



Monitoreo de plataformas digitales sociales en Chile: un examen  
a la industria y sus productos

Tesis para optar al grado de Magíster en Comunicación Política

Rubén Darío Santander Herrera

Profesora guía: Patricia Peña Miranda

Santiago de Chile

2024

## RESUMEN

Tomando como ejemplo paradigmático el caso del llamado “informe Big Data”, que mediante un análisis de las conversaciones en redes sociales pretendió explicar las causas y apuntar a los responsables del estallido social de octubre de 2019 en Chile, este estudio efectúa un examen del sector de las agencias de monitoreo de plataformas sociales (AMP) en el país. Adoptando la perspectiva de la economía política de la comunicación (EPC), la investigación se enfoca en la oferta de servicios y mercancías de estas empresas. Para ello, lleva a cabo un extenso análisis crítico del discurso (ACD), con un encuadre basado en corpus y asistido por herramientas computacionales, sobre los contenidos de los sitios web corporativos de las compañías de este tipo que operan en Chile.

Los resultados de esta aproximación revelan que se trata de una industria culturalmente diferenciada, aunque dinámica y caracterizada por una variedad de orientaciones corporativas. Sus productos, cuya materialidad difusa pone énfasis en la velocidad, vastedad y variedad de los datos que analizan y en sus componentes visuales, aparecen como herramientas fundamentalmente persuasivas dirigidas a contextos gerenciales, en contra de cierta concepción de que se trataría de instrumentos comparables a los de las ciencias sociales. Las AMP se posicionan como especialistas en hacer accesible y manejable un mundo comunicacionalmente caótico, proveyendo a las élites locales de los símbolos socialmente legitimados de (tener) la razón, hoy encarnados en lo que se ha denominado el sublime digital y la ideología del *big data*.

Este trabajo es un aporte a la literatura existente en comunicación política al ofrecer una amplia revisión de las fuentes contemporáneas sobre el campo de las plataformas digitales, el desarrollo de una metodología poco explorada para este tipo de problemáticas y una perspectiva crítica sobre un sector que ha tenido una influencia relevante en la historia comunicacional y política reciente del país.

Palabras clave: Análisis de redes sociales, plataformas digitales, economía política de la comunicación, informe Big Data, agencias de monitoreo de plataformas, minería de datos

# AGRADECIMIENTOS

Para Manuela Badilla y gracias a ella.



# ÍNDICE

<b>INTRODUCCIÓN .....</b>	<b>1</b>
<i>Plan de la tesis.....</i>	<i>1</i>
<i>Un informe extraordinariamente sofisticado.....</i>	<i>3</i>
<i>Un contexto extraordinariamente conmocionado .....</i>	<i>4</i>
<i>La tesis de la intervención extranjera .....</i>	<i>5</i>
<i>La embestida comunicacional .....</i>	<i>7</i>
<i>Hacia la adquisición de las evidencias .....</i>	<i>11</i>
<i>El poder del big data .....</i>	<i>16</i>
<i>El informe de Alto Data Analytics .....</i>	<i>19</i>
Alto Data Analytics: una comprensión incomparable del mundo.....	22
El contenido del informe Big Data.....	24
Las críticas al informe Big Data.....	28
<i>Factores problemáticos de las agencias de monitoreo de plataformas digitales .....</i>	<i>34</i>
<i>Sobre las mercancías analíticas del sector AMP .....</i>	<i>35</i>
<b>1. ANTECEDENTES .....</b>	<b>39</b>
1.1. <i>El sistema de las plataformas digitales: conceptos centrales .....</i>	39
1.1.1. La situación de Internet a fines del primer cuarto del siglo XXI.....	40
1.1.2. La mercancía en el contexto de un capitalismo orientado a lo digital .....	42
1.1.3. Datos, nube y algoritmos .....	44
1.1.3.1. La computación en la nube .....	46
1.1.3.2. Los algoritmos .....	47
1.1.4. <i>Big data y big social data .....</i>	49
1.1.5. La cocción cultural de los datos.....	51
1.1.6. Redes sociales y plataformas .....	53
1.1.7. El sistema de las plataformas y sus características .....	55

1.1.8.	Plataformas publicitarias.....	63
1.2.	<i>Agencias de monitoreo de plataformas (AMP)</i> .....	69
1.2.1.	Tipificación de la minería de datos.....	70
1.2.2.	La labor de las AMP.....	71
1.2.3.	El papel de las interfaces de programación de aplicaciones.....	73
<b>2.</b>	<b>MARCO TEÓRICO</b> .....	<b>77</b>
2.1.	<i>Perspectivas de estudio en torno al uso de los datos sociales masivos digitales</i> .....	77
2.1.1.	Perspectivas administrativas.....	79
2.1.1.1.	Publicidad y marketing.....	80
2.1.1.2.	Microfocalización y democracia.....	82
2.1.2.	Perspectivas críticas.....	84
2.1.2.1.	<i>Big data</i> y ciencias sociales.....	85
2.1.2.2.	Biopolítica del dato.....	86
2.2.	<i>Economía política de la comunicación y la información: un enfoque para el análisis crítico de las agencias de monitoreo de plataformas</i> .....	88
2.2.1.	La mercantilización de la experiencia humana.....	89
2.2.2.	El punto de vista de la economía política de la comunicación.....	90
2.2.3.	Algunas observaciones relevantes realizadas por la EPC.....	92
2.3.	<i>La EPC aplicada al problema del monitoreo de plataformas en Chile</i> .....	95
2.3.1.	La cadena de valor de los datos.....	96
2.3.2.	Sobre el origen de los datos y la pertinencia de la metáfora extractivista.....	98
2.3.3.	La relevancia de las API y el modelo de subvenciones cruzadas.....	101
2.3.4.	Ratings y números.....	105
2.3.5.	Consecuencias sociales de este modelo de producción.....	109
2.3.6.	El otro extremo de la cadena de valor de las plataformas.....	114
2.3.7.	Colonialismo de datos y continuidad colonial.....	117
2.3.7.1.	Consecuencias de este enfoque para nuestra aproximación.....	120
<b>3.</b>	<b>OBJETIVOS DE LA INVESTIGACIÓN</b> .....	<b>123</b>
3.1.	<i>Objetivo general</i> .....	125
3.2.	<i>Objetivos específicos</i> .....	125

<b>4. DISEÑO METODOLÓGICO DE LA INVESTIGACIÓN .....</b>	<b>127</b>
4.1. <i>Enfoque de la investigación.....</i>	127
4.2. <i>Encuadre y procedimientos .....</i>	128
4.3. <i>Análisis crítico del discurso .....</i>	129
4.3.1. Selección del discurso de interés .....	131
4.3.2. Localización y preparación de las fuentes de los datos .....	131
4.4. <i>Criterios para delimitar el sector AMP.....</i>	133
4.5. <i>Procedimientos para la producción del corpus de análisis .....</i>	138
4.6. <i>Análisis del corpus.....</i>	142
4.7. <i>Etapas del análisis de concordancia.....</i>	143
4.8. <i>Triangulación .....</i>	147
<b>5. RESULTADOS.....</b>	<b>149</b>
5.1. <i>Análisis crítico de los discursos en torno a la oferta de las AMP chilenas .....</i>	150
5.2. <i>Problemas durante la preparación de los textos.....</i>	151
5.3. <i>Análisis .....</i>	152
Categoría 1: El producto concede una ventaja por su velocidad o poder predictivo .....	154
Categoría 2: El producto concede una ventaja por la vastedad, profundidad o completitud de la información que abarca .....	156
Categoría 3: El producto permite enfrentar a un nuevo, complejo y desafiante escenario ..	157
Categoría 4: El producto es una empresa exitosa y pionera en el ámbito tecnológico .....	159
Categoría 5: El producto es una empresa dispuesta para cumplir cualquier expectativa de sus clientes.....	161
Categoría 6: El producto permite hacer accionable la información .....	162
Categoría 7: El producto ofrece la capacidad de visualizar la información y mostrar resultados.....	163
Categoría 8: El producto otorga control sobre el poder o atractivo de los números y métricas .....	165
Categoría 9: El producto posibilita el uso de información previamente desperdiciada o extrae un potencial que estaba oculto.....	166

Categoría 10: El producto permite generar perfiles detallados de las personas .....	167
Categoría 11: El producto posibilita incidir directamente sobre la opinión pública o la sociedad .....	168
5.4. <i>Análisis de los discursos en torno a los desafíos técnicos y metodológicos asociados al sector</i> .....	169
Ámbito 1: Explicitación de barreras y dificultades .....	170
Ámbito 2: Exposición de aspectos metodológicos .....	170
Ámbito 3: Relacionado con las cualidades y la calidad de los datos .....	172
Ámbito 4: Relacionado con los valores de la investigación .....	173
<b>6. DISCUSIÓN.....</b>	<b>175</b>
6.1. <i>Volumen, variedad y velocidad</i> .....	178
6.2. <i>Un escenario ominoso</i> .....	179
6.3. <i>De la recuperación de los medios de producción espiritual</i> .....	180
6.4. <i>La relevancia de lo visual</i> .....	183
6.5. <i>El deseo por los símbolos legitimados de conocimiento</i> .....	185
6.6. <i>Sobre los discursos en torno a los desafíos técnicos y metodológicos asociados al sector</i> .....	187
<b>7. CONCLUSIONES.....</b>	<b>189</b>
7.1. <i>Definición y mapeo de las AMP</i> .....	191
7.2. <i>Evaluación crítica de las mercancías analíticas</i> .....	193
7.3. <i>Principales características de las agencias de monitoreo de plataformas en Chile...</i>	198
7.4. <i>La relevancia de este estudio en la era post-API</i> .....	203
7.5. <i>Colofón: aporte metodológico de esta investigación</i> .....	208
<b>BIBLIOGRAFÍA.....</b>	<b>210</b>
<b>ANEXOS.....</b>	<b>235</b>
1. <i>Documentos públicos disponibles relacionados con el informe Big Data</i> .....	235



2.	<i>Registro completo de las unidades significativas consideradas en el análisis.....</i>	236
----	--	-----

## IMÁGENES

Imagen 1.	La Tercera, edición impresa del lunes 28 de octubre de 2019. Las dos notas con supuestas evidencias sobre la intervención extranjera en el estallido social aparecen en la misma página y son de la misma autoría.....	9
Imagen 2.	La Tercera, edición digital del 28 de octubre de 2019. Las dos notas sobre la intervención extranjera en el estallido social perdieron a su autor en las versiones publicadas en el portal de Internet del medio.....	10
Imagen 3.	La Segunda, martes 24 de diciembre de 2019. Primera página del reportaje sobre las gestiones de La Moneda para obtener un informe que respaldara la tesis de la intervención extranjera.....	12
Imagen 4.	Comunicado de Quiñenco publicado en la cuenta personal de Twitter de Andrónico Luksic. ..	17
Imagen 5.	Detalle del sitio web corporativo de Alto Analytics. Captura realizada en 20 de agosto de 2021. ....	23

## FIGURAS

Figura 1.	Cadena de valor de los datos. Elaboración propia.....	98
Figura 2.	Proporciones relativas de las 11 categorías de análisis identificadas de acuerdo a su aparición en el corpus. Elaboración propia.....	154

## TABLAS

Tabla 1.	Conceptos generales y variantes habituales característicos del sector AMP. Elaboración propia. ....	135
Tabla 2.	Mapeo preliminar del sector AMP en Chile. ....	138
Tabla 3.	Detalles del corpus de análisis. ....	141
Tabla 4.	Términos clave para el análisis de concordancia relacionado con las mercancías y servicios del sector AMP.....	144
Tabla 5.	Términos clave para el análisis de concordancia relacionado con los problemas y desafíos de la minería de datos.....	146
Tabla 6.	Detalles del corpus de sitios web corporativos internacionales.....	148
Tabla 7.	Mapeo de empresas del sector AMP.....	193



# INTRODUCCIÓN

## **Plan de la tesis**

En esta introducción analizaremos un caso particular, con importantes connotaciones políticas y sociales en el contexto chileno, para a través de sus detalles aproximarnos desde un punto de vista práctico a nuestro objeto de estudio: las organizaciones que comercializan “monitoreos de redes sociales”. El objetivo de esta sección es permitirnos revelar a estas empresas en sus aspectos generales, pero también abordar algunas de sus particularidades que, a nuestro entender, son centrales para dar cuenta de la relevancia de esta investigación.

En los posteriores capítulos abordaremos el problema general aquí propuesto, admitiendo que se trata de una primera aproximación a un sector que ha sido rara vez analizado de forma específica en la literatura de las ciencias sociales y humanas, así como en la comunicación política. Por ello, la investigación ofrece un énfasis exploratorio-descriptivo, lo que será expuesto con más detalles en la metodología.

En el primer capítulo, Antecedentes, desarrollaremos el contexto de esta investigación. El propósito central de esta sección es definir y delimitar una serie de conceptos y fenómenos para poder afrontar con la mayor claridad posible un área de análisis que presenta amplios espacios de opacidad debido a su desarrollo constante e incontenible, entre otros factores. Partiendo de esta exposición, exploraremos las ideas que sostienen lo que denominaremos sistema de las plataformas digitales, entramado sociotécnico y económico de compleja clasificación. En la parte final de este capítulo, abordaremos nuestro objeto de análisis particular, las agencias de monitoreo de plataformas sociales, algunas de sus características comerciales y cómo localmente se entrelazan, a través de las interfaces de programación de aplicaciones (API), con el sistema de las plataformas de escala global.

El segundo capítulo, Marco Teórico, inicia con el desarrollo del estado del arte de las investigaciones enfocadas en las plataformas digitales y el *big social data*. En dicho apartado pasaremos revista a las diferentes perspectivas de análisis que han buscado aproximarse a este campo desde la investigación en comunicaciones y las ciencias sociales. Esto nos permitirá acotar con mayor transparencia nuestros propósitos teóricos. Habiendo establecido dónde se sitúa este estudio, en la segunda parte del capítulo abordaremos las fuentes teóricas particulares que guían la investigación, que se apoyan esencialmente en lo que se ha denominado economía política de la comunicación (EPC). Tras esto, en el último apartado de la sección, a la luz de los conceptos expuestos, retomaremos el examen de las agencias de monitoreo de plataformas, proponiendo un encuadre teórico para el análisis de sus modos de producción particulares.

Con estos aspectos desplegados, en el tercer capítulo presentaremos los objetivos de la investigación. En el capítulo cuarto desarrollaremos una metodología ajustada a dichos objetivos mediante procedimientos acordes a las características de nuestro objeto de estudio. Como justificaremos más adelante, para efectuar nuestra aproximación empírica hemos optado por apoyarnos en el Análisis Crítico del Discurso (ACD), auxiliados por herramientas computacionales de minado de texto.

En el quinto capítulo daremos cuenta de los resultados de esta aproximación y en el capítulo sexto discutiremos dichos resultados en relación con los preceptos esbozados en nuestro marco teórico.

Finalmente, en el séptimo capítulo, Conclusiones, sintetizaremos nuestros principales hallazgos y delinearemos algunos de los aspectos que creemos merecen ser considerados a futuro para profundizar en la comprensión de las organizaciones que comercializan minería de datos de plataformas digitales a nivel local.

## Un informe extraordinariamente sofisticado

El jueves 19 de diciembre de 2019, el Gobierno de Sebastián Piñera sumó un nuevo elemento a la ya saturada conversación en torno a las movilizaciones sociales que cumplían dos meses agitando las calles de Chile: el concepto de *big data*. Ya se había destacado desde varios frentes el rol de las redes sociales en el periodo, pero hasta ese día nadie había concedido de forma decidida a tales apreciaciones un carácter de prueba pericial o de análisis sistemático con capacidad explicativa del estallido social.

Esa mañana, al ser consultado sobre las afirmaciones del Gobierno respecto a la supuesta participación de grupos organizados en el origen de las manifestaciones, el fiscal nacional Jorge Abbott informó que La Moneda solo le había entregado “alguna información de fuentes abiertas, no de inteligencia” (CNN Chile, 2019b). Inmediatamente, el ministro del interior de ese entonces, Gonzalo Blumel, impugnó lo afirmado por Abbott, señalando: “Eso no es así. Se entregó información extraordinariamente sofisticada, a partir de análisis con tecnología de *big data*, con tecnología de la información, que dan cuenta de antecedentes que son importantes para la investigación penal”. Y al ser interrogado por la demora en la entrega de estos antecedentes a la justicia, Blumel indicó: “Las investigaciones y la información, en particular de inteligencia, es información delicada que se tiene que ir consolidando y transfiriendo a las instituciones de acuerdo a los procedimientos que establece la ley” (CNN Chile, 2019b).

Quedaba servida de esta forma, de cara a la opinión pública, la idea de que el Gobierno tenía en sus manos un “sofisticado” estudio desarrollado sobre cuantiosos volúmenes de datos digitales, sobre *big data*; que este análisis, que había sido elaborado meticulosamente atendiendo a lo delicado de su naturaleza, tenía el carácter de información de inteligencia; y que se trataba de una información tan contundente que por sí sola fundamentaría, en el plano penal, las hasta ese momento arbitrarias afirmaciones del presidente y su gabinete respecto a una supuesta intervención extranjera en la crisis política y social que vivía el país. Inmediatamente, la prensa denominó a este extraordinario documento informe Big Data.

## **Un contexto extraordinariamente conmocionado**

En octubre de 2019, lo que se inició como una serie de acotadas protestas de estudiantes secundarios en Santiago motivadas por el alza en la tarifa del transporte público, desencadenó una seguidilla de masivas movilizaciones en distintos puntos de la capital y en otras ciudades del país en contra del Gobierno. En un fenómeno que tomó por sorpresa a comentaristas y académicos, las demandas ciudadanas rápidamente se diversificaron. En muchos puntos las protestas derivaron en enfrentamientos con la policía y saqueos. El Gobierno respondió con medidas y declaraciones destempladas que agudizaron la situación (Piñera, 2019). El 25 de octubre se desarrolló en Santiago una manifestación que reunió a al menos 1.200.000 personas (Montes, 2019). Marchas masivas se realizaron también en ciudades como Concepción, Valparaíso, Rancagua, La Serena, Punta Arenas y Arica, así como en localidades pequeñas y alejadas de los grandes centros urbanos y en un sinnúmero de emplazamientos por todo el mundo en que los chilenos residentes, muchas veces junto a ciudadanos locales, mostraban su solidaridad con los eventos en Chile. El Instituto Nacional de Derechos Humanos (INDH) informó haber asistido como observador a 345 manifestaciones públicas a lo largo del país (Bravo, 2022).

Durante el mes de noviembre de 2019 se recrudeció la represión policial ante las protestas y también la violencia de algunas manifestaciones. Entre octubre de 2019 y marzo de 2020 se reportaron 33 personas muertas en contextos de protestas, varias de ellas asesinadas por carabineros o fallecidas a causa de sus acciones (El Mostrador, 2020). Hubo miles de heridos, entre los que se cuentan más de 300 personas con trauma ocular producto del impacto de perdigones de goma y metal usados por la policía (Senado de Chile, 2020).

Los motivos de las personas movilizadas durante este periodo eran variados, sin embargo, según múltiples análisis y encuestas de opinión, respondían a causas similares que hundían su origen en la desigualdad estructural del país y las características del Estado subsidiario instaurado por la dictadura militar (Araujo, 2019). El Gobierno de Sebastián Piñera, no obstante, optó por apegarse a una tesis heteróclita.

## La tesis de la intervención extranjera

Detrás de la tesis de la intervención extranjera en el estallido social emerge una serie de supuestos difíciles de probar, pero que han sustentado hasta hoy el imaginario de un sector del país. La lógica primordial que entraña esta tesis es que las manifestaciones iniciadas en octubre de 2019 no serían un fenómeno orgánico, sino un acto coordinado por regímenes extranjeros de izquierda con la intención de desestabilizar al Gobierno de Piñera, ya que su éxito menoscabaría a las alternativas socialistas en el continente y legitimaría el modelo neoliberal vigente en Chile. Complementariamente, se trataría de una suerte de venganza por parte de Venezuela y sus aliados (Rusia, Cuba e incluso Bolivia), el bloque bolivariano, motivada por el “liderazgo” asumido por Piñera en contra del Gobierno de Nicolás Maduro al apoyar a Juan Guaidó y en general a la disidencia venezolana (Dammert y Sazo, 2021).

La narrativa de la conspiración bolivariana era compartida por varios de los gobiernos integrantes de la Alianza del Pacífico para explicar sus propias crisis sociales internas. Frente a las protestas protagonizadas a principios de octubre de 2019 por grupos indígenas en Ecuador que se oponían a los ajustes dictados desde el Fondo Monetario Internacional, el presidente Lenin Moreno denunció —sin pruebas— la presencia de “grupos armados infiltrados” que formarían parte de una conspiración urdida desde Caracas. En Colombia, en tanto, a raíz de las protestas que se desencadenaron a partir del 21 de noviembre de 2019, motivadas por una seguidilla de casos de corrupción y de asesinatos de líderes sociales, el presidente Iván Duque también aludió a un complot bolivariano y expulsó a 59 venezolanos del país por supuestamente haber iniciado las manifestaciones (Robinson, 2019).

Alineado con este relato, Luis Almagro, el entonces secretario general de la Organización de Estados Americanos (OEA), también sin presentar evidencia alguna, señaló en varias ocasiones que las masivas manifestaciones de 2019 en Latinoamérica serían atribuibles a los gobiernos de Nicolás Maduro y de Miguel Díaz-Canel. En un comunicado dado a conocer a la prensa el 24 de octubre, señaló: "Las brisas del régimen bolivariano impulsadas por el madurismo

y el régimen cubano traen violencia, saqueos, destrucción y un propósito político de atacar directamente el sistema democrático” (France 24, 2019).

Asimismo, en Bolivia, tras el golpe de Estado en contra de Evo Morales en noviembre de 2019, la nueva presidenta de facto, Jeanine Áñez, denunció que las protestas indígenas en su contra estaban orquestadas desde el extranjero y expulsó a todos los diplomáticos venezolanos y cubanos (Robinson, 5 de diciembre de 2019). En cada uno de estos países, la tesis del complot chavista permitió justificar el asesinato de decenas de manifestantes y una violencia policial inaudita, como han acreditado múltiples organismos de derechos humanos.

En Chile, la idea de la intervención extranjera se convertiría en una hipótesis sobre la que las autoridades de Gobierno volverían con frecuencia en las semanas que siguieron al 18 de octubre de 2019. Localmente, es posible que esta tesis estuviera alimentada en parte o fundada por completo en un informe de la Dirección de Inteligencia del Ejército (DINE) que le fue entregado al presidente el 20 de octubre, a solo dos días de que se desatara la crisis. Un año después, los detalles de ese informe trascendieron a la prensa. En él se aseguraba que el Servicio Bolivariano de Inteligencia Nacional, vinculado al Servicio de Inteligencia Cubano, había introducido en Chile un “batallón” de 600 agentes clandestinos para llevar a cabo operaciones de insurrección. El comandante de este batallón, según tal informe, sería Pedro Carvajalino, jefe de una organización chavista gubernamental llamada Zurda Konducta. Uno de los problemas con la información recabada por el Ejército era que Zurda Konducta no es una organización paramilitar o de inteligencia, sino un programa juvenil de Venezolana de Televisión (VTV), y Carvajalino era simplemente su conductor. Para la DINE, una de las pruebas de que Carvajalino se encontraba en Chile liderando dicho “batallón” insurgente era una fotografía manipulada, de forma bastante evidente, donde aparece con La Moneda de fondo y que él mismo subió a sus redes sociales (Cossio, 2020).

Pero en octubre de 2019 nada de eso era materia de escrutinio público y dentro de La Moneda la tesis de la intervención extranjera gozaba de buena salud. Ello pese a que, como se ha sabido en los años posteriores, la opinión de las Fuerzas Armadas sobre el origen del estallido



social se distanciaba rotundamente no solo de las evidencias, sino que también de las teorías que manejaban las policías. Así lo ha ratificado posteriormente el entonces ministro del Interior, Gonzalo Blumel:

Los análisis provenientes de las Fuerzas Armadas diferían en ciento ochenta grados de los análisis policiales. Mientras los primeros aseguraban la presencia de agentes foráneos, sin mediar pruebas sustantivas, los segundos descartaron completamente la injerencia foránea, cosa que además parecían ratificar los datos. Menos del 2 a 3 % de los detenidos en las protestas correspondían a extranjeros, cifra muy inferior a la proporción de extranjeros residentes en el país. (Blumel, 2023)

Pese a esto, el Ejecutivo, cuya figura más visible durante esos días era el mismo Blumel, se empeñó en respaldar y tratar de fortalecer a la tesis de la intervención.

## **La embestida comunicacional**

En la madrugada del 28 de octubre de 2019, La Tercera publicó en su sitio web una breve nota titulada “Gobierno rastrea rol de venezolanos en Twitter” (La Tercera, 2019), que luego apareció en la página 5 de su edición impresa firmada por “S. Vayona” (Imagen 1). Pese al sugerente titular, la información que contenía era más bien escasa y confusa. En poco más de 300 palabras se aludía a presuntas “informaciones de prensa que daban a entender que el gobierno de Nicolás Maduro estaba detrás” de una campaña en redes sociales que fomentaba las protestas. Esto habría empujado al Gobierno de Chile a elaborar un informe que buscaba conectar el uso de ciertos *hashtags* en Twitter con una posible coordinación desde Venezuela. Para esto se descargaron “todos los tuits asociados a los hashtags” que provinieran “desde el centro de Caracas y en un radio de 100 millas a la redonda de esa ciudad” (no hay ninguna explicación para el uso del sistema anglosajón de unidades).

Cabe señalar que la geolocalización precisa de las publicaciones está deshabilitada por defecto en Twitter, lo que vuelve a ese metadato no solo muy escaso, sino que poco fiable (Twitter, 2023b). Según precisa Ricardo Baeza-Yates, pese a que se ha documentado que en Twitter la geolocalización explícita puede llegar al 66 %, en su experiencia analizando datos de Chile, el porcentaje no es mayor del 14 % (Baeza-Yates, 2020). El análisis, según la nota de prensa, concluye indicando que “si bien los tuits provenientes desde Venezuela no representan un porcentaje significativo del total de los tuits que marcaron tendencias en Chile, estas fueron utilizadas para aumentar artificialmente las métricas de Chile”. Qué métricas, quién las medía, cuál era el efecto de esas mediciones y en qué medida se aumentaron “artificialmente”, entre otras interrogantes, quedaban para la imaginación del lector.

La misteriosa noticia de La Tercera no indicaba qué repartición gubernamental había desarrollado este informe, ni qué uso pretendía darle el Gobierno, ni cómo había llegado a manos de La Tercera; no contenía otras fuentes, ni consultas a especialistas para poder contextualizar los datos y, desde luego, tampoco estaba firmada por ningún periodista. Al cierre, se revelaba la existencia de un segundo documento que estaría en posesión de La Moneda, el que habría sido entregado por autoridades de Estados Unidos y donde se realizaba el mismo curioso aunque inconducente ejercicio, pero dirigido a cuentas provenientes de Rusia. Nunca más se supo de ninguno de estos dos informes y ningún otro medio aportó detalles adicionales.

Pese a lo llamativo de su contenido, la nota pasó relativamente desapercibida, quizá porque el mismo medio publicó en la misma página de su edición impresa de ese mismo lunes una noticia que tuvo una repercusión mucho mayor: “Policía identifica a uno de los autores de incendios en estaciones de Metro” (Vadoya, 2019), en referencia a uno de los eventos delictivos más graves ocurridos en el contexto del estallido social, la quema de siete estaciones del Metro de Santiago. En la bajada de la nota, también firmada por “S. Vadoya”, se indica que “Según fuentes de inteligencia policial, se detectó participación de ciudadanos cubanos y venezolanos en las manifestaciones” y en el cuerpo de la noticia se reitera esta conexión (Imagen 1). La información, sin duda muy grave, fue reproducida profusamente en diversas redes sociales y por medios chilenos e incluso extranjeros, que daban por sentado que existía evidencia de intervención

extranjera (Infobae, 2019). Sin embargo, durante esa jornada, el fiscal a cargo de la investigación desmintió la información en una entrevista radial e indicó que “no tenemos antecedentes que nos permitan llegar a la identificación de persona alguna en este minuto y mucho menos determinar su nacionalidad” (CNN Chile, 2019a). Tras esto, La Tercera se vio obligada a realizar una aclaración y a modificar la noticia publicada en su sitio web, incluso haciendo desaparecer al autor. Este mismo procedimiento se aplicó a la nota sobre los *hashtags* de Twitter (Imagen 2).



Imagen 1. La Tercera, edición impresa del lunes 28 de octubre de 2019. Las dos notas con supuestas evidencias sobre la intervención extranjera en el estallido social aparecen en la misma página y son de la misma autoría.



*Imagen 2. La Tercera, edición digital del 28 de octubre de 2019. Las dos notas sobre la intervención extranjera en el estallido social perdieron a su autor en las versiones publicadas en el portal de Internet del medio.*

A fines de octubre se sumó otro elemento a alimentar la teoría de la intervención extranjera. Esta vez se trató precisamente de la intervención de una potencia extranjera: el jueves 31 de ese mes, la Casa Blanca informó a través de un comunicado, que circuló en diversos medios y agencias, que el presidente Donald Trump había llamado a su par chileno para manifestarle su apoyo y que en la conversación “denunció los esfuerzos extranjeros para minar las instituciones chilenas, la democracia y la sociedad”, sin hacer mayores precisiones (Cooperativa, 2019). Estas afirmaciones del Gobierno de Estados Unidos estarían conectadas con ciertos análisis dados a conocer por el Departamento de Estado de ese país sobre campañas de influencia provenientes desde Rusia. Ante las consultas del New York Times al respecto, funcionarios de ese país admitieron, no obstante, que la gran mayoría de las publicaciones sobre las protestas parecían ser legítimas, es decir, que no formaban parte de ninguna maniobra comunicacional (Jakes, 2020).

