentidos de la Comunicación*, 20(39), 129-155. <https://doi.org/10.22395/anagr.v20n39a6>

Declaración de Autoría - Taxonomía CRediT	
Autor	Contribuciones
Camila Rojas-Sánchez	Conceptualización, Curación de datos, Análisis formal, Investigación, Visualización, Redacción-borrador original, Redacción-revisión y edición.
Luis Cárcamo-Ulloa	Conceptualización, Adquisición de fondos, Supervisión, Redacción-revisión y edición.
Matthieu Vernier	Curación de datos, Metodología, Software, Validación.

Declaración de Uso de Inteligencia Artificial
Los autores DECLARAN que la elaboración del artículo <i>Cobertura web de la corrupción en Chile: análisis automatizado con inteligencia artificial y lingüística</i> , no contó con el apoyo de Inteligencia Artificial (IA).