Aunque la tesis ya no solo cundía dentro del Gobierno de Chile, sino que era parte del relato habitual de sus partidarios, tanto internos como foráneos, las evidencias para sustentarla estaban resultando más bien esquivas. Pese a ello, el 9 de noviembre Piñera concede una extensa entrevista al medio español El País, en la que declara haber “recibido mucha información, alguna de ella de origen externo, que afirma que aquí hubo intervención de Gobiernos extranjeros” (Montes y González, 2019). Al realizar estas afirmaciones pareciera ser que no estaba dando por auténticos los datos contradictorios que hasta ese momento las instituciones de orden y seguridad le habían presentado, ni el análisis de los *hashtags* de Twitter que según La Tercera el propio Gobierno había realizado, ya que el mandatario además indica con claridad que todavía carece de evidencias: “Yo creo que hoy día hay algo nuevo, distinto a lo que teníamos hace un mes atrás, pero necesito pruebas para poder afirmarlo”.

## **Hacia la adquisición de las evidencias**

Lo que planteaba con convicción Piñera entonces —la tesis de la intervención extranjera— entrañaba la confianza en que encontraría las pruebas que respaldarían sus dichos. Esto es consistente con informaciones que fueron reveladas posteriormente: en ese mismo periodo, el Gobierno se encontraba activamente buscando a alguien que fuera capaz de producir tales evidencias. Alguien que pudiera triunfar donde las agencias de inteligencia del Estado habían fracasado, ya porque sus resultados apuntaban en la dirección contraria a la que perseguía el Gobierno (Carabineros, PDI) o porque sus capacidades estaban resultando insuficientes (DINE, ANI).

El 6 de noviembre, pocos días antes de que apareciera la entrevista a Piñera en El País, los académicos Sebastián Valenzuela (Universidad Católica), Juan Pablo Luna (Universidad Católica), Jorge Pérez (Universidad de Chile) y Bárbara Poblete (Universidad de Chile), investigadores del Instituto Milenio Fundamento de los Datos (IMFD), fueron citados a La Moneda. Habían sido convocados por Jorge Selume, en aquel momento director de la Secretaría de Comunicaciones (Secom), y Magdalena Díaz, jefa de gabinete del presidente. En un reportaje

de La Segunda titulado “Big data: Las gestiones de Selume y Díaz para conseguir un informe” (Violani y Riesco, 2019) los primeros tres académicos revelaron algunos detalles de este encuentro, en que Selume y Díaz les solicitaron confeccionar un documento de “inteligencia” a partir de los datos disponibles en redes sociales para comprobar una “intervención extranjera” en el estallido social. Como señala Pérez en su entrevista, la respuesta de los especialistas fue negativa y tajante: “Le dijimos que la información de redes sociales no puede vincular influencias extranjeras en desórdenes, ni identificar a personas que generan disturbios, ni nada de ese tipo de cosas que era lo que ellos estaban buscando”.

2 **Sociedad** | La Segunda martes 24 diciembre 2019

**Académicos del Instituto Milenio cuentan cómo fue su reunión en La Moneda**

## Big data: Las gestiones de Selume y Díaz para conseguir un informe

**Dos profesores de la UC y uno de la U. de Chile dicen que el titular de la Secom y la jefa de gabinete del Presidente les pidieron comprobar que en el estallido del 18-O hubo intervención extranjera.**

Por Nicolás Violani y Josefa Riesco

**E**n una reunión de una hora, realizada en La Moneda el 6 de noviembre, el titular de la Secretaría de Comunicaciones del Gobierno, el psicólogo y experto en big data, Jorge Selume junto a la jefa de gabinete del Presidente Piñera, Magdalena Díaz, más otras dos personas, le pidieron a académicos del Instituto Milenio Fundamentos de los Datos que a través de información disponible en las redes sociales, les confeccionaran un documento de “inteligencia”, que ayudara a comprobar la existencia de “intervención extranjera” en el estallido social del 18-O y su desarrollo posterior.

En entrevista con “La Segunda”, los profesores remarcan que ellos rechazaron realizar el trabajo y que no son quienes están detrás del informe que el ministro del Interior, Gonzalo Blumel le entregó al fiscal nacional, Jorge Abbott el jueves 19 de diciembre, luego de que el titular del Ministerio Público asegurara el día anterior que no existen indicios de intervención extranjera en la crisis social, contradiciendo así a la vocera de palacio, Karla Rubilar y al propio Presidente.

El Mandatario dio a conocer su tesis el 11 de noviembre en el diario El País de España, donde aseguró que “he recibido mucha información, alguna de ella de origen externo, que afirma que aquí hubo intervención de gobiernos extranjeros. Insistió el lunes 16 de diciembre en una entrevista con Archi, donde aseguró que además “hay tecnología de punta, hay organización militar”.

Hoy el fiscal metropolitano Oriente, Manuel Guerra se refirió al informe entregado y dado a conocer por La Tercera el domingo 22 de diciembre, afirmando que “el aporte de antecedentes resulta bastante escaso para fundamentar la persecución penal y la intervención extranjera”. Desde el Gobierno no se ha dado a co-

**Magdalena Díaz y Jorge Selume.**

**“Magdalena Díaz nos dijo: ‘Voy ser muy franca, yo quiero saber de inteligencia’”.** Sebastián Valenzuela, académico UC.

nocer quién lo encargó ni la firma que lo realizó (ver recuadro). Los académicos consultados para este artículo sospechan que el documento final es de alguna firma que con posterioridad sí aceptó el trabajo. Sus versiones dan luz a un episodio, respecto del cual también consultamos en las oficinas de ministro Blumel y de Jorge Selume y Magdalena Piñera, quienes hasta el cierre de esta edición no se refieren al tema.

**El llamado**

La versión de la reunión del 6 de noviembre es contada por los académicos del centro científico Sebastián Valenzuela (UC), Juan Pablo Luna (UC) y Jorge Pérez (U. de Chile).

“Debe haber sido una semana después del 18 de octubre. Magdalena Díaz me consultó a mí, primero — pues un par de semanas antes habíamos participado en una reunión del Ministerio de Ciencias—, para pedir la opinión de distintos científicos sociales respecto de lo que estaba ocurriendo en Chile. Como yo no entiendo tanto la parte tecnológica ni la parte de circulación de noticias falsas, convoqué al resto de la ‘genie’, cuenta el científico político Juan Pablo Luna.

Díaz y Selume luego recibieron a los académicos, incluyendo también a Bárbara Poblete de la U. de Chile.

“Magdalena Díaz partió muy honestamente comentando que no cachaba mucho de tecnología y que no tenía muy claro por qué estábamos ahí”, relata Valenzuela, quien agrega que ella fue al grano: “voy ser muy franca, yo quiero saber

**Las dudas**

Persiste el misterio respecto de quién realmente encargó el informe que sí se elaboró y de la firma que lo confeccionó. A través de un comunicado de la Subsecretaría del Interior, el Gobierno dice que lo obtuvo de la Agencia Nacional de Inteligencia, organismo que tiene el deber guardar reserva de sus actuaciones, por lo que no puede confirmar ni desmentir nada. El comunicado también descarta pagos con fondos públicos. Aunque el informe dado a conocer por La Tercera contiene una serie de datos generales — como que el 19,3% de los comentarios indagados por el Gobierno tiene origen internacional —, el ministro del Interior Gonzalo Blumel lo validó al decir que lo entregado al fiscal nacional Jorge Abbott el jueves 19 de diciembre contenía “información extraordinariamente sofisticada a partir de análisis con tecnología big data”. Consultamos a Interior por las razones de Blumel para darle ese respaldo al documento, pero hasta el cierre de esta edición no había respuesta.

Imagen 3. La Segunda, martes 24 de diciembre de 2019. Primera página del reportaje sobre las gestiones de La Moneda para obtener un informe que respaldara la tesis de la intervención extranjera.

En esa ocasión, la y los académicos le advirtieron a La Moneda que hasta el momento no había ningún indicio que ellos hubieran podido detectar de una actividad concertada o una coordinación de las comunicaciones relacionadas con las manifestaciones. Al cerrar la reunión, el director de la Secom y la jefa de gabinete de Piñera les pidieron contactos de personas u organizaciones que pudieran estar interesadas en realizar un informe con las características que se habían conversado. Los investigadores, según contaron a La Segunda, les aseguraron que “no se puede hacer inteligencia con datos no estructurados de redes públicas y quien diga que puede hacerlo, vende humo” (Violani y Riesco, 2019).

Lo que Selume y Díaz estaban solicitando no era algo extraordinario, sino un servicio bastante común que con frecuencia se conoce como “monitoreo de redes sociales”, pero que también puede recibir otros apelativos. *Social media analytics*, análisis de redes sociales, escucha social o *social listening*, *social media insights* o *social media monitoring*, se encuentran entre los más habituales. El sector muchas veces distingue entre estas diferentes denominaciones como pasos dentro de sus procesos (Tran, 2020), pero en otras ocasiones las considera y comercializa como actividades diferentes e independientes. No obstante, todas estas operaciones consideran, en última instancia, la realización de alguna forma de minería de datos sobre una o varias plataformas sociales de Internet. Esto es, básicamente, contar interacciones (likes, retweets, etc.) y relacionarlas con los metadatos asociados para extrapolar conclusiones sobre el comportamiento y los intereses de un conjunto de personas usuarias en un rango de tiempo dado.

En pos de la inteligibilidad, llamaremos a todas las empresas que operan en este sector Agencias de Monitoreo de Plataformas (AMP). Usando diversas herramientas, estas empresas ofrecen diferentes análisis y visualizaciones para describir subconjuntos de usuarios y tendencias dentro de las redes sociales, insumos que son particularmente valorados en el ámbito del marketing. Dependiendo de la complejidad de las aproximaciones y las herramientas utilizadas, que pueden incluir procesamiento de lenguaje natural y el uso de grandes modelos de lenguaje (IA), estos servicios pueden variar ampliamente en valor.

Con todo, se trataba de un tipo de producto o servicio al que, ya en 2019, medianas y grandes empresas e instituciones accedían habitualmente de una u otra forma. El mismo Jorge Selume había fundado años antes una compañía (Artool) que proveía asesorías de este tipo a grandes empresas como Banco de Chile y a partidos políticos como Renovación Nacional, por lo que sabía de lo que estaba hablando (Saleh, 2018). Más aún, era un campo bien conocido por el propio presidente Piñera y por sus colaboradores cercanos. En 2011, durante su primer mandato y también en el contexto de masivas movilizaciones ciudadanas y estudiantiles, Piñera contrató a la empresa BrandMetric, una de las grandes compañías de este sector en el país, para prestar servicios a la Secom, lo que en su momento generó un pequeño revuelo (Gatica, 2011).

La reunión sostenida entre La Moneda y un grupo de personas que para las autoridades eran posibles proveedores de este tipo de productos muestra algo que ya se ha observado en otros contextos respecto a la relación entre clientes y agencias de minería de datos de redes sociales. En uno de los capítulos de su libro *Post, Mine, Repeat* (2016), Helen Kennedy realiza una de las raras aproximaciones que existen en la literatura directamente encaminada a describir a estas empresas. Kennedy las llama “compañías intermediarias” y las describe en términos genéricos como *commercial social media insights companies*. En su trabajo, la autora indica que con frecuencia los clientes de estas compañías ignoran las advertencias que se les hacen sobre la calidad de los datos de las plataformas y el alcance de los análisis que se pueden desarrollar a partir de ellos. Los informantes de Kennedy aluden a una especie de autoengaño por parte de quienes compran estos servicios:

La mayoría de mis entrevistados reconocieron que a menudo se encuentran con datos demográficos irrelevantes, limitados o inexactos y con obstáculos relacionados con las



limitaciones de las herramientas, y muchos sugirieron que a los clientes no les importa que esto dé lugar a datos “aproximados.”<sup>1</sup> (Kennedy, 2016, p. 109).

A los consumidores de estos monitoreos no les preocupaba que se les informara que sus preguntas no se podían responder a partir del análisis de los datos de las redes sociales, solamente les interesaba obtener “números para poner en un PowerPoint que puedan enseñar a su jefe.” (Kennedy, 2016, p. 109). En estos casos, que según el estudio de Kennedy parecen ser frecuentes, los trabajadores de estas compañías más que *vender humo* terminan actuando en complicidad con sus clientes para *producir humo*.

Jorge Selume, el director de la Secom durante la crisis del estallido social, era precisamente un personaje con amplia experiencia en este tipo de compañías intermediarias. Por esto, la solicitud que formuló a los académicos del IMFD y la insistencia por obtener determinados números para poder “enseñar al jefe”, pese a las advertencias de los especialistas e independiente de la calidad de los datos, podríamos leerla como un indicio de la forma de relación cliente-agencia que describe Kennedy. Pese a que la negativa de la y los investigadores le enrostró a La Moneda que no estaba apuntando a los proveedores adecuados, tanto el director de la Secom como el propio presidente de la República sabían que existen innumerables organizaciones capaces de suministrar el tipo de producto que buscaban. Desconocemos si con posterioridad a esta fallida reunión el Gobierno acudió a otros académicos o si contactó a alguno de los múltiples proveedores activos en el mundo privado en Chile, pero sabemos que se realizó al menos una gestión adicional para contar con un monitoreo de redes sociales que confirmara la tesis de la intervención extranjera.

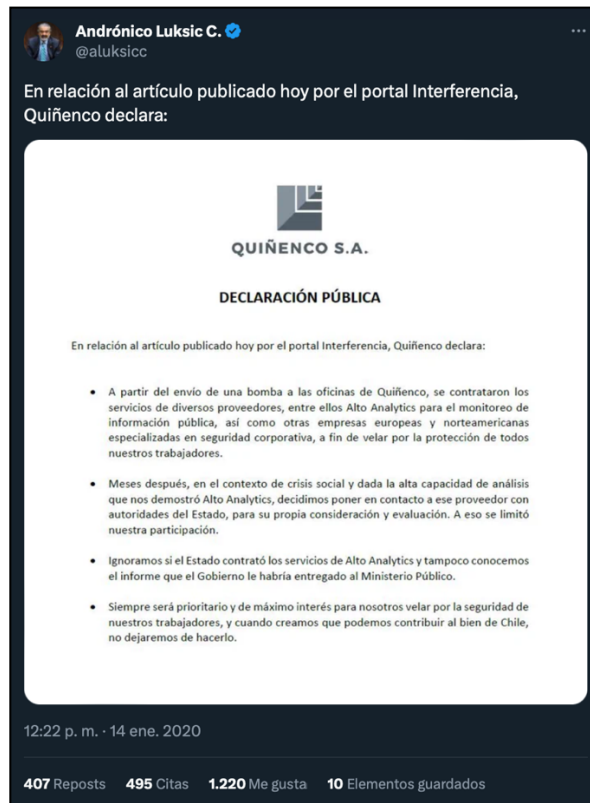
---

<sup>1</sup> Traducción propia del original en inglés. A menos que se indique algo diferente, todas las citas directas a fuentes en inglés que hemos incluido en este texto fueron traducidas especialmente por el autor de este trabajo.

## **El poder del *big data***

El lunes 2 de diciembre de 2019, como reveló el medio Interferencia (Herrero, 2020a; Herrero, 2020b), se llevó a cabo una reunión en las dependencias de la Agencia Nacional de Inteligencia (ANI) en el centro de Santiago. A este encuentro habrían asistido representantes de las tres ramas de las Fuerzas Armadas, de Carabineros e integrantes de la ANI, incluyendo a su director, Gustavo Jordán. También se encontraba Rodrigo Ubilla, subsecretario del Interior en ese momento, y dos ejecutivos del grupo Quiñenco (filial matriz del grupo Luksic, el consorcio empresarial más poderoso de Chile, dueños del Banco de Chile) que a su vez habían formado parte del primer Gobierno de Piñera. Estos eran Rodrigo Hinzpeter, gerente legal de Quiñenco y exministro del Interior y de Defensa, y Mauricio Lob, gerente de Asuntos Corporativos y Comunicaciones del grupo Quiñenco, exdirector de la Secom y exsubsecretario del Ministerio Secretaría General de Gobierno. Junto a estos altos funcionarios, lo más extraordinario de la reunión era la presencia de al menos un experto de la empresa española de monitoreo de redes sociales y “análisis de Big Data” (según su propia jerga) Alto Data Analytics. Se trataba del español José Luis Cansado (La Tercera, 2020), director general de la firma.

En esa reunión, que tuvo una segunda sesión al día siguiente, se habría discutido lo que posteriormente fue conocido como informe Big Data. Los detalles del encuentro, su origen y objetivos precisos, así como las razones de la presencia de altos ejecutivos de una empresa privada, nunca fueron aclarados por las autoridades. Cuando posteriormente, en enero de 2020, se reveló que esta reunión había tenido lugar, Quiñenco afirmó, a través de un escueto comunicado (Imagen 4), que su papel fue exclusivamente poner en contacto a la empresa española con el Gobierno “para su propia consideración y evaluación”, pero que ellos no conocían el informe en cuestión, no lo encargaron ni habían pagado por él (CNN Chile, 2020).



*Imagen 4. Comunicado de Quiñenco publicado en la cuenta personal de Twitter de Andrónico Luksic.*

No obstante, parece claro que tras estos acercamientos del más alto nivel, ese grupo de importantes hombres con dilatadas trayectorias en el Estado, la inteligencia, la defensa y los negocios, se formó una opinión muy positiva sobre las competencias de la empresa española. Lo que sucede posteriormente ese mes de diciembre da cuenta de un Ejecutivo que ya se siente poseedor de las evidencias que necesitaba para demostrar su tesis sobre el verdadero origen del estallido social.

El 11 de diciembre el presidente Piñera graba una entrevista televisiva con el periodista argentino Andrés Oppenheimer para CNN en español. En ella insiste en que ha existido una campaña en contra de su Gobierno, que “en eso ha habido, sin duda, participación de gobiernos e instituciones extranjeras” (CNN Chile, 2019c) y que muchas de las imágenes distribuidas por redes sociales que mostraban los abusos de la policía chilena son falsas. Además, dando a

entender no solo que ya tenía en su poder los antecedentes que Blumel llamaría después “análisis con tecnología de *big data*”, sino que se trataría de una investigación realizada o encargada por su Gobierno, afirma: “Hemos hecho un estudio de millones y millones de comunicaciones a través de redes sociales. Muchas de ellas vienen fuera de Chile. La inmensa mayoría de ellas son para provocar más desórdenes, más manifestaciones, más descontrol, más desobediencia civil” (CNN Chile, 2019c). Aunque se mantiene aferrado a la misma tesis, desde este momento en adelante la presenta como algo respaldado por los datos. Y no cualquier dato, sino que grandes volúmenes de datos, extremadamente sofisticados, obtenidos de las comunicaciones en redes sociales.

Quizá por eso, el 16 de diciembre, cuando se estaban cumpliendo dos meses desde el inicio de la crisis social, Piñera concede otra entrevista, esta vez a las radios de la Asociación de Radiodifusores de Chile (Archi), y decide iniciar su alocución reiterando la tesis de Gobierno sobre el origen del fenómeno. Esta vez, siendo aún más explícito en cuanto al rol de las redes sociales. “Cuando se queman, en forma simultánea, estaciones de metro, supermercados, hospitales, detrás de eso no hay maldad, hay tecnología de punta, un manejo de las redes sociales que vienen fuera de Chile”, acusó en esa oportunidad y conminó a la Fiscalía a hacerse cargo de esos antecedentes “con la mayor diligencia posible” (Diario Uchile, 2019).

A raíz de esta intervención, al día siguiente el fiscal metropolitano Oriente, Manuel Guerra, indicó en entrevista con una radioemisora “somos los primeros sorprendidos con esta información cuando además no nos ha sido puesta a disposición”. Guerra puntualizó que aún no contaba con ningún antecedente concreto “que dé cuenta de algún grupo que haya actuado desde afuera o de un grupo de extranjeros que haya actuado al interior del país para cometer estos delitos” (El Mostrador, 2019).

En su libro *La vuelta larga: Crónica personal de la Crisis de Octubre* (2023), Blumel da a entender que fueron estas declaraciones del fiscal Guerra, “quien se permitió desmentir al Presidente Piñera respecto de una eventual injerencia extranjera en las jornadas de protesta”, las que lo motivaron a hacer público el informe de Alto Analytics, “un estudio que merecía alguna consideración”. Ante la confrontación que se había declarado entre el Poder Judicial y el

Ejecutivo, el exministro del Interior confiesa que el 19 de diciembre de 2019 intentó “zanjar el asunto durante un Consejo de Seguridad Pública que sostuvimos en el Ministerio del Interior, en el que por ley debía participar el fiscal nacional Jorge Abbott, entregándole personalmente una copia del reporte” (Blumel, 2023). Ese día, después de dos meses de buscar indicios y opiniones favorables a la tesis de la intervención extranjera y “con el ánimo, básicamente, de contener las críticas que venían lloviendo sobre el Presidente” (Blumel, 2023), el Gobierno hace pública la existencia de un documento que desde entonces se conocería como el informe Big Data. Pero lejos de poner un punto final a las diferencias con el Ministerio Público, las agudizó y generó lo que en palabras del mismo Blumel fue “un completo bochorno para el Gobierno”.

La pretensión era mostrar, a partir de la minería de datos de redes sociales, señales que apuntaran ya no a indagar las motivaciones de quiénes se manifestaban en contra del Gobierno a través de Internet (lo que en el mejor de los casos permitiría entender a las redes como un “termómetro de la sociedad”, como suelen decir quiénes se dedican a su análisis), sino para conocer a los responsables intelectuales de las protestas y desmanes que habían ocurrido en las calles a lo largo de Chile. Si dicha pretensión se podía concretar gracias al peso de antecedentes “extraordinariamente sofisticados” no harían falta evidencias policiales, la constatación de delitos o el apresamiento de espías y montoneros. Esta era la esperanza explícita que el Presidente y su gabinete ponía en el informe y, al dar a conocer su existencia, Blumel fue enfático en señalar la contundencia que este tendría como medio de prueba. Sin embargo, a diferencia de Sebastián Piñera, que no ocultaba su orgullo respecto a que dicha investigación había sido hecha con los recursos de los que disponía su Gobierno, el ministro, con más prudencia, dejó en vilo su autoría.

## **El informe de Alto Data Analytics**

Es difícil evaluar un informe que muy pocos han visto. El día posterior a que el ministro Blumel revelara su existencia, ya había ciudadanos solicitando conocerlo en virtud de la Ley 20.285 sobre acceso a la información pública. Los argumentos esgrimidos en ese momento por el Ministerio del Interior para negar la entrega del documento resultan, en retrospectiva, un tanto

contradictorios. Frente a la solicitud N.º AB001T0001252<sup>2</sup>, el subsecretario del Interior, Juan Francisco Galli, responde —a través de un oficio— que “el Informe fue puesto a disposición de la Agencia Nacional de Inteligencia por parte de un tercero” y alude a la Ley N.º 19.974 sobre “la obligación de guardar secreto respecto de los antecedentes que obran en poder de los organismos de Inteligencia” (Galli, 2020).

En esta respuesta, el Ejecutivo se muestra enfático en calificar al informe Big Data como un documento de inteligencia, en directa oposición a la opinión del fiscal nacional que, al recibirlo, había declarado que este consistía en “información de fuentes abiertas, no de inteligencia”. Posteriormente, Abbott abundaría en esto, afirmando que “no hemos encontrado ningún valor (en el informe) y en consecuencia no ha sido necesario indagar el origen del mismo. (...) Lo que hace es un análisis del funcionamiento de redes sociales, qué participación tienen, quién aparece más frecuentemente” (Basoalto, 2020).

Al parecer, el ministro Blumel, al menos en un comienzo, coincidía con el fiscal nacional respecto a que el documento entregado el día 19 de diciembre no era información de inteligencia, ya que en esa misma oportunidad Abbott precisó ante la prensa que el ministro se había comprometido a hacerle llegar la (verdadera) información de inteligencia. Según informó La Tercera, ese envío se habría concretado el viernes 20, cuando el director de la ANI se reunió brevemente con el fiscal nacional. De este segundo documento, que sí tendría carácter de información de inteligencia para las partes involucradas, no existe ningún indicio ni se hizo ninguna aclaración posterior. Tampoco se incorporó a las causas penales en curso, ya que el director Jordán no realizó el acto de liberación que exige la ley (Ayala, 2019).

La Contraloría General de la República, por su parte, pese a haber oficiado al ministro Gonzalo Blumel en febrero de 2020 para que transparentara el informe, el día 24 de septiembre

---

<sup>2</sup> Los documentos mencionados en esta sección se incluyen en los anexos de esta tesis.

de 2020 emitió el dictamen E37908N20 respondiendo a un requerimiento del senador Pedro Araya, donde indica:

El Ministerio del Interior y Seguridad Pública manifestó, en síntesis, que el documento a que alude la presentación del rubro fue entregado por la empresa Alto Data Analytics a la Agencia Nacional de Inteligencia, la que, a su vez, lo puso a disposición de esa Cartera de Estado. Añade que consistía en una muestra de un software de análisis de información proveniente de fuentes abiertas, y que luego de recibirla, la envió al Ministerio Público. (Bermúdez, 2020)

Esta descripción, algo más completa sobre las características del informe, no entra en detalles respecto a por qué dicha documentación, una muestra de carácter promocional de un producto de análisis de fuentes abiertas de Internet, debería acogerse a la “obligación de guardar secreto” que impone Ley N.º 19.974. Aunque, desde luego, no existe una obligación legal de detallar los criterios que manejan los responsables de la inteligencia chilena.

Este dictamen de la Contraloría abunda sobre la condición de muestra gratis del informe Big Data:

Luego, respecto de la forma en que la Agencia Nacional de Inteligencia accedió a dicho informe, cabe señalar que este le fue proporcionado por la empresa Alto Data Analytics, en calidad de muestra, sin mediar costo —según indica el singularizado Ministerio—, a fin de mostrar posibles productos a entregar para generar inteligencia. (Bermúdez, 2020)

Resulta curioso que un gobierno decida movilizar su capital político en pro de una tesis controversial (la influencia extranjera en la supuesta planificación de unas protestas) basándose exclusivamente en el presunto conocimiento aportado por un producto entregado de forma no solicitada y gratuita. Esto abre muchas preguntas respecto a la agenda del Gobierno, pero también en lo referente al modelo de negocios de las compañías de minería de datos como Alto Data

Analytics, así como sobre la legitimidad que tienen este tipo de análisis en opinión de algunos tomadores de decisiones.

### **Alto Data Analytics: una comprensión incomparable del mundo**

Alto Data Analytics es una empresa creada en 2012 en Madrid por el español Alejandro Romero, ingeniero de la Universidad de Málaga con un MBA de la escuela de negocios de Henley (Alto Analytics, 2015). Al momento de realizar el informe promocional para la inteligencia chilena en 2019, la descripción de los servicios de la compañía en su sitio web señalaba que entre otras actividades, el “software y los algoritmos de inteligencia artificial de Alto convierten miles de millones de puntos de datos públicos en ciberinteligencia accionable para una mejor y más rápida toma de decisiones” y que “mapea relaciones y patrones digitales, proporcionando una comprensión inigualable del mundo” (Alto Analytics, s.f.). Como otras empresas a nivel global y local, Alto Data Analytics ofrece una variedad de productos abarcados bajo el rótulo impreciso y general de “análisis masivo en tiempo real de fuentes públicas de datos” con lo que aseguran poder entregar a diversas organizaciones “conocimientos críticos para la toma de decisiones”. El propio Romero indicaba en una entrevista concedida en 2016 que su objetivo era “proporcionar inteligencia accionable para la toma de decisiones de nuestros clientes” (Tringali, 2016).

A lo largo de este trabajo analizaremos en detalle los productos que ofrecen las compañías intermediarias a las que aquí hemos denominado AMP. Alto Data Analytics parece constituir un buen ejemplo de la oferta habitual que presenta el sector, por lo que se ofrece como un buen punto de partida. En su sitio, la empresa presenta tres servicios. El primero es Alto Analyzer, descrito como “la herramienta de software premium más innovadora y avanzada para mapear y agrupar conversaciones digitales, conectando miles de millones de puntos de datos de más de 125 países”. Luego está Alto Insights, descrito como un “sofisticado servicio de análisis de datos” a cargo de un “equipo internacional de datos de clase mundial” dedicado a mantener “sus necesidades a la vista para ayudar a convertir el vasto y complejo mundo de los datos públicos en inteligencia valiosa y accionable. Los entregables incluyen informes visuales de alto impacto



adaptados a sus preguntas estratégicas más importantes”. Este producto es, aparentemente, similar a la muestra gratis recibida por el Gobierno de Piñera.

El tercer producto que presenta la empresa es Alto Oculus Virtual Reality, consistente en un “entorno de análisis de realidad virtual patentada (optimizado para los cascos Oculus VR)” que permite “volar a través de visualizaciones interactivas y experimentar de primera mano el poder único de Alto para convertir datos sofisticados en inteligencia accionable”. No queda muy clara cuál podría ser la utilidad real de esta vistosa aplicación, pero sí que el concepto de “sofisticación”, el mismo usado por el ministro Blumel para describir el informe Big Data, es central en la oferta de la compañía.



*Imagen 5. Detalle del sitio web corporativo de Alto Analytics.  
Captura realizada por el autor en 20 de agosto de 2021.*

Junto con estos servicios, Alto Analytics también ofrece lo que denomina “soluciones”, casos de uso en los que sus clientes podrían aplicar las herramientas de la compañía. Estas soluciones incluyen, entre otras, “Reputación de marca y respuesta de crisis”, “Family Offices, protección de la junta directiva y de los ejecutivos de nivel C”, “Investigación de mercado, análisis de la competencia e inteligencia comercial” y “Debate político, opinión pública,

polarización y análisis de disturbios sociales”, último caso ajustado a aquello que ocupó a la empresa en Santiago a fines de 2019 y en cuya descripción detallan que “Alto identifica rápidamente cualquier comportamiento coordinado o no auténtico destinado a alimentar los disturbios sociales, la polarización o la contaminación del debate público” (Alto Analytics, s.f.).

Las “soluciones” que ofrece Alto en su sitio nos permiten hacernos una idea del tipo de clientes al que apunta la compañía. Como es frecuente en el sector, hay una importante orientación hacia el marketing, pero también dirigida a cubrir los intereses de actores poderosos, desde altos ejecutivos hasta instituciones multinacionales y actores políticos. En una entrevista televisiva concedida a RTVE, el CEO de Alto Analytics se refiere en términos generales al impacto que tiene su sector en la política, señalando que sus servicios permiten a los clientes acceder a “una capacidad de entender de forma muy precisa cuál es el debate público. Pero no solo de forma muy precisa, sino de forma muy segmentada” (Alto Analytics, 2016).

### **El contenido del informe Big Data**

Como los expertos hicieron notar rápidamente, cuando una empresa como Alto Data Analytics promete reconocer “cualquier comportamiento coordinado” dirigido a “alimentar los disturbios sociales”, pese al lenguaje hiperbólico, no se refiere a precisar delitos y actividades de insurgencia ni a identificar a sus responsables, más allá de las posibilidades de un análisis sobre los intercambios de carácter público, por muchos “puntos de datos” que registre. Y aunque resulta imposible conocer exactamente qué productos y servicios ofreció José Luis Cansado cuando estuvo en Chile, más allá de lo que prometía Alto Analytics a través de su sitio web corporativo y en las entrevistas a su fundador, sorprende mucho que el Ejecutivo haya tratado públicamente al informe Big Data como un documento de inteligencia. Por ejemplo, Karla Rubilar, vocera de Gobierno en ese momento, afirmó que el reporte presentaba “información clara de organizaciones desde fuera de Chile para incentivar la protesta violenta” (Navarro, 2019). Es llamativo darle este peso a un monitoreo de redes sociales, sobre todo considerando que, como hemos expuesto, este sector y sus productos (así como, sus limitaciones) no eran una novedad ni para Piñera ni para su director de Secom.

Pero, ¿cuál era el contenido preciso de este documento?

La principal fuente para conocer el contenido del informe fue la filtración de este al diario La Tercera y los detalles que el medio entregó sobre él en su edición del domingo 22 de diciembre de 2019. Estos detalles no fueron desmentidos por el Gobierno, que tampoco cuestionó o intentó perseguir a la o los responsables de la filtración de dicho informe a La Tercera y la publicidad de esta información, pese a que se podría haber vulnerado la obligación de reserva que impone el artículo 38 de la Ley N.º 19.974 (misma norma a la que alude Galli al negarse a transparentar dicho documento).

Es importante indicar que en el artículo de La Tercera, que hasta el día de hoy (y hasta donde sabemos) es el único actor que ha tenido acceso al informe fuera de la Presidencia, del Ministerio del Interior y de los participantes de la reunión del 2 de diciembre, se señala sin matices que dicho documento fue elaborado por el Gobierno. Asimismo, se expresa con total claridad que la descripción que se hace corresponde al informe entregado por Gonzalo Blumel al fiscal nacional el día 19 de diciembre (Ayala, 2019) (y no, por ejemplo, al documento que supuestamente le entregó el director de la ANI al fiscal nacional el 20 de diciembre).

El reportaje caracteriza al informe Big Data destacando los siguientes aspectos:

1. Se afirma que dicho informe constituye un “análisis big data” de casi cinco millones de autores únicos y sesenta millones de comentarios.
2. El registro de dichos comentarios se desarrolló entre los días 18 de octubre y 21 de noviembre de 2019. Al parecer, el informe se concentra casi exclusivamente en actividades registradas en Twitter, aunque de manera bastante llamativa no se indica en el artículo.

3. El informe afirma que 31 % de los autores y 19,3 % de los comentarios considerados (no se indican criterios de inclusión y exclusión) “está localizado” fuera de Chile.
4. El grueso del informe, al menos según el reportaje, se ocupa en segmentar a los diversos grupos que contribuyeron a la conversación sobre el estallido social en el periodo analizado.
5. Los autores del reporte identifican cinco grupos de opinión: (a) Cuentas favorables a las movilizaciones, donde destacan a medios como Actualidad RT y TeleSur y a artistas extranjeros como Ismael Serrano y Juan Diego Botto; (b) Perfiles de Venezuela, Argentina, México y España contrarios a las movilizaciones y que compartían alusiones a una supuesta intervención bolivariana en Chile, donde destacan personajes como Patricia Bullrich, Jorge Lanata y María Corina Machado; (c) Los jóvenes, señalado como el grupo mayoritario y favorable a las movilizaciones (no se explica cómo se obtuvo la información etaria de las cuentas, considerando las dificultades que ofrece Twitter para ello). Se afirma que estos perfiles eran “aficionados al K-Pop”, que “registraban bajas referencias a asuntos políticos y sociales” antes del estallido (no se explica cómo se llegó a esta conclusión retrospectiva sobre varios miles o cientos de miles de cuentas, considerando que el análisis considera un registro iniciado solo a partir del día del estallido) y que estaban “altamente influenciados” por medios digitales como Piensa Prensa y ChileOkulto (tampoco se precisa cómo es posible deducir esa influencia); (d) Grupo de “alto impacto”, medido en cantidad de seguidores, que mencionaban la desigualdad como causa de las movilizaciones. Se les apunta como “principales articuladores” del discurso contra Piñera y entre ellos se incluye a futbolistas (Claudio Bravo y Gary Medel), figuras políticas (Pamela Jiles, Camila Vallejo, Karol Cariola, Beatriz Sánchez), periodistas y la cuenta anónima de Twitter @vagoilustrado; y (e) Perfiles que defendían al Gobierno con un impacto mucho menor que los perfiles críticos, entre los que se menciona a figuras de la extrema derecha como Teresa Marinovic, José Antonio Kast y activistas anónimos, así como a cuentas oficiales

de las instituciones de orden y seguridad. En sus mensajes, acusaban a los políticos de izquierda de “no condenar la violencia”.

6. Respecto a la actividad de “bots o usuarios con actividad anómala”, La Tercera señala que el documento identifica a más de once mil de estos perfiles y que “241 de ellos generan 194 comentarios al día”. Los autores del informe sitúan al 52 % de estas cuentas dentro del tercer grupo, la comunidad juvenil.
7. Fuera del ámbito (aparentemente) restringido de Twitter sobre el que se concentra el grueso del reporte, se destaca que varios de los grupos de Facebook usados para organizar manifestaciones durante el periodo, como “Protesta por Chile en Londres” y “Support Chile from Glasgow”, fueron creados desde el extranjero.
8. El informe también menciona que tras el inicio de las manifestaciones se generaron 86 grupos de difusión nuevos en Telegram, 169 grupos en WhatsApp y 338 eventos nuevos en Facebook.

Como vemos, los grupos de usuarios o perfiles descritos por el informe no son entidades comparables, no configuran grupos demográficamente discretos y las características que los delimitan resultan más bien ambiguas y poco reproducibles. Se destaca la falta de aclaraciones sobre la metodología utilizada para establecer categorías geográficas y demográficas, como la edad de las personas usuarias, ya que Twitter no exige la entrega de tales datos y cuando son aportados, lo que hace solo una pequeña cantidad de cuentas, no son en absoluto confiables. Es posible, desde luego, que el reportaje de La Tercera arroje una mirada muy parcial del informe (puesto que condensa supuestamente 112 páginas en poco más de 1600 palabras) y que, en pos de destacar los aspectos más curiosos o noticiosos del documento, haya pasado por alto dar cuenta de los criterios, metodologías y otros elementos que permitirían entender mejor el valor del análisis presentado.

Es llamativo, asimismo, que el informe se concentre de forma tan considerable en buscar y describir actividades en el extranjero, justamente el tipo de vínculo que había perseguido con tanto ahínco el Gobierno durante las semanas previas, pese a que Alto Analytics desarrolló el trabajo de manera gratuita y no solicitada.

Es difícil establecer la relevancia de la cifra entregada por el reporte que afirma que un 31 % de autores de publicaciones (al parecer, en Twitter) estarían ubicados en el extranjero, dado que desconocemos las características del conjunto de los datos considerados en el análisis, así como los criterios y herramientas usados para situar geográficamente a dichos perfiles. Asimismo, tampoco sabemos qué proporción de ese 31 % que tanto se destaca, al menos en el reportaje de La Tercera, corresponde a perfiles críticos al Gobierno de Chile, ya que el informe también habla de un grupo de cuentas foráneas dedicadas a defenderlo y difundir noticias falsas sobre supuestos complots bolivarianos.

Los demás antecedentes que hablarían de una influencia extranjera tampoco parecen cruciales. El medio venezolano TeleSur está asociado desde mucho antes del estallido social en Chile a una línea editorial crítica con la derecha latinoamericana. Medios estatales de países no occidentales, como Actualidad RT (Russia Today), tienen comúnmente una mirada muy crítica sobre los procesos en Occidente, particularmente de aquellos gobiernos aliados a EE. UU. Sin duda es posible etiquetarlos como aparatos de propaganda, pero nada de eso puede ser considerado como un descubrimiento o una sorpresa. Mientras tanto, el punto hecho sobre que los grupos de Facebook que organizan protestas fuera de Chile fueron creados fuera de Chile no conduce a ninguna conclusión que no sea la evidente: los miembros de las colonias chilenas que organizan reuniones en sus respectivos países viven generalmente en esos países.

### **Las críticas al informe Big Data**

La presentación del informe Big Data fue todo lo errática que podría haber sido, pero la recepción pública a la estrategia del Gobierno empeoró todavía más al conocerse los detalles filtrados por La Tercera y nunca desmentidos por La Moneda. Mientras una amplia gama de

políticos, principalmente de oposición, pero también de Gobierno, criticaron la forma y el contenido de lo que se estaba dando a conocer, en la prensa y en las redes sociales cundió la ironía. Para la opinión pública, el informe aparecía, en el mejor de los casos, como intrascendente y, en el peor, manifestaba un desconocimiento profundo no solo del contexto nacional en el que se situaba, sino de las propias dinámicas de las plataformas digitales en la que fundaba su análisis. Tanto las personas usuarias de redes sociales como los especialistas veían en esta ofensiva del Gobierno no tanto mala intención como una ignorancia profunda de los fenómenos propios de Internet y de la tecnología de la que pretendían servirse.

Muy pronto se dio la paradójica situación de que nadie en el Gobierno se quería hacer cargo del informe que en palabras del ministro del Interior y Seguridad Pública era “información extraordinariamente sofisticada” y fundamental “para la investigación penal” (CNN Chile, 2019b). Trascendió a la prensa que desde la Secom se desligaban por completo y señalaban no conocer el documento. El 23 de diciembre, incluso Blumel declinaba hacer cualquier comentario sobre el asunto (Muñoz y Catena, 2019) y el propio Sebastián Piñera no volvió jamás a atribuirse o a atribuirle a su Gobierno la autoría de ese u otro reporte de similar naturaleza. Poco después, en medio de este clima y sin que se aclararan los motivos, renunciaría Jorge Selume a la dirección de la Secom (Catena, 2020).

En paralelo a esta crisis política en el Gobierno, fueron muchos los expertos y académicos que durante los últimos días de diciembre de 2019 y comienzos de 2020 se pronunciaron en alguna medida sobre los aspectos conocidos del informe. En esas intervenciones, buscaban aclarar ante qué tipo de análisis estábamos, en qué consistía su “sofisticación” y qué tanto tenía de *big data*, así como si permitía dar sustento a la tesis de la intervención extranjera que el Ejecutivo estaba tan empeñado en defender.

Uno de los detalles más comentados por los especialistas tenía relación con el uso adecuado del concepto *big data*. Con tal concepto se hace referencia, en términos generales, a “grandes cantidades de datos generados por operaciones computacionales de gran escala con el propósito de analizar y predecir el desarrollo de ciertos aspectos de la sociedad o la naturaleza”

(Fuchs, 2021). César Hidalgo, director del Centro de Aprendizaje Colectivo del Instituto de Inteligencia Artificial y Natural (ANITI) de la Universidad de Toulouse, entrevistado en diciembre de 2019, señalaba que “Básicamente es una palabra amplia que se usa para indicar todos estos datos operacionales que se usan para fines analíticos, aunque su uso primordial sea operacional” (Aguirre, 2019). Su rasgo fundamental tiene menos que ver con el tamaño de los conjuntos de datos implicados que con la capacidad de búsqueda, adición e interconexión de los elementos dentro de estos conjuntos (boyd y Crawford, 2012). Conjuntos cuantiosos de datos de diferentes características son susceptibles de convertirse en *big data* cuando se someten a procesamientos analíticos computacionales que no pueden ser afrontados con herramientas tradicionales.

Según Hidalgo, el llamado informe Big Data consideraba una aproximación menos sofisticada: “lo que hace es muy común: separar los temas en ‘burbujas’, y lo que uno puede hacer con estas redes de comunicación es aislar cada uno de estos grupos, identificarlos y tratar de encontrar cosas comunes con características, demografía y si tuitean sobre un tema específico. Eso permite segmentarlos de forma similar a la segmentación de mercado”. Esto difiere, desde luego, de los análisis de grandes datos que se emplean habitualmente en la gran industria o en la banca, donde existe un procesamiento en tiempo real de datos cuyo volumen, velocidad y variedad es muy superior al tipo de análisis que empresas como Alto Data Analytics desarrollan sobre los datos públicos de Twitter. Por ello, cuando se habla de minería de datos de redes sociales no siempre se alude a *big data* y muchos especialistas prefieren conceptos como el de *big social data*, que refiere a esa parte del *big data* que se relaciona con personas o que describe sus comportamientos e interacciones mediadas por la tecnología en ambientes digitales (Olshannikova, 2017). En este ámbito no es tan importante el tamaño del conjunto de los datos como el tipo de procesamiento y sus fines.

Es por esto que, en referencia al uso del concepto por parte del Gobierno, el académico Jorge Pérez fue tajante: “El término ‘Big data’ no lo uso en ninguna de mis clases. Nosotros decimos: cuando no tienes nada que decir, llenas con big data. Es como un elemento de marketing” (Arce, 2019).



En el periodo en que el informe Big Data era tema obligado de conversación en la prensa, los expertos coincidían en que aquello que se había sido presentado como algo de gran sofisticación era, por el contrario, un tipo de análisis comparativamente simple y relativamente habitual. El problema radicaba en el foco y las expectativas que se habían puesto, por parte del Gobierno o de la empresa que lo había desarrollado, sobre los indicios que supuestamente arrojaba el reporte. El mismo Jorge Pérez indicaba las razones de por qué el principal uso de estos monitoreos son las campañas de marketing: “No es que desde la red social hagas injerencia para que pase algo en el mundo real, lo que hacen es mostrarte, darte una pista parcial y súper sesgada de lo que está pasando afuera” (Arce, 2019). Pérez coincidía con muchos expertos sorprendidos del intento de desarrollar un trabajo de inteligencia usando las comunicaciones de Twitter. Por ejemplo, la investigadora Bárbara Poblete, experta en análisis de redes sociales, señalaba que “en general, cualquier dato que se saque a partir de redes sociales es solo información y esto se transforma en inteligencia al ser contrastado con fuentes externas a la red social” (Siebert, 2019). El decano de la Facultad de Comunicaciones de la Pontificia Universidad Católica de Chile, Eduardo Arriagada, hablando con Radio Futuro, complementaba este punto al indicar que el informe Big Data “está lejos de ser de inteligencia, es un análisis de mucha información, que es la gracia que tienen las redes” (Puebla, 2019). “Este tipo de análisis lo puede hacer cualquier persona porque existen cientos de herramientas para análisis de texto, de redes sociales y cosas para descargar estos datos”, añadía Poblete en otra entrevista. “Esto está al alcance de cualquier persona y los recursos computacionales también están al alcance de cualquiera” (Muñoz y Catena, 2019).

La falta de entendimiento que transmite el informe sobre aspectos que son básicos para los especialistas y que se refleja en las actitudes y observaciones hiperbólicas que Piñera y su Gobierno mostraron sobre él, fue un aspecto que varios expertos destacaron con inquietud. Poblete manifestaba su preocupación respecto a que “esto pueda ser tratado como una fuente absoluta de información. Extraer tendencias no significa que sea cierto en el mundo real, porque a través de las redes sociales se puede inventar lo que uno quiera” (Aguirre, 2019). John Atkinson Abutridy, Doctor en Inteligencia Artificial, indicaba que “si uno bucea más a fondo, se puede

percatar que se está tratando (sin entender) de explorar el lenguaje como generador de acción”, pero “contar palabras sin contexto, realizar estadísticas de cuentas de Twitter con muchos seguidores, u otra actividad en redes sociales, dista bastante de tratar de entender el lenguaje” (Abutridy, 2019). En tanto, Claudio Gutiérrez, Investigador del Instituto Milenio Fundamentos de los Datos, en una columna en CIPER, expresa que el escándalo en torno al informe Big Data “se enmarca en un olímpico desprecio por la rigurosidad científica de parte de quienes gobiernan hoy” (Gutiérrez, 2019).

Las extrapolaciones que habían emergido del informe escapaban a toda racionalidad, al menos desde la perspectiva de los expertos. Mientras el presidente parecía respaldar toda su tesis sobre la intervención extranjera en los “hallazgos” del informe Big Data cuando afirmaba que “hemos hecho un estudio de millones y millones de comunicaciones a través de redes sociales” (CNN Chile, 2019c), las afirmaciones de Gonzalo Blumel no dejaban dudas sobre lo que el Ejecutivo creía ver en el análisis de Alto Data Analytics. Pero para los académicos, la correlación que se pretendía entre comentarios en Internet y acciones delictivas en el mundo real, entre que haya un porcentaje de *tweets* generados desde fuera de Chile y que haya un complot extranjero que sería el causante de las revueltas de octubre, no tenía sentido, manifestaba ignorancia y era irresponsable.

Poblete se expresaba a este respecto señalando que “lo que se presenta en el artículo, no muestra un análisis con la profundidad necesaria, ni con el contraste de otras fuentes como para saber que existe una influencia extranjera” (Siebert, 2019). En la misma línea, el académico Eduardo Engel prevenía: “Que un medio chavista o ruso comente de manera recurrente y obsesiva eventos en Chile no significa que tengan responsabilidad alguna en ellos” (Engel, 2020).

Para quienes habían estado analizando la actividad en las redes durante el estallido social, el informe contradecía algunas de sus observaciones más básicas. Antonio Díaz, director de Espacio Público, señalaba que “hacer un estudio de esta magnitud en el calor de las protestas y los incidentes, es poco útil. Porque no tienes un proceso definido para ir limpiando los datos, que está verificando que la información es la correcta” (Muñoz y Catena, 2019). Bárbara Poblete

expresaba que había visto “varios intentos de hacer pensar que existe influencia extranjera. No lo puedo descartar a futuro, pero no lo hemos podido corroborar en nuestros datos. Es muy fácil falsear la información que aparece en Twitter” (Aguirre, 2019).

En consecuencia, múltiples expertos no tuvieron ambages en calificar el informe como defectuoso o poco serio. En muchos casos, los especialistas apuntaban a evidentes problemas de método al momento de limpiar y manipular los datos. Uno de los aspectos más problemáticos del documento es la sencillez con que despachaba el problema de la geolocalización:

Yo puedo falsear la ubicación, puedo decir, por ejemplo, que estoy en Venezuela y no lo estoy. Eso lo puede hacer alguien que tiene un manejo computacional muy simple, entonces ya el hecho de estar buscando cuentas ubicadas en el extranjero, que declaren que están ubicadas en el extranjero, es complicado porque no hay ninguna forma de asegurar que estén ubicadas allá. (Siebert, 2019).

Para varios de los analistas no quedaban muchas dudas de que el informe Big Data era un documento elaborado ex profeso para apoyar una hipótesis cuya verdad ya estaba decidida y solo se necesitaba conseguir algunos números para ser ofrecida como real. Con la minería de datos de redes sociales “puedo confirmar prácticamente cualquier cosa que yo quiera, y encontrar evidencias de lo que quiera; es información muy manipulable”, explicaba Poblete (Aguirre, 2019). “El Gobierno se metió en una camisa de fuerza. Deslizaron esta hipótesis de intervención extranjera y no les quedó más que encontrar quien les corroboraba esa hipótesis” (Arce, 2019), sentenciaba Pérez. Encontrar una empresa de monitoreo de redes sociales que hiciera ese trabajo a la medida no sería ningún problema, afirmaban ambos académicos: “Les dijimos que habría empresas que les dirían sí, que no comprarán a alguien que diga que sí se puede, porque no se puede” (Arce, 2019); “Ellos (el Gobierno) lo iban a encontrar; y sino, iba a existir alguien que lo quisiera vender. Hay muchas personas que lo pueden hacer. Y así ocurrió” (Aguirre, 2019).

## **Factores problemáticos de las agencias de monitoreo de plataformas digitales**

Si hemos abundado en los detalles del informe Big Data es porque constituye una muestra de la situación actual del sector de las AMP y los desafíos que plantean en el ámbito de la comunicación política. Algunos factores problemáticos a destacar provisionalmente son la opacidad en la que se mueven estas compañías; la completa ausencia de responsabilización de estas y de sus clientes respecto a la factualidad o racionalidad de los productos infocomunicacionales que comercializan y ponen en conocimiento para el debate público; la fe ciega o interesada de los clientes de estas empresas, inclusive cuando se trata de tomadores de decisiones claves para la convivencia democrática; y el desconocimiento generalizado, por parte de la opinión pública y actores centrales en el debate democrático, como la prensa, de lo que podríamos llamar las “circunstancias” (más propiamente: las relaciones sociales de producción) que caracterizan a las AMP, su mercado, sus trabajadores, sus clientes y sus productos.

Este caso, que durante algunas semanas y de forma extraordinaria puso en el centro de la atención mediática a una empresa dedicada a la minería de datos, nos entrega algunos indicios preocupantes del rol y la relevancia que puede haber adquirido este sector, relativamente desconocido para la mayoría de los ciudadanos, entre los tomadores de decisiones políticas en Chile. Para el ámbito de la comunicación política es relevante saber si las características conocidas del informe Big Data son un estándar dentro de la oferta del sector AMP y entender hasta qué punto los equívocos y autoengaños cometidos por el Gobierno de Sebastián Piñera son también posibles en (o constituyen parte de) las relaciones cotidianas entre estas compañías intermediarias y sus clientes habituales. También sería significativo conocer si el sonado caso del informe Big Data dejó alguna lección apreciable en el sector.

El objeto de este estudio son las AMP, sus productos y sus relaciones de producción. A través de una aproximación exploratoria y descriptiva se esbozará una caracterización de este campo, que pese a la crisis del informe Big Data y la amplia cobertura mediática que tuvo, se conserva hasta el momento relativamente opaco de cara a la opinión pública.

Para entender las problemáticas a las que se enfrenta este sector y las que deberían tener en cuenta sus clientes (dentro de los que se incluye organizaciones privadas y públicas), debemos conocer los servicios o productos que ofrecen, así como qué promesa se hacen a través de dichas mercancías, qué expectativas sugieren a sus clientes. En pos de este objetivo, este trabajo se enmarca en términos generales dentro de lo que se conoce como perspectiva crítica de la investigación, en oposición a la investigación administrativa. El fundamento crítico de estas aproximaciones no se dirige solo o principalmente a los medios y sus efectos, sino que a la sociedad que los origina, propaga y sostiene (Jeřábek, 2017). Abordaremos la investigación desde este enfoque general y, en particular, desde la economía política de la comunicación y la información (EPC), el estudio de las relaciones sociales que recíprocamente ayudan a dar forma a la producción, distribución y consumo de bienes, incluyendo mercancías de comunicación (Mosco, 2009). Más adelante profundizaremos en este marco conceptual.

## **Sobre las mercancías analíticas del sector AMP**

En términos preliminares, el elemento que agrupa en un mismo sector los productos ofrecidos por las AMP es el lugar que ocupan dentro de una cadena productiva que deriva de lo que comúnmente se denominan redes sociales. Alegando tener propiedad sobre los datos (Puschmann & Burgess, 2013), este tipo de plataformas digitales otorgan diferentes grados de acceso a ellos a través de sus API (*Application Programming Interfaces*). Es en este nivel donde se establecen las AMP como Alto Data Analytics. Los datos extraídos o minados por esa vía son analizados mediante software y por profesionales especializados que elaboran diferentes tipos de monitoreos, informes o “entregables”. Llamaremos a estos productos mercancías analíticas. Los consumidores de estas mercancías son empresas, gobiernos y tomadores de decisiones a todo nivel.

¿En qué consisten dichas mercancías? Básicamente, se trata de reportes sobre las actividades de las personas usuarias (o un subconjunto de estas) según ciertas “métricas”, como el volumen de menciones, el “alcance” (cuántos usuarios “vieron” una publicación), los

“sentimientos” asociados a sus publicaciones y otras variables en un periodo de tiempo. El conjunto de estos factores es valorado en busca de *insights* (perspectivas, conocimientos) sobre su significado con base en criterios cuantificables para entregar algún tipo de interpretación. Estas pueden tener diferentes alcances, desde estimar la percepción sobre una marca que tiene un grupo de consumidores (¿qué opinan las mujeres de 18 a 25 de clase alta de sobre la nueva campaña de la marca Tampax?) hasta sugerir explicaciones de procesos sociales y políticos (como en el caso del informe Big Data). Estos informes suelen estar presentados en formas visualmente atractivas y contener algunas explicaciones conceptuales básicas, en el entendido de que sus lectores no son expertos. Bajo ese mismo entendido, y de forma problemática como hemos observado en el caso del informe Big Data, suelen no profundizar en la metodología.

Es común que estos documentos concluyan con una serie de recomendaciones de acciones a tomar por parte del mandante. En el caso del informe Big Data, en dicha sección se sugería que era fundamental crear un “lago de datos de análisis continuo”, asignar a un grupo permanente de analistas que desarrollen un monitoreo constante y así generar “un flujo continuo de inteligencia” (Ayala y Ojeda, 2019). En otras palabras, proponía al Estado contratar de forma permanente los servicios de Alto Data Analytics.

La relevancia creciente de las mercancías que ofrecen las AMP como fuente de información sobre la sociedad (Lipschultz, 2019) hace pertinente un estudio en profundidad del sector que las produce. La importancia que para ciertos actores tiene el *big data* como fetiche que brinda legitimidad a la toma de decisiones políticas y estratégicas, no parece apoyarse siempre en fundamentos muy sólidos (como quedó patente en el caso del informe Big Data). La posibilidad de aproximarse a partir de la información de las redes sociales a comprender en algún grado los fenómenos sociales es un asunto no resuelto en el debate académico. Se ha indagado en cómo el análisis de las plataformas puede permitir generar conocimiento sobre los comportamientos sociales (Rogers, 2013), pero hasta hoy pocas investigaciones sobre las actividades *on line* pueden pretender razonablemente hacer afirmaciones sobre las conductas humanas y llegar a conclusiones extrapolables al mundo *off line* (Lazer et al., 2021), como pretendían los promotores del informe Big Data. Sin embargo, la forma en que la industria de la minería de datos en Chile y

el mundo ofrece sus servicios y promociona sus productos, el lenguaje hiperbólico que emplea y la posición de preeminencia social que ha alcanzado el sector tecnológico en las últimas décadas (Williams y Gilbert, 2022), podría llamar a la confusión de ciertos actores, así como de la opinión pública.

Aunque de momento desconocemos de qué manera precisa y hasta qué punto los servicios de las AMP han influido y están influyendo en la toma de decisiones políticas en Chile, observamos que el sector lleva un buen tiempo siendo validado por importantes actores de la élite, incluyendo influyentes consultoras, grandes grupos empresariales y diferentes reparticiones gubernamentales. Sin un escrutinio público adecuado sobre la validez de sus perspectivas y las particularidades de su mercado, los productos de esta industria pueden estar contribuyendo a la cosificación en espacios de poder de ciertas concepciones sobre la sociedad reflejo de sesgos o intereses particulares, lo que puede afectar en último término el ejercicio democrático.





## ANTECEDENTES

En la introducción, usando como ejemplo paradigmático el caso del informe Big Data, hemos buscado poner el énfasis en aquellos elementos que han mostrado ser problemáticos, o que podrían serlo, relacionados con la comercialización de monitoreos y análisis de datos de redes sociales, en pro de explicitar por qué estas agencias deben ser de interés público.

En este capítulo abordaremos los antecedentes de esta investigación. En primer lugar, en la sección titulada *El sistema de las plataformas: conceptos centrales*, nos ocuparemos de trazar en un sentido amplio el contexto sociotécnico e histórico en que se emplazan estas empresas, para situarlas en relación con el resto de los elementos y actores que constituyen lo que en ocasiones se ha denominado capitalismo de plataformas (Srnicek, 2018) o capitalismo digital (Schiller, 1999), entre otras denominaciones. En ese apartado, además, definiremos y acotaremos los conceptos principales que usaremos para referirnos a los diversos elementos del sistema sociotécnico y económico que aborda este estudio.

En la segunda parte de este capítulo, *Agencias de monitoreo de plataformas*, sintetizaremos las características y cualidades conocidas del sector comercial que buscamos caracterizar y que hemos podido recabar a partir de diversas fuentes, incluyendo el caso expuesto en la introducción. Es importante indicar que las investigaciones referidas específicamente al tipo de actores que cubre el presente estudio son escasas y que, en lo que respecta a Chile, esta sería la primera aproximación sistemática a esta categoría de empresas.

### **El sistema de las plataformas digitales: conceptos centrales**

Como hemos podido apreciar a partir de nuestro análisis del caso del informe Big Data, el sistema en el que se inserta la minería de datos de redes sociales es un área no solo tremendamente fértil en la creación y adopción de nuevos conceptos, sino que además estos

conceptos suelen manifestar un carácter polisémico. Dada esta proliferación de términos que desde diversas áreas se han propuesto para dar cuenta de los distintos fenómenos sociales, tecnológicos y comerciales que han emergido en torno a Internet, es absolutamente necesario indicar a qué nos referimos y a qué no cuando hablamos de estos asuntos.

### **La situación de Internet a fines del primer cuarto del siglo XXI**

Como afirma Ben Tarnoff en *Internet for the People* (2022), Internet no solo tiene una materialidad y una historicidad, también es la manifestación de una política (p. xi) y en la actualidad, esa es probablemente su faceta más preeminente. Internet emerge de cara al gran público como uno entre tantos otros elementos de consumo tecnológico que la maquinaria capitalista arrojaba sin tregua a fines del siglo XX, junto al CD-ROM, diversos formatos de audio y video digital, consolas de videojuego, teléfonos móviles, cámaras digitales, servidores abiertos (como los BBS) y un largo etcétera. Un pasatiempo para iniciados que prometía, simultáneamente, una revolución en las comunicaciones, la información, el consumo cultural y los negocios. Gozó durante muchos años de una imagen positiva: a fines del siglo XX e inicios del siglo XXI el tecnooptimismo estaba de moda y muchos entusiastas vislumbraban la gestación de un nuevo tipo de sociedad en la red o a partir de ella.

Pero hacia mediados de la década de 2010 se consolida un cambio en la percepción mayoritaria sobre Internet, que pasa de ser considerada la “respuesta” para crear un mundo más democrático e igualitario (Keen, 2015) a que se comience a afirmar, incluso en espacios que operaron por décadas como evangelizadores de sus virtudes, que Internet estaba “rota” (Baker, 19 de diciembre de 2017). Las muestras de aquello emergían por dónde se mirara. La promesa, mantenida con entusiasmo al menos hasta la crisis de las puntocom respecto a que la red favorecería un capitalismo más justo, se vio contrastada con el peso de una realidad en que el clima de Internet, en lugar de fomentar la libre competencia, había permitido la creación de los mayores y más intrusivos monopolios que ha visto la humanidad.

Por otra parte, las expectativas puestas en el papel de un “periodismo ciudadano” que podría minar o equilibrar el poder de corporaciones mediáticas, pronto fueron más que defraudadas (Curran, 2014). Lejos de democratizarse la información o de generarse un perfeccionamiento del sistema de los medios, en 2024 la crisis del periodismo lleva años profundizándose a escala global. Incluso las grandes corporaciones de medios que, con todo, debían sostener ciertos estándares, han sido desplazadas como protagonistas del entorno informativo por los gigantes tecnológicos que, comúnmente y gracias a un desproporcionado lobby, no le rinden cuentas a ningún Estado. Desde la perspectiva de su rol en el fortalecimiento de la democracia y la transparencia de las instituciones, Internet se ha torcido hacia un panóptico en el que diversos grupos de interés pueden recopilar información y vigilar a los ciudadanos, mientras estos, frecuentemente sin siquiera saberlo, se encuentran a merced de carteles digitales (Bloom, 2019).

No obstante la existencia de un consenso en que Internet se ha roto (y esto quiere decir que, en amplios sentidos, es nociva para la sociedad), ello no ha mermado su presencia predominante en nuestras vidas. Por el contrario, como argumentan Gilbert y Williams (Williams y Gilbert, 2022), las prioridades del libertarismo de Silicon Valley y la penetración ubicua de la computación en la vida cotidiana en los términos impuestos por dicho bloque, incluyendo una carencia casi completa de regulación para las plataformas digitales por parte de los gobiernos, son indicios claros de que vivimos en un mundo en que el sector tecnológico, aliado al financiero, ha triunfado en la batalla por la hegemonía. Esto redundo en que las soluciones propuestas para afrontar la ruptura de Internet rara vez apunten a la raíz del problema: que Internet es actualmente (pero que no tiene que serlo por necesidad) un negocio, que está dominado por corporaciones de interés privado y que funciona en persecución (exclusiva) de beneficios económicos.

Consideradas estas cualidades del sistema económico y social contemporáneo, desde la academia crítica han emergido una serie de términos para describir el sistema político-económico actual, etiquetas que se solapan casi por completo, pero que compiten buscando poner el foco en algún aspecto particular: capitalismo digital (Schiller, 1999), capitalismo de plataformas (Srnicsek, 2018), capitalismo informacional transnacional (Fuchs, 2012b), capitalismo de vigilancia

(Zuboff, 2018), edad del algoritmo (Louridas, 2020), Tecnofeudalismo (Varoufakis, 2024), entre otros. Los promotores de estos apelativos comparten la percepción de que el capitalismo se encuentra en una fase avanzada o tardía en que los modos de producción capitalista han sufrido cambios mediados por la tecnología y en particular por las tecnologías de la información y comunicación. Mejias y Couldry (2019), por su parte, enmarcan sus observaciones sobre el capitalismo con orientación hacia lo digital en un proceso mucho más amplio: la escala temporal para apreciar estos desarrollos debe considerar, señalan, el ciclo global de siglos en que el colonialismo ha estado entrelazado con el capitalismo. En este caso, el apelativo adecuado para enmarcar la economía política contemporánea sería colonialismo de datos.

Con todas sus diferencias, un aspecto relevante que comparten estos conceptos es que emergen en oposición a los enfoques discontinuos (Fuchs, 2012b) que han buscado describir la sociedad contemporánea como algo radicalmente nuevo, en que las relaciones sociales y de producción han alcanzado un estadio cualitativamente diferente. La perspectiva implícita en los conceptos mencionados es dialéctica y enfatiza la continuidad: la sociedad contemporánea es todavía y ante todo una sociedad capitalista orientada a la acumulación de capital. Aquí también enfatizaremos la continuidad dentro de la novedad que plantean. Esta toma de posición faculta el empleo de algunas categorías propias de la crítica marxista a la economía política del capitalismo, adaptando las herramientas y la naturaleza de las evidencias para el análisis de su fase actual.

### **La mercancía en el contexto de un capitalismo orientado a lo digital**

La producción generalizada de mercancías es una de las características centrales de la economía capitalista. Esta mercantilización extensiva hace referencia a la forma en que las fuerzas del mercado dan forma y remodelan la vida ingresando en áreas que, por razones sociales, políticas, culturales y económicas, habían sido dejadas fuera total o parcialmente de este proceso (Mosco, 2009). Esto incluye numerosas áreas de la información y la cultura, y su análisis desde los estudios de la comunicación y los estudios culturales han permitido cartografiar diversas formas en que el capitalismo moldea las instituciones y las relaciones sociales.

De manera que para describir cómo se despliega el capitalismo y cómo se desenvuelve el poder en un determinado entorno o incluso para desarrollar un análisis global de una particular configuración del capitalismo, es útil indagar en los procesos particulares de mercantilización e intercambio que están en juego y predominan.

Las y los autores que afrontan el problema de la mercantilización en el capitalismo orientado a lo digital ponen énfasis en diversos aspectos de las relaciones de producción que caracterizan a la economía fraguada en torno a Internet, como el contenido (Thrift, 2006), la privacidad (Sevignani, 2013), las comunidades online (Kent, 2012), la creatividad humana (Fuchs, 2010) y desde luego la atención y la audiencia (Caraway, 2011; Shepherd, 2013). Dado que lo que está en el centro del negocio de los monitoreos de redes sociales son los datos, aquí nos concentraremos en este aspecto.

Un prerequisite para la mercantilización extensiva en los nuevos espacios creados por Internet es, desde luego, la digitalización. Esto es la transformación de la información creada en una amplia gama de formatos para ser percibida a través de distintos sentidos a un único lenguaje que sea inteligible por las computadoras. Este lenguaje común es el código binario de ceros y unos al que audio, video, texto, sonido, imágenes, software, etc. es transformado para convertirse en comunicaciones electrónicas. Un aspecto central de la digitalización es que procesa y distribuye la información en paquetes discretos, cada uno con una dirección digital que lo identifica antes de ser transmitido. Es en el proceso de digitalización y mediante esta capacidad de adosar a la información diversos datos contextuales o metadata (datos que describen otros datos) que un valor de uso llega a convertirse en un valor de cambio, una mercancía. La digitalización amplía la mercantilización del contenido al ampliar las oportunidades para medir y monitorear, empaquetar y reempaquetar bienes culturales e información personal (Mosco, 2005).

En el contexto del capitalismo contemporáneo, la información personal que fluye por Internet de diferentes formas es ahora un recurso para explotar y mercantilizar (Bloom, 2019). La procedencia y naturaleza de estos datos es variada, pero una importante fuente la conforma el

conjunto de las plataformas tecnológicas conocidas popularmente, en español, como redes sociales y, en inglés, genéricamente como *social media*.

El modelo de negocios original de estas empresas reside, básicamente, en la generación gratuita, no remunerada, de contenidos por parte de las y los usuarios, en retorno por el acceso a la plataforma y sus servicios (Williams y Gilbert, 2022). Este contenido da vida a las plataformas y atrae más usuarios, que despliegan su atención e interactúan de formas diversas con tales contenidos y con otros usuarios. Las plataformas monetizan la atención de sus usuarios vendiendo espacio publicitario altamente segmentado. Simultáneamente, utilizan los datos generados por las interacciones y comportamientos que despliegan sus usuarios y usuarias dentro de las plataformas para mejorar su propio negocio, perfeccionando sus capacidades de mantener e incrementar su base de usuarios. Pero además, tras apropiarse de esos datos, los transforman en un producto en sí mismo para ser empaquetado de diferentes maneras y vendido a actores interesados. Más adelante nos concentraremos en definir las plataformas o redes sociales, pero primero es importante hablar de los datos o la data, como fenómeno social, técnico y económico.

## **Datos, nube y algoritmos**

Los datos son una forma de poder (Iliadis y Russo, 2016). En tanto tal, su registro es una obsesión de los gobernantes. Asimismo, como señala el filósofo canadiense Ian Hacking, clasificar a los sujetos por “tipos” es una necesidad imperial. El imperio clasifica a las personas que le tributan (Hacking, 2015), pero, para ello, primero debe contarlas. Es decir, acumular datos y atribuir metadatos. Para Hacking, la acumulación por sí misma de datos sobre las poblaciones es algo prácticamente inocuo, sin importar el tamaño de la recolección. Lo que resulta subversivo es que el acto de enumerar demanda la creación de tipos de cosas o personas que contar: el conteo tiene hambre de categorías (Hacking, 2007). Estas categorías no son sino metadatos.

Berhard Siegert (2015) analiza la impetuosa obsesión por la recolección de datos que el Imperio español desplegó en el siglo XVI sobre quienes accedían o pretendían acceder a sus dominios de ultramar, momento en que emergen instituciones dedicadas específicamente a hacer

frente a esa tarea (como la Casa de la Contratación de Indias) y se originan nuevas profesiones al alero de las nuevas tecnologías de registro (escribanos, contadores, cronistas). Hoy podemos interpretar esos novedosos procedimientos de la siguiente forma: al acudir a la Casa de Contratación en Sevilla con el fin, por ejemplo, de obtener un permiso para reunirse con un pariente en Chile, el “usuario” debía dar cuenta por primera vez frente a la autoridad de su existencia por escrito (cosa completamente inédita, a menos que se formara parte de la casta gobernante). En ese proceso, su existencia quedaba atada a lo que hoy podríamos considerar una serie de metadatos (limpieza de sangre, religión, edad, historial frente al Santo Oficio, estado marital, temperamento, características físicas) que determinaban sus posibilidades en la arquitectura del poder imperial.

El caso descrito por Siegert prefigura como, independiente de la época, las nuevas tecnologías que agilizan el registro y sofistican el procesamiento de datos son partícipes de la producción de nuevas instituciones y nuevas profesiones que realzan el afán por los datos por parte de aquellos que detentan o pretenden detentar el poder. Situaciones análogas podríamos delinear probablemente a partir de diferentes tecnologías a lo largo de la historia en diversos contextos sociales (pensemos, por ejemplo, en las instituciones, profesiones y los intereses que movilizó la radio). Hoy vivimos uno de esos momentos en que este proceso aparece como patente.

La observación de Hacking, por otra parte, nos empuja a mirar siempre a la acumulación de datos desde la perspectiva del biopoder: ¿quién clasifica a quién y con qué fines? Peter Bloom (2019) invierte esta pregunta para plantear el mismo problema, pero desde la perspectiva de lo que él llama sociedad de monitoreo: en un sistema de constante vigilancia de nuestros pensamientos, acciones y preferencias, ¿qué y quién no es monitorizado y por qué razón? La ausencia de datos es otro indicio de poder, el poder de no ser observado o de permanecer en la opacidad. La voluntad de no ser categorizado.

En la posición social que adquieren los datos en el capitalismo contemporáneo se aglomeran y entremezclan una serie de procesos entrelazados, diferentes pero de naturaleza

complementaria: la concepción del mundo como si estuviera constituido por magnitudes numéricas o estadísticas (la “avalancha de números” de Hacking), la “sublimidad” de lo digital (Mosco, 2005) y el “deseo por los números” (Kennedy, 2016), entre otros; todas maneras de expresar los ribetes de irracionalidad que se aprecian frente a un hambre por las mediciones concebidas como un “sustituto de la verdad” y como parte de la “creación de fenómenos” (Hacking, 1996).

### **La computación en la nube**

Una faceta de esto es que los datos son almacenados, procesados y distribuidos en volúmenes ingentes sin que se hayan formulado las preguntas que dichos datos deberían responder. Se espera que la propia acumulación haga emerger eventualmente respuestas o patrones. Esto ocurre a través de la llamada *cloud computing* o computación en la nube.

La computación en la nube, red de servidores remotos conectados a Internet, lejos de la evocación etérea que produce su nombre, es una parte muy tangible de la industria contemporánea de la informática y las telecomunicaciones, en términos del espacio físico que usan estos servidores y, sobre todo, del impacto energético que provocan. Una reciente estimación sobre este impacto indica que hacia 2027 solo los servidores dedicados al procesamiento relacionado con inteligencia artificial podrían llegar a consumir hasta 134 TWh (de Vries, 2023), lo que es aproximadamente un 170 % del consumo proyectado para Chile en 2023 (Comisión Nacional de Energía, 2022). A 2021, el consumo eléctrico de los servidores de Google era comparable a los de un país como Islandia o Puerto Rico.

Pese a encontrarse prácticamente por completo bajo control privado y estar constreñidas por criterios netamente comerciales, las premisas de la computación en la nube y su asociación con el *big data* se han impuesto como la principal, cuando no la única, forma de resolver un creciente conjunto de problemas, en desmedro de maneras previas de conocer y ver el mundo (Mosco, 2015). Estas premisas han supuesto un cambio radical en la forma en que concebimos socialmente la investigación y el conocimiento. El mito que alimenta esta manera de entender el



mundo es lo que Mosco denomina la “sublimidad” de lo digital, un sentimiento de asombro, admiración, temor y distancia psíquica, que domina las narrativas sobre las tecnologías nuevas y potencialmente transformadoras, pero complejas y difíciles de entender. Este fenómeno, lo vemos repetirse una vez tras otras (Mosco describe algunos ejemplos previos a la aparición de Internet, como el telégrafo, el teléfono, la radio y la televisión) y hoy lo podemos atestiguar con toda su fuerza frente a la irrupción de la inteligencia artificial.

## **Los algoritmos**

Esta forma de relacionarse con la tecnología ha tenido un particular impacto en la última década en relación con el *big data* y con el papel aparejado de los algoritmos en nuestra vida cotidiana. Los algoritmos son instrucciones para hacer algo de una manera específica. Los programas computacionales que están detrás de diferentes plataformas digitales contienen estas instrucciones en un lenguaje comprensible para que las computadoras puedan llevar a cabo determinadas tareas siguiendo los pasos establecidos.

Un algoritmo no produce sus resultados por un acto de revelación. Sabemos exactamente las reglas que ha seguido y los tipos de pasos que tomó. No importa qué tan maravilloso sea el resultado, siempre se puede trazar a sus operaciones elementales. (Louridas, 2020, p. 2)

Pese a que incluso los algoritmos más sofisticados y los grandes datos de los que dependen son al menos parcialmente comprensibles y susceptibles de ser analizados, persiste lo que algunos han llamado sublime algorítmico (Ames, 2018), que se expresa en la percepción de los algoritmos como una especie de entidad divina que se ubica fuera de todo ámbito de responsabilidad.

Así, desde hace más de una década y pese a las críticas y evidencias en contra, se ha consolidado como hegemónica la idea de que los números, aquellos que derivan del registro de

*big data* y su manipulación mediante algoritmos, son capaces de “hablar por sí mismos”, y estarían entonces libres de las ataduras que supondría una teoría o una filosofía.

Olvídese de la taxonomía, la ontología y la psicología. ¿Quién sabe por qué la gente hace lo que hace? La cuestión es que lo hacen y podemos rastrearlo y medirlo con una fidelidad sin precedentes. Con suficientes datos, los números hablan por sí mismos. (Anderson, 2008)

De esta forma de concebir la producción del conocimiento respecto al mundo emergen los parámetros mediante los que son recolectados, analizados y usados los datos en la actualidad. Muchos investigadores de diversas áreas que han incursionado en los estudios basados en *big data* se consideran involucrados en un tipo de “investigación básica” que no necesita atender a aspecto como la ética o inician sus análisis ignorando décadas de recolección de evidencias, así como las discusiones que se han dado en las ciencias sociales y humanas (Ames, 2018).

El sublime digital y algorítmico dan como resultado percepciones generalizadas, atestiguadas en nuestras políticas públicas y en la prensa, sobre la práctica inescrutabilidad de innumerables expresiones de la tecnologías que, dadas sus cualidades, serían imposibles de supervisar o regular y frente a las cuales solo nos queda acatar su poder.

Es esta pretensión la que quedaba de manifiesto cuando Gonzalo Blumel presumía que el Gobierno de Chile estaba en posesión de un “información extraordinariamente sofisticada, a partir de análisis con tecnología de *big data*” (CNN Chile, 2019b): con ello clausuraba todo debate. No hacía falta ninguna racionalidad adicional a dicha etiqueta mágica, *big data*, para poder asegurar que poseía algo similar a la verdad. E incluso cuando los detalles del informe fueron conocidos, cuando los expertos dieron cuenta de su inutilidad y de sus graves falencias metodológicas, el Gobierno insistía en que eran un análisis que merecía ser tenido en cuenta para la investigación. Después de todo, no era el Gobierno el que hablaba, eran los datos y los datos hablan por sí mismos.

## ***Big data y big social data***

Como ya hemos indicado anteriormente, con el concepto de *big data* se hace referencia a “grandes cantidades de datos generados por operaciones computacionales de gran escala” (Fuchs, 2021, p. 50). La alternativa recomendada en español por la RAE para este término es macrodatos, no obstante dado que tanto en el contexto académico como en el uso común se ha asentado el anglicismo, en este trabajo optaremos por hablar de *big data*. Este concepto tiene su génesis en el mundo corporativo de mediados del siglo XX, vinculado a la necesidad de nuevas herramientas para almacenar y realizar búsquedas entre cantidades cada vez mayores de datos de clientes y una evolución hacia un discurso centrado en el uso de la información con fines analíticos, específicamente para el desarrollo de modelos predictivos (Puschmann y Burgess, 2014), lo que da origen a la minería de datos. Actualmente el concepto comúnmente apunta a una amplia variedad de datos recopilados en volúmenes masivos a partir de complejos procesos computacionales aplicados a ámbitos como la minería, la agricultura, la manufactura, las finanzas, la ciencia, las interacciones en línea, la salud, el sector público, el retail, entre otras áreas e industrias.

Según boyd y Crawford (2012), en el uso cultural del concepto intervienen dos creencias interconectadas: que los datos producidos por los usuarios son datos “en bruto”, que son una forma primitiva de datos no refinados, y que los datos en grandes cantidades nos ofrecen un tipo de conocimiento más preciso y exacto que antes no era posible. Pero, ¿qué son “grandes cantidades”, qué significa “gran escala”, qué es lo “masivo”? Como han destacado muchos investigadores e investigadoras, *big data* es, en muchos aspectos, un término deficiente (boyd y Crawford, 2012). Los criterios que componen las definiciones habituales en este campo son altamente inestables dado el dinámico mundo de la informática y se hunden en la opacidad de un lenguaje tan especializado como propenso a los neologismos.

Una definición popular de *big data* ha optado por un enfoque que podríamos catalogar de nemotécnico, hablando de las “tres V” que caracterizan a esta *data*: velocidad, volumen y variedad (Chen, 2018). Algunos autores hablan de las cinco V al sumar veracidad y valor (Yu y

Guo, 2016). Incluso hay quienes postulan siete V (Uddin y Gupta, 2014) . Sin duda, es una increíble coincidencia que todas las características fundamentales de un fenómeno comiencen con la misma letra.

No obstante, ¿qué significan palabras como “velocidad” o “volumen” si no están atadas a una regla, a un estándar que nos indique qué es lo grande, qué es lo veloz, etcétera? Aquello que era grande o que era veloz en 2001 en cuanto al registro de datos, lo era a una escala que para 2011 ya resultaba muy modesta. Incluso aquello que se consideraba un procesamiento de datos “en tiempo real” hace unos años, actualmente se estima como una velocidad que no permite enfrentar los nuevos problemas y aplicaciones que los especialistas en *big data* se plantean.

Es importante destacar lo ambigua de esta definición, ya que en ella se alojan las posibilidades de un uso mañoso del concepto, o con meros fines promocionales. Para el científico de datos no cabe duda qué es *big data*: se trata de una categoría dinámica y la inscripción de un tipo de datos en esta categoría está más orientada por las cualidades y la complejidad técnica de su procesamiento (dirigido a ciertos fines) que por su tamaño. De estos aspectos emerge su complejidad como fenómeno. Así, *big data* es también (y, quizá, ante todo) un fenómeno cultural, tecnológico y académico cuyo rasgo fundamental tiene menos que ver con el tamaño o la velocidad de los datos implicados en un procesamiento dado que con la capacidad de búsqueda, adición e interconexión de los elementos dentro de estos datos en dicho procesamiento (boyd y Crawford, 2012). Estos elementos del *big data*, su relación con prácticas y técnicas, suelen quedar envueltos en poderosas metáforas usadas tanto de forma inconsciente como instrumental por periodistas, ejecutivos de empresas, legisladores y académicos, opacando su condición de constructos sociales y de las consecuencias políticas de ello (Puschmann y Burgess, 2014).

Como señala Vincent Mosco, *big data* es “un mito, una historia sublime sobre cómo conjurar la sabiduría no a partir de la inteligencia defectuosa de los humanos, con todas nuestras conocidas limitaciones, sino a partir de los datos puros almacenados en la nube” (Mosco, 2014, p.193). Este modelo de pensamiento encarna un positivismo digital que impone una forma de

conocimiento por sobre todas las demás, los números como intérpretes exclusivos de la realidad. Se trata de un enfoque compatible con el capitalismo, ya que hay una relación estrecha entre la cuantificación de todo y la mercantilización de todas las cosas (Mosco, 2016).

Así, en tanto fenómeno cultural, muchos aspectos relevantes para nuestro análisis quedan fuera de la mayor parte de las definiciones y usos del concepto *big data*, así como del *big social data*, término usado a veces para referirse a esa parte del *big data* relacionada con los comportamientos e interacciones entre personas en ambientes digitales (Olshannikova, 2017). El positivismo digital suele pasar por alto, como indica Fuchs (2017), las subjetividades, experiencias, normas, valores e interpretaciones de las personas involucradas en la generación y procesamiento de los datos. Los datos que circulan en las plataformas de redes sociales son creados por personas con intenciones totalmente ajenas a su uso como un recurso o mercancía. Dicha *big social data* es producto de la creatividad o del comportamiento humano, no algo “descubierto” por los dueños de las plataformas o por terceras partes, como sugieren algunas metáforas habituales (“Las empresas descubren un tesoro en los datos de internet” (América Economía, 2013).

### **La cocción cultural de los datos**

*Big data*, además, alude a un método, a una perspectiva epistemológica y una cultura que pocas veces se hace cargo de sus propias subjetividades, experiencias, normas y valores. Quienes invocan a los datos y el *big data* en nombre de la objetividad científica con frecuencia ignoran que los datos nunca están completamente “crudos” sino que siempre están “cocinados”. Como indica Gitelman (2013) “la recopilación y gestión de datos presupone interpretación”; en otras palabras, los datos siempre son producidos.

Esto, que puede parecer una perogrullada fenomenológica (todo procedimiento dentro de la cultura alberga las huellas de los seres humanos que se encuentran detrás) es más evidente y menos inmaterial a medida que nos interiorizamos en los procedimientos que desarrollan las empresas tecnológicas a partir del comportamiento de sus usuarios. En 2009, Google anunció la

implementación de una búsqueda personalizada asociada al historial de búsqueda y otros cincuenta y siete indicadores. Esta monitorización de los usuarios, como Google se preocupó de aclarar, no reconocía a sus usuarios, sino a sus navegadores (Raley, 2013). Desde luego, cincuenta y siete indicadores y el historial de búsqueda no constituyen una persona. Pero lo que admitía la empresa al dar cuenta de esta tecnología es que no estaba tan interesada en poseer a las personas como a sus “dobles de datos”, cuerpos espectrales de información (Adey, 2012), montajes parciales de los datos de los individuos que usan sus servicios, ensamblados de acuerdo a criterios de clasificación preestablecidos y algoritmos que les son útiles a las plataformas para sus fines corporativos. Así, las interpretaciones ulteriores que realiza Google sobre el comportamiento de los usuarios en sus servicios no se desarrolla sobre las personas o sobre los datos crudos que emergen de sus acciones en línea, sino en relación a estos “dobles de datos” (Ericson y Haggerty, 2006).

Fuchs (2017) aboga por la necesidad de comprender no solo lo hacen las personas en Internet, sino también por qué lo hacen, cuáles son las implicaciones más amplias y cómo las estructuras de poder enmarcan y dan forma a las actividades en línea. Lo mismo es válido para quienes dan forma a la arquitectura en la que esas personas se ven imbricadas, así como para quienes investigan los comportamientos y motivaciones de las personas en Internet. Particularmente aquellos y aquellas que realizan dichas labores fuera de los controles de la academia y que sin estándares públicamente conocidos dan forma a cuerpos de conocimiento para el mundo privado. Este último grupo es el objeto sobre el que esta investigación centra su atención.

Como recalca Helen Kennedy (2016), para examinar las formas variables en que el poder se construye y pone en práctica, es necesario centrarse no sólo en el poder de los algoritmos y cómo dan forma a los datos, sino también en las personas que interactúan en su vida cotidiana con las tecnologías digitales y el software que articula estas relaciones de poder.

## Redes sociales y plataformas

En el norte global, el concepto en inglés más usado para referirse a los sitios que en la actualidad concentran la mayor cantidad del tráfico en Internet, es el genérico *social media* y, especialmente en la literatura académica, los algo más acotados y un tanto obsoletos *social network sites* (SNS) y *online social networks* (OSN), términos que habitualmente operan como conceptos intercambiables. Históricamente, los tres hacían referencia a lo que originalmente fue popularizado como Web 2.0 (O'Reilly, 2005), una etiqueta de utilidad fundamentalmente propagandística e ideológica, usada para convencer a los inversores de que algo había cambiado radicalmente en Internet luego de la crisis de las puntocom (Fuchs, 2021). Bajo dicho apelativo se agrupaban algunos de los sobrevivientes de la debacle de las empresas de Internet que marcó el comienzo del nuevo siglo.

Timothy O'Reilly es un autor y editor que, con su compañía O'Reilly Media, ha publicado manuales técnicos desde hace décadas que han sido muy influyentes en el ámbito de las industrias tecnológicas. Respecto a la llamada Web 2.0, quizá su aporte más relevante fue poner foco en la posición de los servicios de Internet que eran exitosos, identificándolos como “plataformas”, negocios centrados en la gestión de bases de datos y ubicados estratégicamente entre los usuarios y los proveedores/servidores.

Similar a una llamada telefónica, que ocurre no sólo en los teléfonos a ambos extremos de la llamada, sino en la red intermedia, Google ocurre en el espacio entre el navegador, el motor de búsqueda y el servidor de destino, como un facilitador o intermediario entre el usuario y su experiencia en línea. (O'Reilly, 2005)

Las observaciones de O'Reilly relacionadas con las plataformas continuaron siendo válidas pese a que el concepto de Web 2.0 entró en desuso al ser solapado rápidamente en la jerga común por el de *social media*. En español, en tanto, fue el concepto de *social network sites* el que

se implantó a través de su traducción directa, redes sociales. Se trata de un concepto algo equívoco, particularmente en la literatura académica, donde tiene una larga historia en la psicología y la antropología social (Barnes, 1954), así como derivaciones que van desde la epidemiología a la economía (Wasserman y Faust, 1994).

La amplitud de estos términos y el dinamismo del capitalismo asociado a Internet han impedido una definición formal asentada, por lo que muchos investigadores optan por dar por sentado a qué se refieren cuando hablan de *social media*. Así, a partir de una muestra de artículos provenientes de más de mil revistas académicas relevantes para los estudios en comunicación y medios, Nau, Quan-Haase y McCay-Peet (2022) informaron que el 73 % no contenía ninguna definición y solo el 2 % contenía alguna definición general de *social media*. Por otro lado, Christian Fuchs recoge trece definiciones de *social media*, sin la intención de haber realizado una recopilación exhaustiva y sin decantarse por ninguna en particular (Fuchs, 2021).

En cualquier caso, a lo que apuntaban y apuntan esos conceptos es a describir sitios como Friendster (empresa fundada en 2002), MySpace (fundada en 2003), Hi5 o Facebook (ambas fundadas en 2004), plataformas digitales que, entre otras características, ofrecían la posibilidad de publicar e intercambiar contenidos creados por los usuarios (*User Generated Content* o UGC en inglés).

Al definir este tipo de sitios de Internet, el énfasis en la producción, cocreación y uso creativo de los contenidos disponibles es algo que han destacado varios autores y autoras como su aspecto central. Otros se han inclinado, tanto de forma crítica como descriptiva, por el análisis de los elementos sociales que se desenvuelven en estos espacios, la conexión, la participación y la misma atribución de “social” que se les aplica. Algunos académicos y académicas ponen el foco más lejos, en las condiciones materiales en las que se sustentan y operan estos sitios, como infraestructuras y ecosistemas dentro de un mercado, y sus efectos en la sociedad.

Junto a estas diferentes perspectivas con que se ha analizado el fenómeno, el rango de aquello que se entiende como *social media* no solo sigue siendo muy amplio, sino que además el



dinamismo de Internet dificulta cualquier intento de acotarlo. Comúnmente, las descripciones y categorizaciones de lo que es una red social incorporan plataformas tan distintas como Wikipedia y Twitter, pasando por Facebook, Reddit, Mastodon, Letterboxd, YouTube o Pinterest, sitios que tienen usos, tecnologías y filosofías muy diferentes.

Es inoficioso definir las plataformas *social media* por lo que ellas dicen ser, por sus funcionalidades o por el uso que le dan las personas. Ya que ellas no están en control de la mayor parte de los aspectos que determinan los modos en que las plataformas lidian con sus comunidades, con sus datos y con cómo se toman las decisiones que se toman, las plataformas sociales no pueden clasificarse de acuerdo a lo que las personas que las usan *quieren* hacer en ellas. Dado que nuestra investigación se acota a un sector comercial que emerge como una ramificación de aquello que hacen las plataformas sociales, categorías que nos permitan reconocer tanto sus cualidades infraestructurales como sus modelos de negocios nos serán de mayor utilidad.

Dejando a un lado el énfasis en los contenidos o las expectativas, podemos decir que una plataforma es una arquitectura digital programable diseñada para organizar las interacciones entre usuarios, incluyendo usuarios y usuarias finales (personas), entidades corporativas y organizaciones públicas (van Dijck, Poell y Waal, 2018). Las plataformas se orientan a la recolección sistemática de datos, su procesamiento algorítmico automatizado y la monetización de estos datos mediante diferentes modelos. Pese a que se plantean como intermediarios, son más bien mediadores, ya que “dan forma a la realización de actos sociales en lugar de simplemente facilitarlos” (van Dijck, 2013).

### **El sistema de las plataformas y sus características**

Es gracias y a través de la computación en la nube y el uso de variados algoritmos que opera, provee sus servicios y obtiene sus ganancias el sistema de las plataformas digitales. Llamaremos así al conjunto interrelacionado de empresas que actualmente dominan las comunicaciones mediadas por Internet y que configuran una red de infraestructuras digitales que

permiten la interacción entre usuarios, usuarias y grupos, poniendo a su disposición herramientas más o menos flexibles para generar diferentes tipos de transacciones de información.

Identificar o clasificar los diferentes tipos de plataformas que participan de este sistema en relación a sus propiedades no es un ejercicio trivial. Una taxonomía adecuada nos permitirá analizar de forma más clara las relaciones sociales de producción en que se ven envueltas o que determinan las plataformas digitales, así como poner foco en algunas de las propiedades emergentes del sistema en su totalidad. Además, un consenso en torno a una sistemática de las plataformas (entendida como el estudio de la clasificación, la evolución y las relaciones entre las plataformas) puede ayudar o guiar a los gobiernos a actualizar sus marcos regulatorios respecto a los nuevos actores o al papel que han adquirido viejos actores en el mercado y en la sociedad. Esto queda enunciado, por ejemplo, en la propuesta de la Unesco respecto a la gobernanza de las plataformas digitales, porque una apropiada identificación de aquellas entidades que estarán sujetas a futuras regulaciones locales implica, entre otras cosas, una evaluación adecuada de los riesgos que suponen para los derechos humanos y para las instituciones democráticas (Unesco, 2023).

El desafío, no obstante, ha demostrado no ser menor: por un lado, puede resultar estéril en importantes aspectos, dada la rápida transformación del ecosistema tecnológico caracterizado por desarrollos difíciles de predecir; por otro, la opacidad de las grandes corporaciones multinacionales dificulta la investigación y la comprensión profunda de muchos de los elementos clave que caracterizan a las plataformas en un momento dado.

Se han propuesto varias taxonomías para las plataformas digitales. Como suele pasar, la manera en que se organizan las entidades tiene menos que ver con su materialidad que con los fines del clasificador. Este tipo de taxonomías de fenómenos tan tremendamente complejos y dinámicos no se alejan de lo que Levi-Strauss entendía como “ciencia de lo concreto”, clasificaciones que priorizan las exigencias intelectuales de sus clasificadores antes o en lugar de algún criterio de necesidad (Lévi-Strauss, 1964). Aquí lo relevante es situar los fenómenos de una

manera que los vuelva accesibles y proclives a responder las preguntas que nos hemos hecho, usando las herramientas que tenemos a nuestra disposición.

Diferentes clasificaciones habituales, particularmente en el contexto del marketing, se inclinan por enfatizar las “necesidades” que satisface cada sitio (Zhu y Chen, 2015) o su principal formato medial (texto, video, audio, mensajes, interacciones personales, etc.) (Burgess y Green, 2018). Particularmente este último enfoque entrega listas amplias, pero no muy específicas, ya que muchos sitios se plantean como de propósito general, permitiendo con el tiempo múltiples formatos, y el énfasis de cada red social ha demostrado ser un asunto dinámico (por ejemplo, la importante inclinación de Instagram por el video sobre la fotografía, especialmente desde 2016, los esfuerzos de Twitter a partir de 2021 por popularizar sus salas de audio, etc.).

Kaplan y Haenlein (2010) clasifican los sitios *social media* de acuerdo a ciertas propiedades que identifican como fundamentales: “presencia social” (definido como el contacto acústico, visual y físico que el medio permite entre dos personas), “riqueza medial” (concepto basado en el supuesto de que los medios difieren según la cantidad de información que permiten transmitir en un intervalo de tiempo determinado y que esa diferencia haría a algunos más eficaces y menos ambiguos que otros), “autopresentación” (alude al supuesto de que en cualquier tipo de interacción social las personas tienen el deseo de controlar las impresiones que otras personas se forman de ellos) y “autorrevelación” (la revelación consciente o inconsciente de información personal). Así, de acuerdo a la cantidad relativa (“baja”, “mediana”, “alta”) de presencia social y riqueza medial que ofrece una plataforma, frente a la cantidad relativa (“alta”, “baja”) de autopresentación y autorrevelación que imponen estas, los autores identifican seis tipos de redes sociales.

Ya que nuestro interés aquí está puesto en describir y analizar de forma crítica una ramificación comercial del sistema de las plataformas (las agencias de monitoreo de plataformas o AMP), clasificaciones como la propuesta por Kaplan y Haenlein, cuyo objetivo expreso es analizar “oportunidades que surgen” para las empresas en “el campo de las redes sociales en

rápida evolución“ (Kaplan y Haenlein, 2010) y que ponen en el centro solamente las expectativas de las y los usuarios, entendidos meramente como audiencias, no nos serán de gran utilidad.

Esto es porque las expectativas de las audiencias, así como las intenciones declaradas por los dueños de las plataformas, nos informan de manera insuficiente o equívoca sobre las relaciones sociales en que se ven involucrados las y los usuarios y sobre los objetivos reales de las empresas. Así, por ejemplo, una de las categorías descritas por Kaplan y Haenlein, “mundos sociales virtuales” (cuyo ejemplo más conocido es Second Life) está definida por la “libertad” que otorga a los y las usuarias, planteándose como un espacio neutral, como si no hubiera una arquitectura que guiara los comportamientos y los datos a ciertos fines. Según los propios responsables de Second Life, su plataforma se diferencia de un MMOG en que “no hay ningún conflicto fabricado, ningún objetivo establecido” (Kalning, 2007). No obstante, esta afirmación es fácil de refutar (y de pasada, entender por qué clasificaciones como las de Kaplan y Haenlein son insuficientes): la motivación principal de la existencia de Second Life siempre ha sido ganar dinero. De acuerdo a Linden Labs, la empresa dueña de la plataforma, ingresan 650 millones de dólares anuales en “artículos” a través de su mercado virtual y otras transacciones (Parkin, 10 de junio de 2023) y ello determina, sin duda, las cualidades reales de la plataforma.

En *The culture of connectivity* (van Dijck, 2013), la académica de Países Bajos José van Dijn toma un camino un poco más directo y opta, en principio, por dibujar solo cuatro categorías generales: 1) Social networking sites (SNSs), cuyo carácter primordial es promover el contacto interpersonal entre individuos o grupos y que fundamentalmente coinciden con las redes sociales/*social media*. ; 2) Sitios de contenido generado por usuarios (UGC), donde incluye a YouTube, Flickr, Instagram, Myspace, GarageBand y Wikipedia, entre otros; 3) Sitio de intercambio comercial y marketing, cuyo objetivo principal es el intercambio o la venta de productos, como Amazon, eBay o Craigslist; 4) Juegos y sitios de juegos, como FarmVille, The Sims o Angry Birds (según sus ejemplos).

Como vemos, a través de estas categorías nuevamente se opta por englobar sitios con arquitecturas y modelos de negocios muy diferentes, pero que en términos de expectativas de los

usuarios son similares. Así, tenemos entre los sitios de UGC a plataformas tan disímiles como Wikipedia (sin modelo de negocio), Instagram (donde el componente de contacto interpersonal, que define a los SNS es al menos tan importante como el componente de UGC) y GarageBand (que es un software de creación de música sin componentes sociales). En la categoría de sitios de juegos incluye plataformas como FarmVille, que junto con ser un juego de simulación incorpora muchos elementos de interacción social entre sus participantes, y Angry Birds, un juego tradicional que no requiere ningún tipo de interacción en línea entre los jugadores.

Pese a las inconsistencias que puede ofrecer esta sencilla clasificación, la autora es clara en recalcar que los límites entre las categorías de plataformas no son fijos y que existe una batalla continua entre las diferentes empresas por apropiarse de los distintos segmentos que constituyen la sociabilidad en línea. En dicho sentido, las categorías que ofrecen cobran sentido al entender los fines analíticos del trabajo de van Dijck. Su interés primordial es abordar particularmente aquellas plataformas SNS y UGC que son el objeto principal del libro. Desde ahí, van Dijck pone en duda los supuestos sobre la preeminencia de las interacciones sociales que propician las plataformas, cuya forma de presentarse oculta el importante rol mediador de las máquinas y sus dueños.

El foco, entonces, para una categorización que nos ofrezca una perspectiva más reveladora de la configuración del poder en las redes sociales, debe apuntar tanto al nivel sociotécnico (tecnología, contenidos, usuarios y sus hábitos) como a las plataformas como estructuras socioeconómicas con una gobernanza, modelos de negocio y propietarios con determinadas características. Esto lo aborda van Dijck de forma más clara en *The Platform Society* (van Dijck, Poell y Waal, 2018), trabajo donde, para clasificar las plataformas digitales, establece una distinción más general, pero a la vez más precisa, entre plataformas infraestructurales y plataformas sectoriales.

- Las **plataformas infraestructurales** proveen los servicios básicos de información y comunicación sobre las que se construyen otras plataformas y son, por lo tanto, más influyentes. Muchas de ellas son propiedad de los gigantes de la tecnología, en particular las cinco grandes corporaciones estadounidenses: Alphabet, Amazon, Apple, Meta y Microsoft.

Estas empresas, señala van Dijck, operan como “guardianes” de los flujos de datos en Internet, su gestión, procesamiento, almacenamiento y canalización.

- Las **plataformas sectoriales** son aquellas que prestan servicios en un sector o nicho particular, como las noticias, las finanzas, el retail, el transporte, la comida, la salud, la educación, entre otras. Algunas de las plataformas sectoriales más conocidas carecen de activos materiales, no tienen prácticamente empleados y no ofertan productos, contenidos o servicios tangibles. Solamente conectan a usuarios individuales y proveedores de servicios. van Dijck distingue a este subgrupo como “plataformas conectivas”.

La distinción entre plataformas infraestructurales y sectoriales tampoco es fija, sino que más bien existe una dinámica constante que las impulsa hacia la integración. La segunda propuesta de van Dijck va en la línea de una taxonomía enfocada en la identificación de los mecanismos subyacentes de las plataformas como entidades partícipes de un entorno socioeconómico y político particular y la conformación recíproca entre sus actores.

Dentro de las propuestas de taxonomía que abordan la economía política del sistema de las plataformas y que buscan responder no sólo la pregunta “¿cómo son las plataformas y cómo se diferencian unas de otras?” sino que, sobre todo, “¿cómo y en relación con qué actores operan las plataformas en contextos específicos?”, se encuentra la clasificación desarrollada por el canadiense Nick Srnicek.

Srnicek propone una distinción que considera cinco tipos de plataformas digitales. La primera es similar a lo que van Dijck reconoce como plataformas infraestructurales, mientras que las cuatro siguientes serían subcategorías de plataformas sectoriales:

1. **Plataformas de la nube.** Similar al concepto de plataformas infraestructurales usado por van Dijck, son las propietarias de hardware y software que ofrecen en arriendo a negocios que dependen de la nube. De esta manera, empresas como Netflix no requieren poseer y mantener la potencia computacional que exige la gestión de sus servicios, sino que arrienda dicha infraestructura a AWS (Amazon Web Services).

2. **Plataformas industriales.** Son aquellas que producen el hardware y el software que se necesitan para transformar la manufactura industrial tradicional en procesos automatizados conectados por Internet cuya promesa de incrementar la productividad y la competitividad. Sus dos principales exponentes a nivel global son Siemens o General Electric.
3. **Plataformas de productos.** Son aquellas que generan ganancias al transformar bienes tradicionales en servicios digitales y cobrar por ellos un alquiler o una suscripción. Este es el caso de Spotify o Netflix, entre muchas otras compañías.
4. **Plataformas austeras.** Intentan restringir al mínimo sus activos y obtener ganancias mediante una extrema reducción de costos que incluye altos niveles de explotación laboral aprovechando espacios desregulados del mercado del trabajo a nivel global. Es el caso de compañías como Uber, Rappi o Airbnb.
5. **Plataformas publicitarias.** En esta categoría encontramos a aquellas que habitualmente llamamos redes sociales, empresas como Twitter, Facebook, Snapchat, Weibo, LinkedIn, TikTok o VK y también a Google Search, Youtube y Reddit, entre muchas otras, que generan ingresos mediante la extracción de datos de las actividades *online* de sus usuarias y usuarios (*big social data*), el análisis de dichos datos para producir segmentos de audiencia y la subasta de espacio publicitario para anunciantes basándose en esas métricas.

Un problema de la tipología de Srnicek es que extiende la categoría de plataforma a casi cualquier negocio imaginable solo por el hecho de que hoy en día casi cualquier negocio puede ser considerada una fuente potencial de datos. Esto es una consecuencia de su esfuerzo por componer un análisis totalizador en torno a su concepto de “capitalismo de plataformas”. Otro problema es que no nos informa en qué lugar quedan aquellas aplicaciones similares a las plataformas publicitarias, pero que carecen (permanentemente o de momento) de un modelo de negocios o que poseen una estructura no centralizada, como Mastodon o Wikipedia. Asimismo, deja afuera a plataformas con modelos de negocio basados en la comercialización de los datos de sus usuarios, pero que carecen (de momento) de un modelo de negocios publicitario, como WhatsApp.

Uno de los focos relevantes de esta clasificación es el análisis de las plataformas en cuanto a su negocio primordial. De ello resulta que, al igual que en las otras tipologías presentadas más arriba, algunas plataformas parecen ajustarse a más de una categoría. En este caso, por ejemplo, junto con ser plataformas publicitarias, Google y Facebook también participan en el sector de las plataformas de la nube. No obstante, las reestructuraciones de estas compañías en años recientes parecen darle la razón al instrumento analítico de Srnicek, en tanto que las grandes corporaciones se han ido desagregando de acuerdo a algunos de los aspectos que destaca el canadiense. Así, en 2015 Google pasó a llamarse Alphabet, conglomerado que reúne bajo su alero a varias compañías, dentro de cuya estructura los negocios de Google Search y Google Cloud Platform, que es la plataforma de la nube de la corporación, quedan claramente delimitados. Algo similar ocurre en lo que antiguamente era Facebook, que hoy en día es solo una de las compañías dentro de la corporación Meta Platforms, Inc. creada en 2021. En síntesis, la evolución de la gobernanza de estas corporaciones tiende a replicar mejor las observaciones de Srnicek que distribuciones como las de Kaplan y Haenlein, lo que nos da una pista respecto a cómo mirar el fenómeno si nuestro objetivo es un análisis que tome en cuenta sus relaciones sociales de producción.

La taxonomía de Srnicek presenta algunas similitudes con la que tomó como fundamento el reciente Reglamento de Servicios Digitales de la Unión Europea (UE). En primer lugar, el trato para estas empresas es a través del concepto de plataforma, entendida como “servicios de intermediación en línea” (Granada, ), lo que coincide con las nociones tanto de Srnicek como de van Dijck (y de O'Reilly). Sin embargo, la UE se concentra sobre todo en la relación de los consumidores con los servicios, antes que en las cualidades particulares de los servicios o las características materiales de los negocios que hay detrás. Desde esta perspectiva, el reglamento considera cuatro grandes categorías (European Commission, 2023): 1) Servicios intermediadores que ofrecen infraestructura de redes (proveedores de acceso a Internet, registradores de nombres de dominio); 2) Servicios de alojamiento de datos (computación en la nube y alojamiento web); 3) Plataformas en línea (mercados en línea, tiendas de aplicaciones, plataformas de economía colaborativa y plataformas de redes sociales); y 4) Plataformas y buscadores en línea de tamaño muy grande (utilizadas por al menos 45 millones de usuarios mensuales o “más del 10 % de los



450 millones de consumidores de la UE”, entre las que se cuentan Facebook, Amazon Store, Apple AppStore, Google Maps, Booking, Alibaba, YouTube, Twitter, TikTok, Bing y Google Search, entre otras), que son objeto de normas especiales (Unión Europea, 2022).

La categorización de la UE incorpora dos elementos que son fundamentales para la comprensión efectiva del sistema de las plataformas: resalta el papel de los proveedores de Internet y pone foco en la penetración (cantidad de usuarios) que tiene cada plataforma. Este último es un factor que también subraya la Unesco en sus Directrices para la gobernanza de las plataformas digitales (Unesco, 2023), donde establece que el tamaño, alcance y la cuota de mercado (antecedentes de la “probabilidades de influir en una parte significativa de la población”) son criterios básicos para el reconocimiento de las compañías a las que deben orientarse las normativas en jurisdicciones específicas. Aunque estos aspectos no son centrales para nuestro marco de análisis, son relevantes para dar cuenta de que el establecimiento de regulaciones que garanticen algún grado de transparencia y responsabilización de estos negocios es algo apremiante.

### **Plataformas publicitarias**

Para nuestros fines, la taxonomía de Srnicek se presenta como la más útil. La categoría plataformas publicitarias se ajusta casi a la perfección a aquellas plataformas que son la fuente de los datos que las AMP usan como base de su propio negocio. Si bien se realiza minería de datos de otras plataformas fuera de esta categoría, se trata en esos casos de aproximaciones no comerciales, no orientadas a la venta de informes y monitoreos a terceros actores.

Hasta aquí nos hemos referido indistintamente a estas plataformas publicitarias como redes sociales o *social media*. Asimismo, hemos resaltado la importancia del concepto de plataforma, debido a que expresa una posición estratégica y una condición material para estas empresas. Y aunque el término plataforma publicitaria señala relativamente bien lo que constituye dicha categoría, falla en varios aspectos. En primer lugar, al rebajar el componente social de estos sitios a un elemento secundario, borra de un plumazo un énfasis en los elementos

sociotécnicos que han sido largamente tratados en la literatura y que si bien merecen ser morigerados en su relación con los componentes socioeconómicos y políticos, no parece apropiado hacerlos desaparecer. Ello, además, rompería la vinculación del concepto con sus denominaciones habituales en la vida cotidiana, en la prensa, en la legislación inclusive (redes sociales, *social media*), lo que atenta contra su comprensión y su potencial como instrumento analítico. Es por ello que en este trabajo hemos optado por hablar de “plataformas sociales publicitarias” o simplemente “plataformas sociales”, concepto que conserva el énfasis en lo social y que además ya se usa como alternativa a “redes sociales” en español, aunque de forma menos habitual.

La tarea de las plataformas sociales publicitarias, servicios y aplicaciones construidos en torno a la convergencia de contenidos compartidos, comunicación pública y conexiones interpersonales (Burgess y Green, 2018), consiste en posicionarse a sí misma entre las personas y como el espacio en que tienen lugar sus actividades. Esto les confiere un acceso privilegiado a los datos que tales actividades e interacciones producen. Pero además de extraer y registrar dichos datos, las plataformas necesitan someterlos a un análisis, clasificarlos, relacionarlos, convertirlos en suma en *data* “trabajada”. Esto generalmente ocurre mediante el procesamiento automatizado llevado a cabo por los algoritmos de una máquina.

Como hemos indicado antes, los algoritmos son el conjunto de instrucciones predefinidas para el manejo de determinada información y su utilización en diferentes tareas. La aplicación del concepto a tareas computacionales relacionadas con el procesamiento de datos presupone que estos se presentan en un estado inicial y que son sometidos a una serie de operaciones hasta llegar a un estado final que provee una solución a un problema. Es preciso resaltar que aunque los algoritmos operen de forma automática y se presenten como neutrales, cuando no naturales o necesarios (en el sentido de ser la mejor respuesta posible a la resolución de un problema dado), son creados por personas y decisiones corporativas, de manera que en su producción la ideología y las nociones culturales son codificadas en los artefactos tecnológicos (Timcke, 2021). Así, el procesamiento de los datos desde un estado inicial a uno final, implica un diseño, una serie de decisiones que quedan estampadas en dicho resultado final.

Ilustremos esto con un ejemplo sencillo. La plataforma X/Twitter rellena de contenido un *timeline* llamado “Para ti” personalizado para cada usuario. Según indican, para hacerlo usan “publicaciones de cuentas” a las que dicho usuario sigue, así como con “publicaciones recomendadas”, además de publicaciones promocionadas (X, s. f.). Esta línea de tiempo no se presenta en orden cronológico, sino que se configura de acuerdo a un algoritmo. Desde luego, X no revela más que generalidades sobre ese algoritmo:

Seleccionamos cada publicación utilizando una variedad de señales, incluyendo lo popular que es y cómo la gente de su red está interactuando con ella. (...) Te recomendamos publicaciones basadas en a quién ya sigues y a los temas que sigues (...) Al clasificar una respuesta más alta, consideramos factores como si el autor de la publicación original ha respondido, si una respuesta es de alguien a quien sigues o si la persona es un suscriptor de X Blue. (X, s. f.)

Con esta información, nos podemos imaginar cómo podría operar un algoritmo sencillo para dar forma al *timeline* de un usuario al que llamaremos USER:

**Paso 1:** mostrar en orden aleatorio todos los contenidos que aún no hayan aparecido en el *timeline* de USER, que hayan sido publicados en las últimas 5 horas por los perfiles a los que USER sigue y que hayan tenido 5 o más interacciones (retweets, likes o comentarios).

**Paso 2:** intercalar entre cada 2 publicaciones mostradas en el *timeline* de USER un contenido que haya tenido más de 100 interacciones de cuentas a las que USER no sigue.

**Paso 3:** intercalar cada 8 publicaciones mostradas en el *timeline* de USER un contenido con más de 45 interacciones publicado por un suscriptor de *X Blue*<sup>1</sup> que sea seguido por al menos 5 de los perfiles a los que sigue USER.

---

<sup>1</sup> Servicio “premium” de X que implica un pago por parte del usuario o usuaria para obtener algunas funciones adicionales.

**Paso 4:** intercalar entre cada 10 publicaciones mostradas en el *timeline* de USER un contenido promocionado.

**Paso 5:** volver al paso 1.

Lo que resulta transparente a partir de la clase de parámetros que X/Twitter explicita en su sitio es que hay una enorme serie de decisiones (corporativas) implícitas en cada uno de los módulos de la aplicación. En muchos casos estas decisiones solo admiten variables de carácter subjetivo que alguien de la empresa debe tomar en atención a ciertos fines. Esto es lo que da forma a la arquitectura de la plataforma. X/Twitter nos muestra contenidos “populares” o “relevantes”, y esa relevancia o popularidad se mide numéricamente a partir de parámetros creados por la propia compañía: *retweets* o *likes*, número de respuestas, velocidad con la que el contenido obtiene interacciones, número de usuarios que revisan el perfil de quien escribió un *post* luego de que en su *timeline* se imprimió ese *post*, cantidad tiempo que pasan otros usuarios leyendo un hilo, popularidad medida en número de seguidores de los perfiles que interactúan con un *post*, etc. No hay nada de natural en un *retweet*: es una creación específica de la plataforma, de probada efectividad como mecanismo para lograr que las personas produzcan contenidos más llamativos y pasen más tiempo en la red social. Lo que en la jerga de las plataformas (y de las AMP) se conoce como *engagement*.

Al medir la “relevancia” de una publicación de acuerdo a estas diferentes métricas de “interacción”, lo que se afirma implícitamente es que para que en X/Twitter algo sea relevante, debe provocar una reacción medible según dicha métrica en el lector. De nada le sirve a dicha plataforma que si usted ve un post muy interesante, apague el teléfono, se lo comente a sus amigos y durante dos horas conversen sobre dicho contenido. Claramente fue relevante para usted (produjo interés e interacciones sociales), pero no fue relevante para X/Twitter. Lo que X/Twitter busca es que las personas usuarias pasen la mayor cantidad de tiempo en su aplicación produciendo la mayor cantidad posible de datos (interacciones y contenidos) por segundo. Tales intenciones, que son imperativos que manifiestan el carácter eminentemente comercial de la plataforma, quedan inscritas en los algoritmos.

Este procesamiento, en el caso de las plataformas sociales publicitarias, convierte datos asociados a la experiencia digital de las usuarias y usuarios en productos con valor comercial, en mercancías. La forma que adquieren estas mercancías son variadas y depende tanto de las cualidades de cada plataforma, del tipo de datos que están orientadas a extraer, de las características de sus clientes y de sus modelos de negocios.

En primera instancia, la principal mercancía es un espacio publicitario que asegure al comprador la atención del público al que se quiere dirigir. Esta es la mercancía primordial de la industria dedicada a la publicidad: la audiencia como mercancía. En la década de 1970, Dallas Smythe propuso, partiendo de una teoría realista y objetivista, que aquello que se comercia con el mercado de los anunciantes de los medios de comunicación de masas no es contenido, sino la atención del público, en contra de la idea de que el producto comercializado estaba conformado por entidades mentales subjetivas como mensajes, información, imágenes, significados, entretenimiento, manipulación, etc. (Smythe, 2006).

Los medios digitales que operan sobre Internet amplifican los elementos centrales del argumento de Smythe. Los datos digitales, cuyas propiedades metadadales permiten medir, rastrear y monitorear con precisión cada transacción informacional, pueden ser usados para refinar como nunca antes el proceso de entregar audiencias multimediales a los anunciantes (Mosco, 2009). Las compañías que lidian con audiencias digitales pueden empaquetar y reempaquetar usuarios y usuarias en innumerables formas específicas para que reflejen las diversas categorías demográficas y de consumo que el mercado requiera, si acaso saben extraer los datos precisos (si adquieren una posición de poder adecuada para hacerlo).

Por ejemplo, Facebook permite delimitar la publicidad de un candidato político a su circunscripción electoral aproximada y sólo mostrar determinados anuncios a ciertos perfiles demográficos (por ejemplo, sólo a mujeres de determinado rango de edad con determinados intereses “declarados” dentro de la plataforma). Este tipo de operaciones de registro y clasificación es capaz de generar un creciente número de productos más o menos sofisticados. Así, una plataforma puede tenerle a usted al mismo tiempo dentro de un *target* tan general como

“perfil activo de la plataforma” y al mismo tiempo en otro tan particular como “mujeres transexuales de 20 a 35 años con varias mascotas” y comerciar su atención con anunciantes completamente diferentes en cada caso. En ese sentido, no solo tenemos un “doble de datos” alojado en las plataformas, sino que múltiples y variados.

Estos segmentos o clasificaciones demográficas, socioeconómicas, psicológicas y psicográficas de grandes volúmenes de usuarios y usuarias son la principal mercancía que plataformas sociales publicitarias como X/Twitter, Google o Facebook comercian a través de sus portales X Anuncios, Google Ads y Anuncios en Meta, respectivamente. La propuesta comercial es relativamente sencilla, pero requiere de un gran volumen de datos de cada usuario para funcionar y un enorme poder de automatización. Facebook lo resume así:

Cada vez que hay una oportunidad de mostrar tu anuncio a alguien de tu público en las plataformas de Meta (como Facebook, Instagram y Messenger), el anuncio entra en una subasta en la que se determina si esa persona debe ver tu anuncio. El objetivo de una subasta es simple: mostrar tu anuncio a las personas del público con más probabilidades de interesarse en él. (Meta, s. f.)

El factor clave de esta descripción es “las personas del público con más probabilidades de interesarse en él”, cuya realización en la práctica implica un conocimiento profundo y amplio de nuestros dobles de datos.

Pero además, el patrimonio de datos de las plataformas puede usarse para el análisis de tendencias de consumo globales y locales, en marcos temporales de diferente extensión; sondeos de la reputación institucional, corporativa o personal; monitoreo de fenómenos comunicacionales en tiempo real; alertas sobre crisis comunicacionales; análisis de sentimientos asociados a eventos particulares; estudios dirigidos a perfilar y predecir mediante correlaciones el comportamiento de diferentes actores sociales; el seguimiento a la actividad de usuarias o usuarios específicos; entre otros.

Así, por ejemplo, el análisis combinado de las horas que miles o millones de usuarios y usuarias emplean mirando *gameplays* en Twitch (cruzando datos de detallados perfiles personales, incluyendo múltiples aspectos demográficos, su historial de actividades y compras, horarios de conexión, geolocalización, juegos favoritos e interacciones) puede ser convertido en una valiosa mercancía para las compañías que desarrollan videojuegos o que venden productos dirigidos a niños, niñas y adolescentes.

No obstante, aunque plataformas digitales como Facebook, X/Twitter y otras ofrecen interfaces para que sus personas usuarias puedan realizar monitoreos básicos del desempeño de sus propias publicaciones, en general no ofertan servicios especializados de análisis de sus datos para responder preguntas específicas de terceros. Como hemos señalado, el negocio principal de las plataformas publicitarias es la subasta de espacio publicitario. Lo que venden a los anunciantes no son los datos mismos, sino la promesa de que su *software* de análisis conectará de manera eficiente los anuncios con los usuarios correctos, con su *target*. Así, la transformación de los datos de las plataformas sociales publicitarias en productos como los mencionados cae sobre todo en manos de compañías intermediarias.

A este tipo de productos contruidos a partir de los datos de las plataformas sociales los llamaremos, en términos genéricos, mercancías analíticas. Un ejemplo paradigmático de estas mercancías y al que le hemos dedicado el grueso de la introducción es el conocido informe Big Data desarrollado por Alto Data Analytics para el gobierno de Piñera.

## **Agencias de monitoreo de plataformas (AMP)**

En el mercado chileno, las empresas que producen estas mercancías comúnmente se describen a sí mismas como proveedoras de monitoreos de redes sociales, entre otras denominaciones similares. Engloban actividades que en la literatura académica se ha asociado a la minería de datos de plataformas sociales. Aquí, en pro de la claridad, llamaremos a estas compañías agencias de monitoreo de plataformas (AMP).

Existen diferencias y matices entre los monitoreos que realizan las AMP como Alto Data Analytics o Brandmetric y la minería de datos de plataformas sociales, como un área que engloba muchos procedimientos para diferentes fines. Pero en términos generales, en ambos casos se trata, básicamente, de contar interacciones en plataformas como X/Twitter o Facebook, asociarlas con metadatos y extrapolar a partir de diversos procesamientos cierta clase de conocimiento sobre las actividades de usuarias y usuarios, humanos o no.

### **Tipificación de la minería de datos**

La minería de datos de plataformas sociales, según detalla Helen Kennedy (2016), puede desarrollarse a través de cuatro vías o categorías:

- **Minería “dentro de la plataforma”**: herramientas e interfaces que las propias plataformas ofrecen a sus usuarios y usuarias con datos básicos sobre el desempeño de sus publicaciones y otros factores.
- **Minería “gratis y simple”**: aplicaciones gratuitas en línea que ofrecen la posibilidad de consultar métricas de diversas redes sociales. Son similares a las herramientas integradas en las propias plataformas, pero con algunas ventajas menores, como visualizaciones más llamativas. Estas aplicaciones recurren a las API, por lo que son sensibles a sus limitaciones (tema que abordaremos más adelante), lo que explica por qué esta categoría, tal como la describió Kennedy en 2016, haya quedado prácticamente obsoleta en 2024.
- **Minería “comercial”**: la llevada a cabo por agencias de monitoreo que exploran las principales plataformas sociales publicitarias y otros espacios “sociales” de la web. Para hacerlo, estas empresas utilizan una variedad de tecnologías y métodos. Se trata de servicios ofertados como útiles para un amplio rango de propósitos, generalmente costosos y que incluyen el trabajo de analistas humanos además de procesos automatizados.
- **Minería “gratis y compleja”**: en esta categoría Kennedy incluye desarrollos y herramientas de monitoreo que se han creado con el fin de investigar las plataformas sociales, generalmente en contextos académicos, como NodeXL o Gephi.



Queda fuera de esta clasificación la minería de datos de plataformas sociales que realizan grandes corporaciones multinacionales y los gobiernos y sus aparatos policiales y de inteligencia. Aquí tampoco nos interesa la minería de datos que algunas organizaciones, como empresas o partidos políticos, desarrollan internamente. Nuestro foco estará puesto solo en lo que hacen las compañías intermediarias que desarrollan lo que Kennedy llama minería “comercial”.

## **La labor de las AMP**

Como observábamos a partir del trabajo de Siegert (2015), la adopción de nuevas tecnologías de registro y procesamiento de datos viene asociada a la creación de nuevas organizaciones y profesiones. Las agencias de monitoreo de plataformas sociales (AMP) y los consultores y especialistas en *big data* son algunas de las novedades que traen consigo las tecnologías de comunicación digitales que emergieron durante las últimas décadas del siglo XX y que hoy son ubicuas. En este apartado desarrollaremos los principales aspectos conocidos sobre estas empresas, su trabajo, sus productos y su posición en el mercado de los datos de las plataformas.

Como ya hemos establecido, nuestro objetivo aquí es profundizar en el análisis de lo que Kennedy llama “minería comercial”. Es la misma Helen Kennedy (2016) quien ha desarrollado la única aproximación con algún grado de profundidad al papel de estas empresas intermediarias, a las que califica como *commercial social media insights companies*. Estas empresas, que analizan la actividad en las redes sociales a solicitud de clientes de pago, con frecuencia ofrecen variados servicios y productos. Sus tareas pueden pasar por identificar cuántas personas están hablando sobre un tema o una marca, quiénes están hablando y qué dicen (por ejemplo, ¿cómo ha sido recibido por los jóvenes el nuevo modelo de teléfono de Samsung?), encontrar líderes de opinión (*influencers*) sobre determinados temas o conversaciones clave (¿quién está detrás de las publicaciones más populares en contra del gobierno?), extraer y cruzar información demográfica sobre las personas que participan en conversaciones en las plataformas sociales (¿cuáles son las conversaciones más relevantes de esta semana para las mujeres de las comunas con alto poder adquisitivo de Santiago?), entre otras.

A 2023 los conceptos de “sector”, “industria” o “mercado” de monitoreo, análisis *social media* o *social media data mining* son frecuentemente usados por algunos de sus propios actores de manera informal. Aquí proponemos que en la actualidad podemos considerar al conjunto de las empresas de monitoreo y análisis de plataformas sociales como un sector comercial claramente definido, así como una industria (Black, 2002), ya que se trata en términos generales de una actividad inserta en la economía cuyo propósito es transformar insumos similares (en este caso, los datos de las plataformas) en un grupo de productos relacionados (diferentes tipos de mercancías de comunicación). Detallar las particulares de sus procesos productivos, de su oferta y de las relaciones sociales de producción que caracterizan a esta industria es parte de este estudio.

Las AMP operan en un mercado caracterizado por una fe casi ciega en el poder del *big data* y los algoritmos. En esta ideología, orientada por lo que se ha caracterizado como sublime digital y algorítmico, domina la creencia de que los datos recopilados en las plataformas sociales pueden proporcionar información detallada sobre algo a lo que antes prácticamente no existía el acceso directo: la opinión y los sentimientos públicos. En este creciente nicho, diferentes compañías y organizaciones públicas y privadas (incluyendo el Estado de Chile, como hemos visto en la introducción) están interesadas en saber qué se dice sobre sus productos y acciones en las redes sociales y qué sentimientos se expresan. La aspiración es saber quién dice, piensa o siente “qué”, “cuándo”, “dónde” y hasta “por qué” en la sociedad, pero sin invertir recursos, tiempo y fe en costosas, lentas y anticuadas ciencias sociales.

Las empresas que caracterizan a este mercado otorgan una variada cantidad de nombres a sus quehaceres: análisis social media o de redes sociales, *social listening* o escucha social, *social media reporting*, *social media analytics*, inteligencia social, entre otros. Estas agencias acceden al *big social data* principalmente (aunque no exclusivamente) a través de las API de las plataformas para desarrollar sobre esos datos tareas de búsqueda, recopilación, procesamiento, análisis, interpretación y almacenamiento sistemático de contenidos generados por usuarios y usuarias (Sheldon, Rauschnabel y Honeycutt, 2019). Con ello elaboran diversos productos analíticos,

generalmente “monitoreos” o “informes” cuyos “entregables” pueden adoptar diversas formas, desde sencillos reportes escritos a complejos *dashboards* (aplicaciones) en línea, generalmente con un fuerte énfasis en la visualización de datos.

Estas mercancías analíticas puede enfocarse en una o varias plataformas sociales, mostrando el conteo de veces que se ha visto determinado contenido (perfiles, post, *hashtags*, temas, etc.), sus interacciones (*clicks*, *likes*, *retweets*, respuestas, etc.) y la relación de esos contenidos con los metadatos asociados que sean de interés, para con ello extrapolar conclusiones sobre comportamientos y tendencias de un conjunto de usuarios en un rango de tiempo dado. Mediante diversas herramientas, las AMP ofrecen diferentes análisis y visualizaciones, insumos particularmente valorados en el ámbito del marketing. Dependiendo de la complejidad de las aproximaciones y las herramientas utilizadas, que pueden incluir procesamiento de lenguaje natural y el uso de grandes modelos de lenguaje (IA), el costo de estos servicios puede variar ampliamente.

Uno de los aspectos críticos de estas mercancías analíticas y que las convierte en un objeto apropiado para un análisis desde la comunicación política, es que se ofertan como “inteligencia” o “conocimiento accionable” (*actionable knowledge*). Este es un concepto proveniente del *management* que apunta a la extracción de información relevante para ser puesta en práctica, tomar decisiones gerenciales, mejorar los resultados comerciales y que sea presentada de una manera que facilite su uso en la vida real (Arefin, 2023). Es decir, se trata de mercancías cuyo contenido es prescriptivo (en oposición a descriptivo) y que otorga poder a su poseedor (Argyris, 2003), la capacidad de intervenir y transformar los entornos.

### **El papel de las interfaces de programación de aplicaciones**

Como hemos indicado, el grueso de los datos que se emplean en estos monitoreos provienen de las interfaces de programación de aplicaciones, o API, de las plataformas sociales. Amazon Web Services las describe así:

Las API son mecanismos que permiten a dos componentes de software comunicarse entre sí mediante un conjunto de definiciones y protocolos. Por ejemplo, el sistema de software del instituto de meteorología contiene datos meteorológicos diarios. La aplicación meteorológica de su teléfono “habla” con este sistema a través de las API y le muestra las actualizaciones meteorológicas diarias en su teléfono. (AWS, s. f.)

Las API (públicas) suelen estar diseñadas para fomentar la elaboración de software por parte de desarrolladores independientes que “agreguen valor” a la plataforma. Una API expone solo los objetos o acciones que el desarrollador externo necesita, lo que simplifica la programación. Ejemplos de usos comunes de las API de algunas plataformas sociales son la posibilidad usar las credenciales de Facebook para crear una cuenta en otro sitio web o poder acceder a Twitter mediante un cliente o app no oficial, como Twittrific (característica eliminada por esa compañía a principios de 2023). Aunque la principal motivación de las empresas al desarrollar sus API es incrementar la producción de nuevas ideas y atraer nuevos usuarios a partir de una inversión relativamente modesta, estas también han servido y en ocasiones se han adaptado especialmente para que investigadores de diverso tipo accedan a datos de las plataformas.

Los datos que se recuperan mediante las API de las plataformas sociales publicitarias son tratados con frecuencia como “fuentes públicas” o “fuentes abiertas” (p. ej. Monaco y Arnaudo, 2020). Así se puede corroborar en las declaraciones en torno al caso del informe Big Data (p. ej. Basoalto, 2020). Este lenguaje no solo da a entender que tales datos no tendrían dueños o que sus dueños habrían renunciado a sus derechos sobre ellos: además implica que tales datos no han pasado por las manos de actores interesados, que no han sido manipulados. En cierto modo, se procura transmitir que se trataría del dato en su estado *natural*, no filtrado ni procesado. Dicha forma de hablar de los datos provenientes de las API opaca las cualidades fundamentales de estos y cada una de las decisiones humanas que dan pie a su elaboración.

Es importante recalcar esto porque gran parte de los discursos en torno al *big data* está orientado a esconder a los intermediarios, resaltar la presunta neutralidad de los números y el valor de la “desintermediación” (Sadin, 2018). Los algoritmos de las plataformas como Instagram o la rusa VKontakte no capturan la totalidad de los datos posibles de sus millones de usuarios ni registran datos al azar, sino que seleccionan de un ingente número de entradas (*inputs*) aquellos datos que benefician más al negocio de la plataforma. Asimismo, diferentes API (y cada plataforma puede ofrecer una variedad de ellas) imponen distintas limitaciones a los datos que un investigador cualquiera puede recopilar y, por lo tanto, determinan el tipo de preguntas que se pueden hacer en relación con ellos (Matei, Jullien y Goggins, 2017). En tal sentido, es innegable que los datos que utilizan las agencias de monitoreo son datos ya procesados o semiprocesados y que tal condición es constitutiva de su naturaleza. Ahondaremos en esto más adelante.

Hasta aquí hemos desarrollado los conceptos centrales pertinentes para referirnos tanto al sistema de las plataformas digitales como al entorno de la minería de datos. En el siguiente capítulo abordaremos las perspectivas desde las que se han estudiado estos fenómenos y delinearemos nuestro marco teórico.



## MARCO TEÓRICO

Para exponer las complejidades de la red de relaciones económicas, políticas y sociales en las que se sitúan los fenómenos que aborda este trabajo, en la sección *Perspectivas de estudio en torno al uso de los datos sociales masivos digitales*, presentaremos algunas de las principales líneas de análisis desde las que se han abordado, en el campo de las comunicaciones y las ciencias sociales, los análisis y la crítica a las estructuras que conforman el sistema de las plataformas.

En el siguiente apartado de este capítulo, titulado *Economía política de la comunicación y la información: un enfoque para el análisis crítico de las agencias de monitoreo de plataformas*, desarrollaremos el estado del arte de las ópticas teóricas que guían esta investigación, la teoría crítica y la economía política de la comunicación (EPC), con relación al abordaje de nuestro campo de investigación.

En el último apartado de este capítulo, *La EPC aplicada al problema del monitoreo de plataformas sociales*, desarrollaremos la propuesta de encuadre teórico para nuestra aproximación. En esta sección y a partir de los antecedentes planteados a lo largo del capítulo, situaremos a las AMP de acuerdo a su posición en el sistema de las plataformas y en relación con sus clientes. De dicho esquema general se desprenden las problemáticas particulares y las preguntas cruciales que orientan esta investigación, lo que a su vez circunscribe las herramientas, métodos y fuentes que emplearemos para el análisis del sector.

### **Perspectivas de estudio en torno al uso de los datos sociales masivos digitales**

En 2006, Facebook abrió su plataforma fuera del entorno estudiantil a cualquier usuario de más de 13 años. Desde ese mismo año, debido a la introducción de su *News Feed* (que mostraba todo lo que habían hecho en la plataforma los contactos de cada usuario o usuaria) y, al

año siguiente, con el lanzamiento de *Beacon* (que mostraba los productos auspiciados que los amigos de las y los usuarios habían comprado), la compañía no ha dejado de recibir críticas debido a su aparente desprecio por la privacidad, el flexible uso que hace de los datos que recoge y la opacidad respecto a las decisiones detrás de sus algoritmos. Cuando la compañía salió a la bolsa en 2012, más allá de haber retirado algunas características polémicas, ninguna de las críticas de fondo se habían resuelto. Desde ese momento hasta 2023, cuando alcanzó los tres mil millones de usuarios a nivel global, Facebook ha estado en el centro de varios escándalos por mal uso de datos privados (ur Rehman, 2019), innumerables campañas políticas abusivas (Shane e Isaac, 2017), junto a algunos genocidios (Mozur, 2018) y golpes de Estado (Schaffar, 2016). Probablemente, hasta la compra de Twitter por parte de Elon Musk a fines de 2022, Facebook era claramente el epítome de lo que estaba roto en Internet.

No obstante, hasta el Reglamento de Servicios Digitales de la Unión Europea (Unión Europea, 2022), aprobado en octubre de 2022 por el Consejo Europeo y que entrará en completo vigor en 2024, ninguna legislación extensiva había obligado a las plataformas sociales a ofrecer algún grado de transparencia sobre la forma en que operan sus algoritmos. Pese a que Google comenzó a transformar a sus usuarios en mercancías en el año 2000, cuando lanzó Google AdWords (Google, 2012), hasta la actualidad los esfuerzos por regular a estas empresas y hacerlas más transparentes han sido, por decir lo menos, tímidos.

La academia ha recorrido un camino opuesto al del legislador. Pese a la opacidad de las plataformas, investigadoras e investigadores de diversas áreas han pavimentado el camino para entender las virtudes y los peligros de estas tecnologías que, ancladas en Internet, tomaron al mundo por sorpresa y rápidamente, como argumentan Gilbert y Willliams (2022), permitieron a dos poderosos bloques ideológicos, el sector tecnológico y el sector financiero, actuando a nivel global, imponer sus términos y valores.

En ese periplo, las tendencias de las investigaciones han pasado de un optimismo a veces inmoderado a un pesimismo exacerbado y han transitado por diferentes énfasis disciplinares. En esta sección presentaremos algunas de las principales perspectivas de análisis desde las que se



han abordado, en las comunicaciones y las ciencias sociales, el estudio y la crítica a las estructuras que conforman el sistema de las plataformas.

### **Perspectivas administrativas**

En 1941, Paul Lazarsfeld denominó “investigación administrativa” en el campo de los estudios de los medios y la comunicación a aquella investigación instrumental que se preocupa predominantemente del desarrollo, planificación o administración de actividades de comunicación y sus tecnologías con el fin de hacerlas más eficientes y efectivas ). Este planteamiento se presenta en oposición a las investigaciones de carácter crítico, aquellas que abordan los problemas históricos, sociales y culturales más amplios que plantea la comunicación mediada tecnológicamente, apoyándose en la filosofía, la teoría, la crítica y el análisis cualitativo (Fuchs, 2017), a menudo desde la perspectiva del interés público. Así, mientras un enfoque reflexiona sobre los medios considerando las formas en que existen actualmente como algo dado, el otro busca explicar por qué adquieren esas formas y cómo podrían llegar a tener otras configuraciones más proclives al bien común.

El propio Lazarsfeld consideraba que las investigaciones críticas y administrativas podían complementarse mutuamente, aunque su énfasis estaba puesto en usar la investigación crítica como una fuente de ideas para ser explotadas financiera e ideológicamente por la corriente administrativa principal (Jensen, 2020). Aquí, en cambio, debemos nutrirnos de la investigación administrativa desarrollada sobre las nuevas tecnologías de comunicación debido a que estas aproximaciones han predominado desde un principio en la indagación sobre las prácticas de los actores dominantes de Internet, un campo que durante mucho tiempo se ha tendido a considerar como eminentemente técnico. Esa pretensión, la de hablar de las tecnologías de la información y comunicación en Internet, de *big data*, algoritmos, *cookies*, métricas, etc., como quien habla de bujías, válvulas, velas y pistones, es aún imperante en buena parte de la literatura del área. La atención a estas fuentes es, por lo tanto, reveladora de ciertos velos significativos y nos permitirá entender mejor los imperativos materiales y comerciales que están inscritos en los conceptos que usamos cotidianamente para hablar sobre las plataformas digitales.

## Publicidad y marketing

Según Olshannikova et al. (2017), los campos que se han ocupado del estudio, definición y usos del *big social data* son áreas de la informática con una eminente orientación comercial como la computación social (*social computing*), que introduce variables relacionadas con las prácticas sociales en el diseño de sistemas computacionales y aplicaciones; la ciencia de grandes datos (*big data science*), que provee las herramientas para manejar y procesar este tipo de datos; y la analítica de datos (*data analytics*), disciplina cuyo foco es el análisis computacional sistemático de datos para la extracción e interpretación de patrones “significativos” y su aplicación en la toma de decisiones gerenciales; además de las ciencias sociales computacionales (*computational social science*), campo que utiliza *big social data* para llevar a cabo investigaciones sociales, habitualmente de mercado.

Dado este contexto, probablemente la publicidad y el marketing son las disciplinas que más se han nutrido y han aplicado el conocimiento y las aproximaciones que han emergido de estos cuatro campos, debido a que están más directamente relacionadas con el negocio principal de las plataformas sociales publicitarias. De estas áreas surgen las descripciones más difundidas de los aspectos básicos del monitoreo *social media* y la “analítica social” (Kelsey, 2017).

Un sector estrechamente ligado a los mencionados es la comunicación estratégica, que ha sido definida como el uso intencionado de la comunicación por parte de una organización para cumplir su misión (Hallahan, 2007). Su aproximación a este tipo de datos en principio tiene que ver con conocer y medir la presencia *online* de diferentes actores y grupos de interés (empresas, marcas, personalidades, políticos, entre otros), mejorar la percepción que la audiencia *online* manifiesta de ellos y enfrentar las crisis comunicacionales en los medios sociales digitales (Canhoto, 2015). Algunos autores han señalado que las relaciones públicas se enfrentan a una revolución irreversible provocada por el *big data* (Wiencierz y Röttger, 2019), aunque parece ser una revolución más bien limitada: la mayor parte de la literatura que se ha producido en este campo en torno al *big data* está ligada a la comunicación de marketing, donde se apuesta por el uso de los datos de las plataformas para mejorar la microfocalización de clientes y la

“cocreación” de productos y de información, orientada según intereses corporativos (Wiesenberg, 2017). Sin embargo, también existen unas pocas aproximaciones que han analizado las relaciones públicas dirigidas por datos desde una perspectiva crítica (Gregory y Half, 2020).

Con el desarrollo de las plataformas sociales de Internet y su penetración extensiva en la población, emergen enfoques que buscan poder levantar datos significativos sobre la sociedad a partir del comportamiento en línea de las personas (Rogers, 2013). La constatación de que sitios como Twitter y Facebook eran relevantes foros de discusión y difusión de ideas políticas y el hecho de que muchas de las conversaciones que se volvían relevantes allí repercutían en las agendas de la prensa y de las autoridades, derivó en una percepción de necesidad de parte de las instituciones políticas por recolectar, monitorizar, analizar y visualizar información relevante de redes sociales de forma continua (Stieglitz y Dang-Xuan, 2013). Para ello, estas organizaciones frecuentemente han replicado las estrategias desarrolladas en el marketing, la comunicación estratégica y las relaciones públicas, enfocándolas en el análisis de las acciones y percepciones, pero también hacia la intervención y la voluntad de incidir.

Estas prácticas, habituales para las grandes y medianas empresas que dependen en parte de la percepción social de sus marcas, han permeado al sector público. Ahí, donde antiguamente un Estado desarrollaba campañas comunicacionales esporádicas y focalizadas, usando principalmente medios de alcance general como la radio, ahora realiza una producción permanente de propaganda, unas veces con alcance general, otras altamente segmentada. Reparticiones que nunca habían visto la necesidad de comunicarse con una audiencia, optan por enfoques disruptivos, interactivos y constantes de comunicación estratégica. En Chile, un caso paradigmático de este cambio de foco es el de la Contraloría General de la República y su personaje Contralorito (Cárdenas, 2018).

Paralelamente, surgen perspectivas que se concentran en aprovechar los nuevos medios para intentar producir un tipo de conocimiento que tradicionalmente era objeto de disciplinas de las ciencias sociales, como la antropología, la sociología o el trabajo social. Significativos estudios exploratorios de este tipo han sido desarrollados para orientar políticas públicas (Bright

y Margetts, 2016), mejorar la respuesta ante desastres naturales (Sakaki, Okazaki y Matsuo, 2010; Wang y Ye, 2018), construir datos sociodemográficos en ausencia de censos o encuestas tradicionales (Cesare, 2018; Brandt, 2020), entre otros fines.

Estas aproximaciones al *big social data* suelen concentrarse en la medición y análisis de los datos relacionados con eventos comunicativos *online* tratándolos como datos que efectivamente dan cuenta de situaciones *offline*, habitualmente simplificando el locus de análisis en consistencia con el tipo de dato disponible, abstrayéndolo de su contexto económico, político, social o cultural de producción. Aunque sus problemas no son los mismos que atañen a las aproximaciones de índole comercial, se trata de estudios instrumentales cuyo objetivo es explorar asuntos puntuales y desarrollar, planificar o sostener alguna actividad comunicativa dirigida a un fin particular o conocer sus efectos. En ese sentido, todavía pueden agruparse en lo que en comunicación se conoce como perspectiva administrativa de la investigación, en tanto no ponen en el centro de su análisis la cuestión de la validez de los datos recopilados ni, de forma más general, se preguntan sobre cómo el poder incide o moldea determinados medios.

### **Microfocalización y democracia**

Dentro de estos enfoques, que presuponen un cierto grado de traducibilidad entre los comportamientos y acciones *online*, medidos a través de metadatos asociados a las interacciones, y los comportamientos *offline*, que están asociados a complejos contextos socioculturales y económicos, también podemos incluir un polémico tipo de investigación administrativa: el rastreo de usuarios y usuarias para, mediante un conocimiento detallado de sus comportamientos en línea basado en el análisis de millones de interacciones y metadatos, crear perfiles hipersegmentados y poder depurar mensajes microfocalizados, buscando mejorar la eficacia de la propaganda política.

Durante años, las herramientas publicitarias de las redes sociales han permitido segmentar de forma relativamente fina a las audiencias mediante una caracterización demográfica que incluye edad, género, educación, nivel socioeconómico e interés en temas específicos, para

ofrecerles publicidad que sea *significativa*. Significativo en este contexto quiere expresar que sea de “interés del consumidor”, una consigna agresiva que han repetido majaderamente las plataformas, pero que en la realidad deriva en algo fundamentalmente diferente. Por un lado, implica que el anunciante es capaz de llegar específicamente a su público y no desperdiciar recursos hablándole a individuos fuera de sus segmentos objetivos. Por otro, implica que la publicidad será significativa no porque sea de interés de quien la recibe, sino porque la máquina se la presenta a quien es más susceptible de interiorizar el mensaje (comprar el producto, votar por el candidato, etc.). Cathy O’Neil ejemplifica esto con el caso de las universidades privadas con ánimo de lucro de Estados Unidos, cuyo público objetivo son los sectores más vulnerables. En tales casos, “publicidad significativa de interés del consumidor” implicaba segmentar los anuncios hacia personas de las zonas más pobres del país, que previamente hubieran hecho clic en anuncios que ofrecieran préstamos de consumo, entre otros factores. Gracias a este tipo de estrategias, entre 2004 y 2014 se triplicaron las matrículas de ese tipo de instituciones (O’Neil, 2018).

Las arquitecturas creadas para satisfacer las orientaciones comerciales que priman en Internet han conformado un sistema en que terceras empresas son capaces de generar segmentaciones altamente detalladas (o al menos eso prometen), incluyendo lo que se ha denominado *psychographic targeting*, que busca distinguir a las personas usuarias sobre la base de sus rasgos de personalidad, deseos y motivaciones. Esto se consigue a partir del cruce de una cantidad y diversidad mayor de datos y la capacidad de asociarlos a identidades específicas gracias a normas permisivas y protocolos poco seguros de las plataformas sociales, así como a métodos ilegales o reñidos con los acuerdos de uso de los sistemas (e.g. *web scraping*) y técnicas de desanonimización.

El caso más conocido de este tipo de aproximaciones es, desde luego, el de la empresa inglesa Cambridge Analytica, que se hizo público en 2018 (Gibney, 2018). A través de una aplicación desarrollada sobre la API pública de Facebook llamada “thisisyourdigitallife”, que incluía un cuestionario de personalidad, el equipo de Cambridge Analytica aprovechó las brechas de la plataforma para recopilar datos de las cuentas de los participantes y también de sus amigos,

alcanzando de esa forma a administrar datos de alrededor de 87 millones de usuarios. A través de un algoritmo, la empresa desarrolló su producto: un canon de treinta y dos tipos de perfiles psicológicos (o psicográficos) a la venta para su uso en campañas políticas (ur Rehman, 2019). Hasta ese momento la microfocalización estaba orientada por “votante probable” y la innovación de Cambridge Analytica fue cruzarla con rasgos de personalidad y psicológicos, inspirándose en un trabajo publicado en 2015 por Youyou, Kosinski y Stillwell (2015). De esta manera, campañas como la de Donald Trump pudieron focalizar a un extremo inédito los anuncios personalizados que hacían llegar a probables votantes, explotando los temores y sesgos particulares de cada perfil.

Nunca estará claro el efecto exacto que esta estrategia tuvo en esas elecciones (o que ha tenido en campañas posteriores), pero hay cierto consenso en que jugó un importante papel.

### **Perspectivas críticas**

Desde las ciencias sociales y los estudios de la comunicación, por otro lado, la investigación de los fenómenos en torno a Internet tuvo en un principio una orientación primordialmente cualitativa y estuvo dirigido fundamentalmente a investigar a los usuarios y usuarias, sus prácticas en los espacios de interacción virtual y a las comunidades que se conformaban en línea (Kollock y Smith, 2002). La comprensión de lo “virtual” o del “ciberespacio” como algo aparte del resto de la vida social, independiente de la “vida real” (Miller y Slater, 2020), que primó en los años en que se fue popularizando Internet, ha permeado hasta algunas de las representaciones que son habituales en la actualidad y aún puede incidir con fuerza en fenómenos de la comunicación política.<sup>1</sup>

---

<sup>1</sup> Así, por ejemplo, cuando el gobierno de Chile dio a conocer en 2019 el informe Big Data desarrollado por Alto Analytics, sus autoridades no consideraban las observaciones realizadas dentro de las plataformas sociales como algo complementario y entrelazado con el resto de la vida social, sino como la única vida social relevante, ya sea porque estimaban que la vida “virtual” era un reflejo preciso pero aislado de la vida “real” o porque era la única esfera

Por otra parte, cuando se abordaban procesos más amplios, o de nivel macro, asociados a Internet, en muchos casos primaba un tecnooptimismo en cuanto a las potencialidades democratizadoras y los nuevos balances de poder que las redes podrían traer consigo. Se realzaban cualidades como la horizontalidad de las comunicaciones, la interacción y la información libre, bajo una concepción de Internet como algo sustancialmente diferente a los medios de comunicación de masas (Castells, 2001).

### ***Big data* y ciencias sociales**

Es especialmente durante la última década cuando diferentes perspectivas académicas sobre Internet y, en particular, sobre las plataformas sociales, han incorporado técnicas y metodologías de análisis de datos masivos. Estos enfoques han adoptado métodos digitales para explorar desde esa perspectiva aspectos más amplios de la vida social contemporánea y no solo fenómenos estrictamente *online* (Snee et. al., 2015). Este tipo de trabajos se distinguen de los enfoques administrativos en que habitualmente admiten la necesidad de integrar diferentes disciplinas para obtener una comprensión más significativa del mundo social y, en general, prestan atención a las limitaciones de estas nuevas fuentes de datos (Schroeder, 2018). Con ello comienzan a primar puntos de vista que propugnan el desarrollo de nuevos e innovadores planteamientos para dar sentido al mundo social, incorporando los datos de las redes sociales, antecedentes que por sí mismos, en comparación con los datos demográficos, históricos o sociológicos, a menudo están descontextualizados.

Dentro de esta línea se han analizado las relaciones entre la adopción global de los medios digitales y el declive de la democracia, utilizando una multitud de variables producto de observaciones en línea (discurso de odio, desinformación, etc.) para ponerlas en relación con

---

susceptible a ser medida en los términos del *big data*. A esta perspectiva que plantea una separación radical entre el mundo físico y el mundo digital se la ha denominado “dualismo digital” (Jurgenson, 24 de febrero de 2011).

datos sobre procesos políticos (polarización, participación, confianza en las instituciones, etc.) obtenidos a partir de estudios transversales, encuestas panel, experimentos de campo, entre otros métodos tradicionales (Lorenz-Spreen, 2023).

Aproximaciones híbridas como las que buscan desarrollar técnicas para estimar aspectos demográficos a partir del cruce de datos obtenidos de las plataformas e información censal o datos oficiales (Sloan, 2013), cuyo objetivo es revelar indicadores que no han sido directamente entregados por las y los usuarios, plantean importantes preguntas éticas. Estas se relacionan con la capacidad de los algoritmos de producir información potencialmente sensible, la validez de los consentimientos atados a los términos de servicio de las plataformas y, en general, con la idea de la privacidad (Williams, 2017). Se trata de brechas que han dejado obsoleta la figura del anonimato en Internet, caracterizada por la clásica viñeta de Peter Steiner (en Internet “nadie sabe que eres un perro”). En todo el mundo, la policía y diversas agencias de inteligencia han sucumbido al sublime digital y a través del monitoreo de las plataformas sociales han descubierto nuevos espacios para ejercer control (Trottier, 2015).

### **Biopolítica del dato**

En atención al tipo de problemáticas que estas y otras aproximaciones traen aparejadas, un número cada vez mayor de investigadoras e investigadores comienzan a dirigir su mirada al dato mismo y los contextos y contornos en que son producidos, apropiados, analizados y usados en relación con el poder, con el capital y con la democracia. De estas miradas críticas, una de las perspectivas más influyentes que emerge es la asociada a los estudios de control y vigilancia relacionados con las tecnologías de la información e Internet (Lyon, 2003; Ball, Haggerty y Lyon, 2012; Marx, 2016). Este enfoque se puede remontar al trabajo del inglés radicado en Australia, Roger Clarke, quien a fines de la década de 1980 acuñó el concepto de *dataveillance* (vigilancia de datos) para describir la transformación que apreciaba en las prácticas de vigilancia hacia los ciudadanos gracias al incremento del rastro de datos que estos producían. Esta línea de análisis ha tenido un amplio impacto en la comprensión contemporánea del papel de las



plataformas y a partir de ella se han acuñado conceptos populares como el de “capitalismo de vigilancia” (Zuboff, 2018).

Shoshana Zuboff, profesora de la Escuela de Negocios de Harvard, define el capitalismo de vigilancia como un “nuevo orden económico que reivindica la experiencia humana como materia prima gratuita para prácticas comerciales ocultas de extracción, predicción y venta”, un sistema que se apodera de la experiencia humana y la convierte en datos sobre el comportamiento para predecir nuestro futuro en beneficio de otros. Algunos autores han disputado el énfasis de Zuboff de la vigilancia por sobre el capitalismo (Morozov, 2019), acento que deja de lado y opaca varios otros procesos de explotación de la información dentro del capitalismo contemporáneo que también son importantes para caracterizarlo. Para Fuchs (2021), Zuboff desestima además el rol del trabajo en el capitalismo contemporáneo y, por lo tanto, el papel de la explotación de los seres humanos que producen los datos, los contenidos y las interacciones sociales en línea que constituyen las bases materiales de la producción de las mercancías que caracterizan el sistema de las plataformas. Couldry y Mejias (2021), en tanto, resaltan las limitaciones que tiene dicho encuadre al poner el foco en los ciudadanos de Estados Unidos y Europa, mientras conceptualiza la datificación y sus demonios como una aberración de las democracias occidentales, desestimando los largos procesos coloniales que han tenido y tienen lugar en el resto del mundo.

En el caso del capitalismo de vigilancia y en general de los estudios de control y vigilancia del capitalismo con orientación hacia lo digital, no son solo los datos generados por usuarios y ofrecidos a terceros por las plataformas sociales los que están en juego, sino que los datos de un amplio rango de sensores dedicados a la captura de lo que Zuboff llama “excedentes conductuales”: teléfonos inteligentes, cámaras, televisores, relojes, prendas de vestir, videojuegos, etc. La mayor parte de estos escenarios de captura de datos escapan al objeto de la presente investigación. No obstante, entre los problemas que resaltan estos trabajos que son relevantes para cualquier análisis de los nuevos medios digitales figura el que los contornos de la acumulación de datos son desiguales y lo que se registra, cuantifica y procesa reproduce las desigualdades e inequidades de la sociedad; y que los mecanismos internos de las tecnologías y

sus efectos son con frecuencia presentados como inexplicables, inextricables o como cajas negras (Thatcher, 2021).

Tales aproximaciones, al igual que este trabajo, se enmarcan en términos generales dentro de lo que muchas autoras y autores llaman perspectiva crítica de la investigación, en oposición a la investigación administrativa que hemos descrito más arriba. Esta amplia gama del estudio sobre los fenómenos de la comunicación, la información y los medios busca explícitamente ponerlos en relación con sus condiciones sociales, culturales, históricas, económicas y políticas de producción, desde posturas teóricas comúnmente vinculadas al interés público.

## **Economía política de la comunicación y la información: un enfoque para el análisis crítico de las agencias de monitoreo de plataformas**

En 2019, en su libro *Monitored. Business and surveillance in a time of big data*, Peter Bloom se refería al estado actual del capitalismo con las siguientes palabras:

El elixir del big data, las tecnologías inteligentes, la inteligencia artificial (IA) y los algoritmos ocultos son fuerzas poderosas, casi mágicas, que son incomprensibles y, no obstante, la clave de nuestra salvación. Nuestro anhelo por esta liberación tecnológica está íntimamente ligada a una necesidad creciente de recolectar más datos sobre nosotros mismos y los demás. (Bloom, 2019, p. 42)

Esta hambre insaciable por datos, que Kennedy formula como “deseo por los números” (Kennedy, 2016), es la expresión hegemónica de un capitalismo que transforma su necesidad voraz de ganancias en una demanda infinita de monitoreo digital a través de una variedad creciente de sensores. Parece haber un giro cultural: el apetito por los datos ya no es exclusivo de los gobernantes, ni se hace a espaldas de los gobernados. Se ha democratizado la percepción de que en el *big data* está la “clave de nuestra salvación”. Para Bloom, es la descomposición social asociada a los reordenamientos neoliberales lo que ha provocado una percepción generalizada de

caos, lo que paradójicamente alienta la necesidad de un mayor monitoreo en pro de ofrecerles a las comunidades una renovada “seguridad ontológica” (Bloom, 2019, p. 43).

La salvación mediante los datos está relacionada con una plétora de dispositivos (en su sentido literal): tecnologías inteligentes, tecnologías móviles, tecnologías de autoseguimiento (*self-tracking*), etc. No obstante, es innegable el lugar central que las plataformas sociales han tenido en la difusión de estos modelos ideológicos y que tienen en la configuración de la sociedad monitorizada que ha sido moldeado por el capitalismo.

### **La mercantilización de la experiencia humana**

Ya que nuestra tarea aquí es describir cómo opera el sector de las AMP, debemos primero entender cómo se inserta la experiencia humana en el sistema de las plataformas, qué implicancias tiene su transformación a dato digital y qué caracteriza a dichos datos; en definitiva, cuál es la relación de estos datos con la experiencia humana capturada. En segundo lugar, nos proponemos entender cómo ese dato, elaborado por las plataformas sociales publicitarias para sus fines comerciales, puede ser reutilizado de diferentes modos.

Gandy (2011) ha destacado que, debido a que la información es intangible, es también muy difícil de manejar, controlar o limitar el acceso a ella. Las tecnologías detrás de las plataformas sociales son en gran parte tecnologías perfeccionadas para la gestión de información personal, por lo tanto, se encuentran en una posición ideal para apropiarse y establecer control sobre esta “mercancía extremadamente problemática” (Gandy, 2011). Como ha demostrado el capitalismo en innumerables ocasiones, una de sus características fundamentales como formación social es la transformación generalizada de valores de uso en valores de cambio, extendiéndose en áreas que previamente ocupaban esferas diferentes a la del mercado.

Este mismo proceso se ubica en el centro de las dinámicas que se inician con las actividades humanas en las plataformas sociales. En un proceso de mercantilización inmanente, los mismos valores de uso dan origen a diferentes valores de cambio: por una parte, los datos

permiten la producción de aquellas mercancías que son relevantes para el negocio principal de las plataformas sociales publicitarias, la producción de audiencias altamente segmentadas. Aquí nos interesa una dinámica de mercantilización que ocurre de forma paralela utilizando los mismos datos y que culmina con la producción de las mercancías que comercian las AMP. Por lo tanto, es ineludible acudir a una perspectiva teórica que nos posibilite encuadrar el proceso de elaboración social de dicha mercancía. En los siguientes pasajes desarrollaremos un marco teórico con este foco en mente.

### **El punto de vista de la economía política de la comunicación**

Con frecuencia, los portavoces de Facebook, X/Twitter o LinkedIn ponen énfasis en una narrativa que dibuja a las plataformas como una herramienta completamente neutral para “conectarse” y “expresar la individualidad”. En oposición a estas pretensiones corporativas y en contraste con la orientación instrumental que el análisis administrativo les otorga a las plataformas, el enfoque seleccionado para afrontar la presente investigación es el de la economía política de la comunicación y la información (EPC), también llamada economía política crítica de la comunicación (Hardy, 2014). Se trata de un amplio campo donde se intersectan la economía, la ciencia política, las comunicaciones y los estudios culturales, una ciencia integral de la sociedad (Knoche, 2021) que problematiza la injusticia e inequidad en los sistemas de poder y es crítica del determinismo tecnológico y de la visión optimista sobre el potencial emancipador de Internet (Wasko, Murdock y Sousa, 2014). Esta perspectiva es crítica de las deficiencias del capitalismo y las élites de poder y sus fundamentos se remontan a la revisión de la teoría marxista desarrollada por la Escuela de Frankfurt (Hardy, 2014). Particularmente relevante para este punto de vista es la comprensión del conocimiento como algo que no está eximido de valores, sino que trae consigo inscritos los valores de la cultura y los sujetos que lo producen.

Para la economía política crítica de la comunicación, las diferentes maneras en que se organizan y financian los sistemas de comunicación en una sociedad tienen implicaciones en el rango y la naturaleza de los contenidos que se distribuyen y para la forma en que estos son consumidos y usados. Ha sido ampliamente discutido que esta perspectiva debe evitar caer en un

determinismo económico (Golding y Murdock, 1997); el punto de esta orientación está en afirmar que los productos de las industrias mediales son al mismo tiempo económicos y culturales (Hardy, 2014). Por ello, para poder aproximarnos a las dimensiones simbólicas de la producción y circulación de significados (que permiten a las personas entender, comunicarse y actuar en el mundo) debemos también comprender las variables económicas de esa producción.

Esta línea de investigación tiene una larga historia, que se remonta a los años inmediatamente posteriores a la Segunda Guerra Mundial en Estados Unidos. Allí la figura que sentó las bases de este tipo de aproximaciones fue el ya citado Dallas Smythe desde su posición en el *Institute for Communication Research* de la Universidad de Illinois. De esta tradición también forma parte Herbert Schiller, destacado investigador que desarrolló uno de sus más importantes trabajos de campo analizando el sistema de los medios bajo el gobierno de Salvador Allende en Chile (Schiller, 2019).

Otras investigadoras e investigadores que se inscriben total o parcialmente dentro de la EPC, incluyendo sus ramas europeas, son Vincent Mosco, Robert McChesney, Dan Schiller, Graham Murdock, Peter Golding, Janet Wasko, Jonathan Hardy, Armand Mattelard, Oscar H. Gandy y Helena Sousa, entre otros. Asimismo, sobre las bases de la EPC se ha desarrollado una economía política crítica particularmente enfocada en Internet y los nuevos medios digitales, entre cuyos exponentes están Christian Fuchs, Mark Andrejevic, Stefania Milan, Scott Timcke, Manuel Castells, Nick Srnicek, José van Dijck, Nick Couldry y Ulises Mejias, por mencionar a algunos.

La influencia de esta corriente en Latinoamérica ha configurado la llamada economía política de la información, la comunicación y la cultura, que a la tradición crítica sobre los medios incorpora el debate político y económico sobre el desarrollo y subdesarrollo, el decolonialismo y otras tradiciones particulares (Becerra y Mastrini, 2014; Bolaño, 2019).

## **Algunas observaciones relevantes realizadas por la EPC**

El canadiense Vincent Mosco define de forma general a la economía política de la comunicación como el “estudio de las relaciones sociales, particularmente las relaciones de poder, que mutuamente constituyen la producción, distribución y consumo de recursos, incluyendo recursos de comunicación” (Mosco, 2009).

Para el inglés Jonathan Hardy (2014), la EPC aplicada al estudio de los medios está constituida por tres áreas principales. Primero, el estudio de cómo operan las industrias de la comunicación, con foco en la propiedad, las finanzas, los modelos de negocio, los mecanismos de soporte (como la publicidad) y el efecto que tienen las políticas gubernamentales. La segunda área es la influencia y las consecuencias de las diferentes formas en que se organizan los medios: comerciales, estatales, públicos y sus combinaciones. El tercer tema principal es la relación entre contenidos mediales, sistemas comunicacionales y la estructura social en un sentido amplio. Es decir, para entender una configuración medial específica, la indagación requiere comprender cómo se producen y distribuyen sus productos en una sociedad dada o qué lugar tienen en una estructura particular de dominación.

Así, uno de los focos que caracterizan los estudios englobados bajo esta orientación está puesto en cómo las mercancías de comunicación son producidas, distribuidas y consumidas y a qué relaciones de poder dan forma dichos bienes. Para esto, es importante entender a los medios, incluyendo los medios sociales digitales agrupados en las plataformas sociales publicitarias, como un problema político de la sociedad. Bajo esta óptica, las medidas políticas, subvenciones, estructuras e instituciones creadas para controlar, dirigir y regular los medios tienen un rol central en el carácter del sistema mediático (McChesney, 2013) y de sus relaciones sociales de producción. Esta observación tiene una importancia fundamental, ya que encadena cualquier análisis descriptivo desarrollado bajo esta perspectiva con una intencionalidad prescriptiva: es un tipo de conocimiento que admite y fomenta el cambio social.

De esta forma, según el académico y activista Robert McChesney (2013), la EPC tiene como gran objetivo resolver la problemática de los medios a través de la creación de un sistema mediático más apto para la promoción de los valores democráticos. Para avanzar hacia ese objetivo, es crucial conocer de qué forma operan los sistemas mediáticos existentes. La emergencia de los nuevos medios digitales y lo dinámico del sector plantean un serio desafío a la capacidad de las personas de adaptarse, comprender los modos en que operan y ser críticas respecto al funcionamiento de estos. Lo que es extensivo a las y los investigadores de la comunicación y la sociedad.

No obstante, siguiendo la perspectiva de la EPC, en el ámbito de los medios digitales de Internet, importantes trabajos críticos se han concentrado en el rol dominante de las grandes corporaciones detrás de las principales plataformas sociales y sus relaciones con el poder militar y político (Mosco, 2017; Timcke, 2017), en el análisis del ecosistema en que se desenvuelven las plataformas (van Dijck, 2013) y de sus infraestructuras (Mosco, 2015), en los procesos de mercantilización de diferentes aspectos de la experiencia humana relacionados con las actividades de los y las usuarias en los medios digitales (Cohen, 2008; Kang y McAllister, 2011; Shklovski, 2009 ; Seignani, 2013), en el problema y las características del trabajo digital (Caraway, 2011; Fuchs, 2012a; Nixon, 2014; Postigo, 2016), así como en las relaciones y tensiones entre la economía política de las plataformas, su arquitectura y diseño algorítmico y la censura y la protesta social (Dencik y Leistert, 2015). Conectada con esta última, una importante línea de investigación se ha orientado al registro y control constante al que son sometidos los usuarios y usuarias de las plataformas y la vigilancia como “clasificación social” (*social sorting*), un medio para verificar identidades pero también para asignarles valor (Lyon, 2003; Scholz, 2013; Bekkers, 2013; entre otros).

Reveladoras de los procesos que están en curso en el capitalismo orientado hacia lo digital son las observaciones de Mejias y Couldry sobre el colonialismo de datos. El énfasis de estos investigadores no está puesto en la explotación en el contexto digital (Terranova, 2000), sino en otro factor clave de la economía política marxista: la apropiación originaria. Desde esta

perspectiva, la escala temporal de observación de los fenómenos que configuran el capitalismo contemporáneo debe ser mucho más amplia y abarcar todo el ciclo global de colonialismo entrelazado con capitalismo que ha dado forma a las actuales relaciones Norte-Sur globales. La apropiación de la experiencia humana en general, para su mercantilización a través de los mecanismos que disponibilizan las tecnologías en la actualidad, no sería sino otra etapa de la apropiación colonial (Mejias y Couldry, 2019).

Bajo el colonialismo de datos, en lugar de la apropiación de recursos naturales y trabajo que caracterizó al colonialismo histórico, lo que ahora está siendo apropiado es la vida humana misma a través de su conversión en datos. Es la vida humana, en su sentido más amplio, lo que se convierte en un insumo directo para la producción capitalista de mercancías. El capitalismo de datos se caracteriza por el surgimiento de un nuevo tipo de relaciones sociales, las “relaciones de datos”, cuya aparente naturalidad enmarca a los datos como algo de lo que el capital puede apropiarse con total legitimidad, otorgándole estabilidad al sistema colonial (Couldry y Mejias, 2020).

Diferentes vertientes de la EPC han explorado y discutido con bastante profundidad los macroproblemas que emergen del sistema de las plataformas y han intentado ponerse a la par de los desarrollos tecnológicos y las rápidas modulaciones y transformaciones que las grandes corporaciones protagonizan continuamente dentro del neoliberalismo. Sin embargo, los trabajos que involucran específicamente a los actores intermediarios que aquí nos interesan son relativamente escasos y aquellos que recurren a la recolección de datos empíricos, más aún. Incluso en el ámbito de la prensa, la presencia de estas compañías de monitoreo de redes sociales es casi inexistente, salvo por contadas ocasiones (el caso de mayor relevancia es el que detallamos en la introducción de este trabajo), pese a la proliferación de estos servicios y a que incluso los medios de comunicación recurren a ellos habitualmente.

En la siguiente sección, tomando como base las investigaciones disponibles en el campo de la EPC y algunas fuentes heterodoxas, retomaremos el examen de las AMP, proponiendo un



encuadre teórico para el análisis de las relaciones sociales de producción que caracterizan al sector.

## **La EPC aplicada al problema del monitoreo de plataformas en Chile**

Al interrogar a las AMP sobre las promesas explícitas e implícitas en sus mercancías analíticas, los límites de sus facultades explicativas y predictivas, así como sobre las relaciones sociales de producción que constituyen al sector, buscamos aproximarnos a una lectura sobre sus potencialidades efectivas y los posibles desafíos que presentan para la sociedad democrática.

Para hacerlo, nuestra propuesta de investigación reconoce como aproximación principal el análisis de la oferta de servicios y productos que estas compañías ofrecen en Chile. La conceptualización que hacemos de estas compañías intermediarias destaca, asimismo, su posición dentro de una cadena global de producción de valor, derivada del tratamiento de los datos por parte de las plataformas sociales.

En el siguiente apartado describiremos la configuración general que tiene esta cadena de creación de valor, es decir, el recorrido a través del sistema de las plataformas de los datos de la actividad humana digitalizada, desde su origen o captura, hasta su transformación en mercancías analíticas por parte de empresas del sector AMP.

A continuación analizaremos algunas de las principales tensiones conceptuales y consecuencias de este modelo. Aquí discutiremos la pertinencia de las metáforas extractivistas, profundizaremos en el relevante rol de las API y relacionaremos las mercancías de las AMP con productos análogos del sistema de los medios de comunicación de masas. Tras esto, ahondaremos en algunas de las consecuencias sociales que se desprenden de los antecedentes recogidos.

Finalmente, en los últimos dos apartados de este capítulo, abordaremos el rol de la clientela del sector AMP y explicitaremos el papel del colonialismo de datos para nuestro marco teórico-conceptual.

### **La cadena de valor de los datos**

El foco de esta investigación apunta a una parte de la cadena productiva que caracteriza al sistema de las plataformas sociales publicitarias. Como ha quedado definido, estas son arquitecturas digitales en línea programadas para permitir que dos o más personas o grupos interactúen. Estas interacciones son una fuente de datos. El procesamiento de estos datos conforma el centro del negocio de las empresas propietarias de dichas plataformas.

La cadena de valor de las plataformas se inicia, entonces, con la experiencia humana digitalizada. Este proceso se produce a través de lo que Mejias y Couldry (2019) llaman “relaciones de datos”, nuevos tipos de relaciones humanas que permiten la extracción de datos para su mercantilización, aspecto que enmarcan en el proceso general de colonialismo de datos. En estos procedimientos de enajenación, las plataformas no se limitan a “grabar” y almacenar: sus sensores y algoritmos tienen un rol activo en la selección de los datos y creativo en la atribución de metadatos. Estos datos constituyen un recurso que adquiere en un primer lugar un valor de uso: las empresas propietarias de las plataformas usan dichos datos para perfeccionar sus productos. Este perfeccionamiento consiste, en términos generales, en conseguir que las y los usuarios permanezcan más tiempo y entreguen más datos a los sensores de las plataformas. Hasta ese momento, es irrelevante el trabajo que pudiese estar inscrito en dichos datos, o en otras palabras, no nos encontramos frente a mercancías (Marx, 2008).

Pero simultáneamente se desarrollan procesos tendientes a dotar de valor de uso social a dicho recurso, los que habilitan la transformación de estos datos en valores de cambio. Este es el proceso mediante el cual las plataformas sociales publicitarias crean dobles de datos y, con ello, segmentos de audiencia disponibles para ser comercializados en el extenso mercado de la

publicidad en línea. En este sentido, las plataformas son tanto un medio de comunicación como un medio de producción (Fisher, 2016).

No obstante, hay un remanente de datos que tiene otro destino, que pareciera no poseer un valor de uso inmediato para las plataformas digitales y, sin embargo, estas hacen una importante inversión de capital para volverlo disponible, como recurso, a través de sus API. Las compañías de monitoreo de redes sociales obtienen mediante estas API los datos que minan y analizan. Una misma plataforma digital puede ofrecer diferentes API y cada API ofrece diferentes parcialidades de datos y metadatos. Existen otras fuentes de datos disponibles, desde luego, pero las API son la fuente que caracteriza al sistema de las plataformas.

Las empresas que ofrecen servicios de minería de datos se ubican en este eslabón de la cadena de valor del sistema. Puede existir en este punto un intercambio comercial entre las plataformas, como proveedoras de los datos empaquetados de diversas maneras, y las agencias, pero no es la norma del mercado. Durante muchos años, lo habitual ha sido que las plataformas otorguen acceso gratuito a sus API, las que se diseñan adaptadas a diversos casos de uso. Los escenarios de uso para investigadores generalmente están excluidos de cualquier tipo de pago.

Al no existir un intercambio económico, en la práctica pareciera que las empresas mineras de datos acceden a estos como si se tratara con recursos o materias primas. Explotan estos recursos mediante diferentes herramientas de “minería” y a través de procesos más o menos estandarizados producen nuevos bienes. Estos ingresan al mercado transformados en mercancías analíticas: informes, minutas, presentaciones, *dashboards*, etc. Debido a sus altos costos, en términos generales, y a lo especializado de sus fines presuntos, estos productos son patrimonio de un mercado relativamente exclusivo formado por empresas de tamaño medio a mayor, como bancos, mineras y universidades, organizaciones de la sociedad civil bien financiadas, gobiernos e instituciones del Estado, partidos políticos, consultoras y figuras públicas, entre otros.

La Figura 1 presenta de forma esquemática la cadena de producción de valor que acabamos de describir. En los siguientes apartados discutiremos algunos de los nudos críticos de esta cadena.

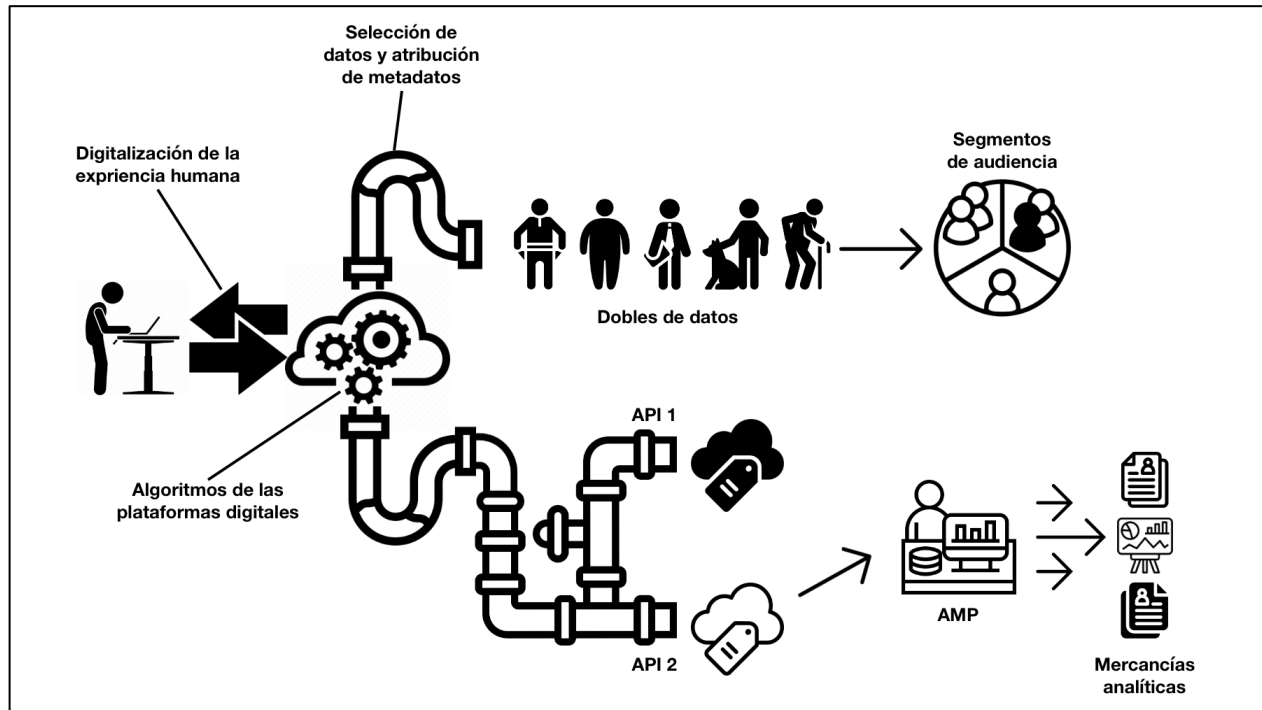


Figura 1. Cadena de valor de los datos. Elaboración propia.

### Sobre el origen de los datos y la pertinencia de la metáfora extractivista

Srnicek (2018) afirma que los datos, sin importar su posición dentro del sistema de las plataformas, son un recurso natural o “materia prima” (*raw material*, en el original). El autor canadiense es explícito en señalar que el concepto lo usa en el mismo sentido que lo hace Marx, a saber, como un objeto que “ya ha pasado por el filtro de un trabajo anterior” (Marx, 2008). Sin embargo, se trata de un término que en este contexto carga consigo un sentido equívoco significativo y crucial para la comprensión del papel de los datos.

Hoy es un lugar común afirmar que los datos son el “nuevo petróleo” (Bhageshpur, 2019), lo que resulta una idea extraña cuando consideramos a los datos y la información como cosas intangibles, pero más admisible cuando pensamos en estos como objetos con capacidad de intervenir y afectar al mundo (Hacking, 1996). Srnicek secunda esta idea y sostiene que, como el petróleo, los datos son objetos preexistentes en la naturaleza y que su fuente natural es la actividad humana. Sin embargo, sostener que los datos masivos que provienen de las actividades de las personas usuarias de redes sociales poseen una materialidad previa a su procesamiento por parte de las plataformas es, cuando menos, controversial. Estos no son desprendidos de la naturaleza por medio del trabajo para, por intermedio de ese filtro, poder pasar a ser considerados materia prima en un sentido marxista. Su materialidad como recursos, en cambio, es consustancial a su captura: el dato es tanto enajenado de la experiencia humana como creado a través del procedimiento de registro mediado por sensores y algoritmos. Este procedimiento no es un simple acto de grabar y almacenar, es un acto creativo de atribución de metadatos que, entre otras cosas, permite la elaboración de infinidad de dobles de datos y la producción de segmentos de audiencia comercializables.

Que las plataformas jueguen un rol activo en la creación de los datos no significa que deje de ser un procedimiento de apropiación esencialmente ilegítimo de la experiencia humana. Sin embargo, implica que dicha expropiación no constituirá una cartografía que reproduzca el territorio de lo humano en su totalidad, sino que un mapa que solo demarca ciertas vías de esa experiencia enajenada y en el que la toponimia ha sido alterado según criterios opacos.

Estos nos lleva al segundo problema del concepto de “materia prima”. En un sentido literal, el dato más “crudo” (*raw*) al que podemos acceder dentro del sistema de las plataformas ya es un recurso semiprocesado, en que ya está inscrita una intención, filtro o propósito por parte de la plataforma y sus intereses comerciales. Al momento en que las empresas intermediarias recogen dichos datos de las plataformas, no estamos ante nada parecido a un material crudo (*raw material*). La pretendida pureza de los datos es otra faceta de la apología a la falsa desintermediación.

Pero si bien la experiencia humana no está constituida por datos esperando a ser explotados, como propone la analogía con las materias primas (profundamente enraizada en el lenguaje de cualquier discusión sobre las plataformas e Internet, incluyendo esta sobre la “minería de datos”), se trata de una retórica digna de al menos un breve análisis.

En lo sustancial, lo que presenciamos es la prolongación de un discurso que se aferra a un proyecto ideológico de larga historia. El tropo de los “recursos naturales” ha servido en el pasado para justificar la extracción de aquello que se tiene por “primitivo” y “no refinado” (Crawford, 2021). Se trata de expresiones que buscan poner de manifiesto un estado de incompletitud. Como han enfatizado Ulises Mejias y Nick Couldry (2020), el concepto mismo de *raw material* hunde sus raíces en el pasado colonial, contexto en el que implica que algo está disponible para su explotación por parte de alguien, un sujeto con coordenadas geográficas, raciales y de género bien definidas. La experiencia humana no es una materia prima, pero el sistema de las plataformas ha perpetuado y potenciado la analogía, legitimando con ello la apropiación y recreación de los datos humanos que caracteriza al capitalismo contemporáneo.

La metáfora extractivista, asimismo, nos permite corroborar lo pertinente de la reflexión planteada por Ludwig, Hilborn y Walters (1993) respecto a los recursos, cuando señalan que “es más apropiado pensar que los recursos gestionan a los seres humanos que lo contrario” porque mientras “más grandes e inmediatas sean las perspectivas de ganancia, mayor será el poder político usado para facilitar una explotación ilimitada.” (Ludwig et al., 1993) Estos autores ejemplifican su punto con la fiebre del oro: cuando aparecen perspectivas de ganancias grandes e inmediatas, los políticos y los gobiernos tienden a aliarse con grupos de interés para facilitar dicha explotación. La constante histórica del capitalismo es que los recursos son inevitablemente sobreexplotados, exponiendo en su camino a los seres humanos a diversos perjuicios.

La idea de una “fiebre de los datos” no es nada fuera de lo común (p. ej. Limón, 2018), como tampoco lo son las alianzas público-privadas para explotarlos con la mayor intensidad posible. Mientras que gran parte de los desarrollos tecnológicos, incluidos aquellos relacionados con el *big data*, han sido subvencionados por los gobiernos, políticos y autoridades en todo el

mundo realizan gestiones en busca de suavizar regulaciones medioambientales o reducir los aranceles a las grandes corporaciones tecnológicas para que, por ejemplo, estas instalen sus data centers dentro de sus territorios (Méndez, 2023).

### **La relevancia de las API y el modelo de subvenciones cruzadas**

Luego de la creación de los datos, junto con su empaquetamiento para los fines del negocio central de las plataformas publicitarias, una etapa fundamental de esta cadena de valor es el momento en que un volumen variable de dichos datos, en condiciones también variables, de acuerdo a normativas y criterios internos de las plataformas, se disponibilizan para su descarga por terceros actores a través de las API. Diferentes entidades acceden a los datos por esta vía, desde investigadores académicos hasta profesionales del marketing, pasando por *influencers* y máquinas que se autoentrenan (como ChatGPT). La descarga y procesamiento de estos datos por distintos métodos puede estar orientada a variados fines. Es por ello que las plataformas suelen contar con una variedad de API de acuerdo a diferentes escenarios de utilización.

van der Vlist, Helmond, Burkhardt y Seitz (2022) destacan que aunque las API pueden parecer objetos técnicos “microscópicos”, son elementos centrales de las infraestructuras materiales de las plataformas y articulan los procesos de datificación. La racionalidad explícita en Silicon Valley que justifica la implementación extensiva de API públicas es fomentar al ecosistema de desarrolladores independientes para que “agreguen valor” a la plataforma a través de la creación de nuevas aplicaciones que operen en torno a ella, fomentando la innovación, la creación de funcionalidades complementarias y la atracción de nuevos usuarios y usuarias (Bonasio, 2014). Dada la centralidad de la tecnología de las API públicas en la cadena de valor de los datos sociales y el relevante rol que tienen para el sector de las AMP, en este apartado evaluaremos brevemente sus lógicas implícitas.

Como hemos dicho, las plataformas publicitarias producen primordialmente audiencias y atención altamente segmentada y en ese proceso generan un cuantioso remanente de *big social data*. Uno de los usos que hacen las plataformas de estos datos masivos es dar acceso a ellos (o al

menos, a una porción de ellos) a través de sus API. Esto plantea un problema interesante, ya que la mayoría de las veces este acceso se entrega de forma gratuita a terceras empresas, como las AMP. Estas compañías lucran generando mercancías derivadas (como el informe Big Data) a partir de recursos que han obtenido gratuitamente.

Hasta aquí hemos sostenido que la mayor parte de las decisiones que dan forma al sistema de las plataformas están orientadas por sus premisas comerciales. ¿Es posible que en este punto de la cadena productiva haya algo “gratis”? ¿Las plataformas se conforman con fomentar la actividad de desarrolladores independientes, mientras de forma paralela miles de empresas del sector AMP sacan provecho de “sus” datos? ¿Existe una lógica comercial ahí que no estemos viendo? Trabajar con datos de este tipo y de esta magnitud implica un proceso de depuración complejo y crecientemente costoso (Matei, Jullien y Goggins, 2017) a medida que alrededor de las API se van acumulando desarrolladores, investigadores y usuarios y usuarias. ¿Qué motiva la cesión gratuita de un recurso que es extremadamente oneroso de sostener? Una posibilidad es que estamos frente a un ejemplo de subvención cruzada.

La estrategia de subvenciones cruzadas es una práctica común en el capitalismo contemporáneo (Srnicek, 2018). Consiste en la financiación de una parte no rentable de una empresa por una parte más rentable de la misma (Rutherford, 1995). En el caso de plataformas publicitarias como Facebook, la programación y mantención de las API sería una rama del negocio subvencionada por otras ramas de la compañía para permitirle entregar acceso gratuito a una parcialidad de los datos a terceros actores. El objetivo de ceder a un público tan extenso un tipo tan valioso de activo podría estar ubicado en la tentativa de otorgarle valor a los productos generados por aquellas ramas comerciales que sí cobran por sus bienes, en particular aquella dedicada al negocio principal de la empresa: la subasta de espacio publicitario.

De esta manera, cuando una corporación como Alphabet o Meta ofrece a ciertos actores, entre los que están las AMP, la descarga “libre” de un determinado conjunto organizado de datos procesados (cuyas posibilidades están prefiguradas por la compañía en relación con sus propios fines) lo que propicia con ello es un mercado, una necesidad y una teoría sobre lo que es



relevante y lo que no. Lo que Google, Facebook o X/Twitter entregan “gratis”, en otras palabras, son categorías sobre cómo medir el éxito, el impacto y la influencia dentro de sus plataformas. Todos factores, por supuesto, que se encuentran en el centro de su propio negocio.

En principio, las modulaciones que permite realizar una API facultan a las plataformas tanto a facilitar el desarrollo de un ecosistema de aplicaciones a su alrededor como a mantener el control sobre este en beneficio de sus propios intereses. Las API de plataformas como Facebook han evolucionado desde interfaces simples hasta complejos acuerdos de gobernanza, donde las especificaciones técnicas de las API sirven para hacer cumplir las políticas y estrategias de las plataformas (van der Vlist, 2022).

Al hacerlo, las plataformas sociales publicitarias difunden una determinada manera de organizar el mundo que es compatible con sus modelos de negocios. Así, por ejemplo, la versión vigente de la API de búsqueda de Twitter hasta julio de 2020, solo permitía buscar en una muestra de tweets “relevantes” publicados en los últimos siete días. Además, la documentación oficial especificaba que dicha API no pretendía ser una fuente exhaustiva de *tweets*, que no todos los *tweets* se indexarían ni estarían disponibles para su búsqueda (Twitter, s. f.). Esto, sin detallar los criterios que delimitaban tales conjuntos “relevantes”. Asimismo, hasta julio de 2020 no era posible usar la API de Twitter para acceder a la información de los resultados de las encuestas en los tweets ni recopilar los “hilos” de conversación, pese a que la primera función había sido incorporada casi cinco años antes y la segunda dos años y medio antes (Twitter, 2020). Es decir, mediante el uso de su propia API de búsqueda, hasta ese momento no era posible generar ninguna clase de conocimiento respecto a maneras asentadas de comunicarse dentro de la plataforma.

En suma, dependiendo de la API específica usada para recabar información desde una plataforma, el momento en que se accediera a dicha API y el nivel de acceso solicitado (algunos gratuitos, otros pudiendo implicar un costo significativo), pero también de acuerdo a modulaciones mucho más lejos del alcance de investigadores e investigadoras regulares (como la definición de lo que es una muestra “relevante” para la API de búsqueda de Twitter), las propiedades específicas y el contorno de los datos, que supuestamente deberían ser referenciales

de la experiencia humana de un cierto conjunto de sujetos, podrían adquirir formas muy diferentes.

La opacidad de las API es una *feature*, un atributo distintivo de estas, no un *bug*. Las plataformas como X/Twitter ofrecen un mínimo de información sobre los mecanismos de muestreo de datos que usan sus API. Asimismo, dentro de Términos de Servicio, se prohíbe a las y los usuarios “investigar, escanear o probar la vulnerabilidad de cualquier sistema” dentro de la plataforma, lo que en teoría impide a terceros intentar entender los entresijos de sus API. Pese a ello, se ha informado de variaciones sustanciales respecto a los datos que se pueden obtener según se use una u otra API de dicha plataforma (Pfeffer et. al., 2023).

En estos contextos, es tanto o más relevante aquello que se incluye en estas muestras, lo que conforma nuestros dobles de datos públicos, como lo que queda afuera. Los datos recopilados vía las API no solo dejan afuera lo que sus algoritmos consideran no relevante, también quedan fuera los referentes directos de las actividades que constituyen la experiencia humana capturada: los contextos espaciales y temporales compartidos de los usuarios y usuarias en una conversación, los “links” que los humanos hacemos con aquello que está afuera de la pantalla y que, no siendo necesario nombrarlo ya que es patrimonio común de un momento dado, queda fuera del registro.

Son los datos disponibles lo que da forma a las respuestas que pueden obtener los investigadores que realizan minería de datos, pero esos datos también delimitan lo que puede ser preguntado. Y es claro que no todas las preguntas están disponibles. Además de las modulaciones comerciales arbitradas por los algoritmos, las API pueden plantear otro tipo de distorsiones a la pretensión de neutralidad del *big data* al establecer sesgos regionales que reproducen inequidades económicas, históricas, geográficas, raciales y sexogenéricas (Wiggers, 2020). En esta línea, por poner un ejemplo, a 2024 la API para investigadores que ofrece la popular plataforma china TikTok, solo está disponible para académicos de Estados Unidos y Europa (TikTok, s. f.).

Las plataformas sociales publicitarias se han promovido a sí mismas como espacios neutros, campos de sociabilidad igualitarios y no demarcados, a la vez que ofertan como datos “crudos” o “en bruto” a los datos semiprocesados que entregan a través de sus API a diversos actores, como los protagonistas del sector del monitoreo de redes sociales. Sin embargo, las plataformas como sus datos encarnan una política particular. Aunque se presenta como algo técnicamente establecido, constituido por herramientas y normas presuntamente carentes de motivaciones humanas, pretendidamente precisas y mecánicas, se trata de un reparto de lo sensible histórico particular que orienta determinadas fuerzas productivas y relaciones sociales de producción.

### **Ratings y números**

Los datos dentro de las plataformas digitales tienen la propiedad de ser pluripotenciales: a partir de los mismos datos se pueden producir una variedad de mercancías. Asimismo, el curso que sigue el dato al interior y fuera del sistema de las plataformas corresponde a una forma de mercantilización inmanente, es decir, de cómo las mercancías producen nuevas mercancías (Mosco, 2009).

Algo similar ocurre en los medios de comunicación de masas. Lo inmaterial y esquivo de la mercancía primordial de dichos medios (¿los mensajes o la atención de las audiencias?) es un efecto del dinamismo de una formación social, el capitalismo, insaciable por abarcar cada espacio de la vida humana. En 1984, Eileen Meehan propuso un tercer contendiente como mercancía fundamental para el caso de la televisión: los ratings, la transformación de los espectadores (o, más precisamente, de la conducta o interacción de los espectadores frente a un televisor) en una serie de números que representan audiencias verificadas.

Varias características emparentan a los ratings televisivos con las mercancías analíticas de las AMP. En ambos casos se trata de productos de empresas dedicadas específicamente a la producción de dichas mercancías, lo que implica prácticas, culturas e infraestructuras particulares. Los clientes consideran ambos tipos de productos como “mediciones imperfectas

pero básicamente científicas sobre el comportamiento de las audiencias” (Meehan, 1984). En ambos casos la forma esencial del producto es un informe o presentación cuyo foco son los números.

Las críticas que se le hace al rating televisivo como instrumento suelen concentrarse en aspectos como el tamaño de la muestra y la recolección de los datos, incluyendo los problemas de transparencia. No obstante, Meehan señala que estas críticas solo serían válidas si las empresas que los elaboran y comercializan realmente produjeran investigaciones en lugar de mercancías, artículos fabricados para ser vendidos en un determinado entorno empresarial. Para el primer caso, se aplicarían las reglas de la investigación en ciencias sociales y estadísticas; para el segundo caso, las normas que están en juego son las de la efectividad, la productividad y la rentabilidad.

El argumento de Meehan es que la evidencia muestra que los ratings no son otra cosa que herramientas manipulables según los intereses de las cadenas de televisión para justificar sus políticas programáticas. El sostenimiento de técnicas de medición inadecuadas e imprecisas por parte de las empresas que calculan el rating ocurre porque tales deficiencias son útiles para sus clientes. En ese sentido, se identifican claramente con una mercancía: “los ratings per se ya no deben ser tratadas como informes sobre el comportamiento humano, sino más bien como productos, como mercancías moldeadas por exigencias comerciales y estrategias corporativas” (Meehan, 1984, p. 221).

Las semejanzas entre el mercado de los ratings televisivos y el sector que aquí nos ocupa se vuelven más evidentes al comparar el trabajo de Meehan con los hallazgos de Helen Kennedy. En su libro de 2016, *Post, Mine, Repeat*, la académica de la Universidad de Sheffield dedica el quinto capítulo, *Commercial Mediations of Social Media Data*, a estas empresas.

Ya hemos detallado algunos de los aspectos de dicha investigación en las páginas previas. Kennedy, junto a cinco asistentes, desarrolla probablemente la primera aproximación empírica al sector, llevando a cabo catorce entrevistas entre 2012 y 2013 a trabajadores de compañías de

Reino Unido, Estados Unidos, España, Noruega y Sudáfrica que entre sus servicios ofrecían minería de datos de plataformas sociales. A diferencia de la presente investigación, Kennedy pone su foco principalmente en los “mineros de datos”, antes que en el sector mismo o sus productos, discutiendo los problemas éticos y los valores de los trabajadores de este sector, confrontándolos con aspectos como la vigilancia, la invasión de la privacidad y la transparencia. No obstante, de su análisis extrae importantes indicios que serán útiles para nuestra investigación.

Para indagar en qué hacen estas compañías y sus trabajadores, Kennedy recurre fundamentalmente a los testimonios de sus entrevistados. Entre sus hallazgos podemos destacar que quienes trabajan con los datos sociales consideran un problema central su calidad y precisión, que se presentan como desestructurados, confusos, metodológicamente defectuosos y volátiles. Se trata de un material en que datos relevantes se mezclan fácilmente con datos irrelevantes (un ejemplo sencillo son las búsquedas que incluyen conceptos comunes, como en el caso de marcas como Apple). Como respuesta a estas deficiencias, según Kennedy, la oferta de las compañías se concentra en resaltar el tamaño y velocidad de sus datos.

Es también muy valiosa la información que Kennedy obtiene respecto a los clientes de las AMP. Uno de sus descubrimientos más significativos tiene relación con la actitud de estos cuando se les advierte sobre las deficiencias de los datos y se les transparentan las limitaciones de las herramientas de minería de datos. Muchos de los entrevistados, según recoge la investigadora, sugirieron que a sus clientes no les importa que el resultado sean datos toscos, aproximativos o de baja resolución (*rough data*). Frecuentemente, pareciera que es suficiente con entregarle un número a los clientes, independiente de su exactitud o de si el *big social data* usada está limpia o es pertinente. Por ejemplo, en referencia a las expectativas de su clientela, un testimonio señala:

Si esos datos son exactos o no, es irrelevante. Solo quieren algunos números para poner en un PowerPoint y mostrárselos a su jefe. Si alguien pregunta “¿has puesto ojo en las redes sociales?”, puedes responder “sí, somos 36 esta semana”. Y esa es una solución muy atractiva. (Kennedy, 2016, p. 109)

Aparentemente, lo que se pide a los datos y a sus productores en estos casos no es evidencia, sino connivencia para sustentar cierta especulación (financiera, industrial, reputacional, política, etc.). Así, por una u otra vía, al menos de forma ocasional, estas empresas minan y filtran datos para ofrecerles a sus clientes exactamente lo que ellos quieren ver. De esta forma, grandes datos referencialmente imprecisos son la base de un tipo de conocimiento que puede estar siendo relevante para la toma de decisiones en ciertos contextos.

Los hallazgos de Kennedy muestran cómo, incluso hacia el final de esta larga cadena de valor de los datos, que se inicia con las actividades humanas en las plataformas y recorre dicho sistema, estos continúan siendo moldeados por decisiones e interpretaciones humanas, invisibilizadas por diversos mecanismos antes de presentarlos como “conocimiento accionable”. Tal como en el caso de los ratings, descrito por Meehan, en lo que respecta a los productos que comercializan las AMP se aprecia una sujeción a las reglas del mercado antes que a cualquier estándar de la investigación. El debate en torno a la validez referencial de esas mercancías es suprimido en medio de una avalancha de números y el deseo insaciable por ellos, la ilusión de objetividad ligada al sublime digital y el poder persuasivo de factores como la visualización de datos. La “ceguera” de los clientes ante las carencias técnicas y metodológicas que pudieran tener los datos y los instrumentos para interpretarlos, al igual que en el caso de la industria televisiva, bajo el prisma de las observaciones de Kennedy se muestra como algo constitutivo del sector y característico de las expectativas de sus clientes.

No se trata de preocupaciones triviales. Puede ser admisible un grado amplio de incertezas en disciplinas transparentemente sometidas a la crítica sectorial o social o en campos que no tienen un peso tan sustancial en la toma de decisiones políticas, que no afectan directamente a personas o que se someten al escrutinio público de aquellos a quienes afectan. Pero este no es el caso cuando se trata de las agencias de monitoreo de plataformas como la que confeccionó el informe Big Data para el Gobierno de Chile.

## **Consecuencias sociales de este modelo de producción**

Entre 2022 y 2023 la inteligencia artificial (IA) y los grandes modelos de lenguaje como los desarrollados por OpenAI (ChatGPT, DALL-E), Anthropic (Claude), Google (Bard, hoy Gemini) o Meta (LLaMa) se transformaron en uno de los principales temas sociales a nivel global. Una tecnología que hasta hace unos lustros habría sido recibida con una oleada de tecnooptimismo, originó en cambio un diluvio de pesadillas tecnodistópicas en las que la recolección masiva de datos personales opera a través de un “pragmatismo despiadado” en pro de alimentar la insaciable voracidad de las computadoras (Crawford, 2021), mientras estas se dedican a crear imágenes, videos y audios indistinguibles de las imágenes, videos y audios que estábamos acostumbrados a dar por “reales” (Satariano y Mozur, 2023).

Al mismo tiempo, la compra de Twitter por parte del multimillonario y activista de extrema derecha Elon Musk cambió radicalmente las condiciones de uso y el modelo comercial de la plataforma que durante más de 15 años se había consolidado como el terreno por antonomasia para la minería de datos. Una red social que por más de una década había sido criticada por poseer una arquitectura que posibilitaba la rápida distribución de mentiras y discursos de odio, y había implementado numerosas políticas para contener ese problema, de un día para otro tomó el camino opuesto: se deshizo de su equipo de moderación y comenzó a fomentar la creación de contenidos incendiarios, sin importar si estos son falsos o peligrosos (Goggin, 2023).

Pese a los enormes desafíos que estos dos procesos paralelos entrañan para la democracia, según un sinnúmero de observadores (Manheim y Kaplan, 2019; Elliott, 2022; French, 2022; Buhmann y Fieseler, 2023), los cuestionamientos y temores rara vez alcanzan al capitalismo contemporáneo ni al hecho de que las actividades cuestionables en esas dos esferas, Twitter y la IA, tengan que ver con la persecución de mayores ganancias. Ni OpenAI ni Google pondrán en pausa sus desarrollos en tecnología de IA, ni Musk parece estar dispuesto a dar el brazo a torcer

en su esfuerzo por convertir a Twitter, ahora llamado X, en una app que “lo haga todo” (Heath, 2023). Todo esto nos habla, por una parte, de la resiliencia de estas estructuras, invulnerables a la crítica, pese a los peligros e inconvenientes que pueden traer aparejadas. Por otro lado, cuando se tolera sin más que el espacio de lo público esté moldeado por distorsiones informacionales que son producto de prerrogativas comerciales, queda al descubierto la indulgencia que prima en nuestra sociedad hacia aquello que no es “referencialmente preciso”, hacia lo falaz.

Hemos ahondado en el recorrido de los datos dentro del sistema de las plataformas para dar cuenta de algunos de los entresijos de dichas arquitecturas, de los cuellos de botella en los que indudablemente se toman decisiones que lejos de permitir que “los números hablen por sí mismos”, acabando con la teoría y volviendo obsoleto al método científico, hacen que los números hablen por otros. En una conversación con la revista Science, un psicólogo dedicado a la minería de datos, con relación a la plataforma Twitter bajo el control de Musk, afirmaba que “esto revela el problema fundamental de si las sociedades democráticas pueden permitir que multimillonarios excéntricos controlen los espacios públicos” (Kupferschmidt, 2023).

La respuesta más directa a dicha interrogante es que las sociedades democráticas ya lo permiten y que el sistema de las plataformas ya está controlado por esa clase de multimillonarios, muchos de ellos menos llamativos que Musk, pero no menos excéntricos. Seres humanos que en los casos más extremos pueden atiborrar las plataformas de las que son dueños con sus propios *post* aprovechando la opacidad de los sistemas (Schiffer y Newton, 2023), pero que regularmente distribuyen su ideología de formas más sutiles a través de lo que Gilbert y Williams (2022) denominan “poder de plataforma” (*platform power*).

Los datos que circulan dentro y alrededor del sistema de las plataformas parecen enunciar y ocultar lo que otros, que permanecen como no responsabilizables, quieren que sea enunciado o callado. En el mundo metrificado que han creado, esos sujetos de poder se consideran exentos de metadatos. Así, una de las primeras medidas de Musk como dueño de X/Twitter fue censurar una cuenta que monitorizaba los metadatos de su jet, mientras que en Silicon Valley los ejecutivos de



las principales corporaciones tecnológicas envían a sus hijos a escuelas Waldorf sin acceso a Internet (Richtel, 2011).

Las decisiones inscritas en los algoritmos ayudan a dar forma a las teorías sobre lo relevante que las personas que usamos las plataformas interiorizamos. Lo que vemos en cada uno de los rincones del sistema de las plataformas son indicios dignos de una investigación agnotológica. La agnotología es el estudio de las causas de la ignorancia deliberada o culturalmente inducida para, por ejemplo, vender un producto o influir en una opinión, a través de datos científicos inexactos o engañosos.

La agnotología no describe la existencia de incógnitas sistémicas, sino más bien la producción de incertidumbre sistémica a través de un debilitamiento de los procedimientos intelectuales utilizados en la creación de conocimiento. Emplea los mismos procedimientos epistemológicos para crear conocimiento y comprensión, pero con el efecto opuesto. (Betancourt, 2016, p. 101)

Los procedimientos agnotológicos en el capitalismo contemporáneo, la producción de incertidumbre sistémica, dan como resultado la incapacidad de establecer la factualidad de cualquier contenido. Tanto la evidencia empírica como aquello que era socialmente consensuado como legítimo pierden su valor epistémico predominante. Y frente a ese escenario, una forma de reconquistar cierta seguridad ontológica es aferrarse a sustitutos de la verdad como los números y el *big data*, que entregan una ilusoria sensación de objetividad.

Cualquiera que se aproxime tímidamente al mundo del *big data* tal como es relatado en los medios de comunicación de masas, sabe que allá afuera, para la mayor parte del mundo, los datos son neutros, todopoderosos y sublimes. Desde luego que encontrará tímidas advertencias en el camino, notas al pie previniendo los matices. Pero es cuestión de ver, por ejemplo, las campañas políticas para entender el magnetismo del sublime digital y la fe en los datos. En su programa de gobierno para las elecciones presidenciales de 2021 en Chile, el candidato de la extrema derecha, José Antonio Kast, proponía “mejorar sustancialmente la respuesta” a las

emergencias “integrando tecnologías y data” (Kast, 2021), sin entregar más detalles. El programa de otro candidato, Sebastián Sichel, mencionaba cinco veces la palabra “data” o “big data”, en referencia a tecnologías que proponía implementar en diferentes ámbitos (Sichel, 2021). En tanto, el programa del candidato Gabriel Boric (Apruebo Dignidad, 2021), actual presidente, mencionaba también el *big data* en una ocasión, aunque prometiendo regularla en el contexto de la protección de la privacidad en línea.

La conceptualización predominante sobre los grandes datos como un artilugio técnico neutral, enajenado de sus condiciones culturales, económicas e históricas, tiene consecuencias sobre la manera en que son estudiados y empleados. Como hace notar Mosco (2015), implica una deriva hacia dos vías de análisis que están lejos de ser ideales: por un lado, que mediante el *big data* se traten problemas complejos acudiendo a sobresimplificaciones; por el otro, que se opte por ignorar y dejar de lado todos aquellos problemas que ameritan un tratamiento cuidadoso relacionado con la subjetividad. Bajo esta epistemología, quedarían fuera de nuestro alcance todos esos problemas para los que se desarrollaron herramientas “poco sofisticadas” como la etnografía o las entrevistas en profundidad. Incluso las encuestas y experimentos de campo, pero también la recolección de datos y testimonios y el uso de técnicas criminalísticas que caracteriza la labor, por ejemplo, de las policías.

Esto es lo que estaba ocurriendo cuando el Gobierno de Chile decidía en 2019 poner en el centro de sus decisiones políticas, comunicacionales y de seguridad nacional una mercancía analítica elaborada mediante la descarga y el análisis de datos (fundamentalmente) de Twitter: por una parte, sobresimplificaba asuntos tremendamente complejos, como son las movilizaciones sociales, las crisis políticas y las actividades delictivas; pero además optaba por ignorar deliberadamente todo análisis tradicional.

Pero no es solo la representación del *big data* como algo transparente y no distorsionado lo que constriñe las preguntas y las respuestas posibles. Todas las actividades que se dan bajo el capitalismo contemporáneo orientado a lo digital están encauzadas hacia un único objetivo: las ventas exitosas, la ganancia. La competencia en los mercados globales impulsa continuamente la

búsqueda de formas de producción cada vez más baratas (Pfeiffer, 2022). Desde este punto de vista, entre los estudios basados en datos que prometen lecturas sobre fenómenos sociales, aquellos amparados en la ideología del *big data* aparecen como ampliamente más competitivos (en términos de su costo e inmediatez) que las intervenciones que ofrecen disciplinas de las ciencias sociales que hasta hace relativamente poco dominaban tales mercados. También juega un papel en esto la percepción de que las encuestas por muestreo son un instrumento insuficiente en comparación con la minería de datos de las redes sociales, que permitirían salvar la distancia que hay entre una “simple” muestra y una población. Así lo explicitaba en 2010 el fundador de la empresa chilena Brandmetric:

Nosotros competimos con las encuestas, pero con la ventaja de que podemos entregar resultados mucho más rápido, al instante, de hecho y con muestras mucho más grande. Por ejemplo, la encuesta CEP se hace con 1.500 personas, en cambio nosotros revisamos unos cien mil usuarios de Twitter. (BioBioChile, 2011)

Otro factor muy relevante es la manera en que se presentan y visualizan los datos de las plataformas sociales cuando se transforman en mercancías analíticas: los y las investigadoras de las plataformas no suelen someter sus datos al tipo de análisis multivariante habituales en las ciencias sociales, sino que los muestran visualmente a través de ejercicios de análisis de agrupamiento (*cluster analysis*), lo que hace que los resultados sean accesibles a una amplia audiencia (Savage y Burrows, 2007). Y así como *big data* es un concepto de moda y un aparato discursivo, también lo son las visualizaciones que caracterizan la presentación de los datos que resultan de su análisis, parte central de las mercancías analíticas que comercian las AMP.

Tal como señalan Kennedy y Engebretsen, las visualizaciones no solo facilitan la comprensión de los datos, también la persuasión. Estas visualizaciones son soportes semióticos, estéticos y sociales, pero a la vez generan respuestas emocionales, lo que es consistente con las retóricas de los procesos agnotológicos y con la hegemonía de la mistificación de lo digital. Asimismo, no es difícil imaginar como estas visualizaciones se pueden transformar en vehículos para la producción de apofenia, fenómeno psicológico que consiste en percibir patrones donde en

realidad no existen (boyd y Crawford, 2012). Enormes cantidades de datos pueden ofrecer conexiones que irradian en cualquier dirección y las imágenes pueden permitir orientar tales conexiones hacia representaciones falaces.

### **El otro extremo de la cadena de valor de las plataformas**

Para explorar cómo las distorsiones y sesgos que hemos examinado en este capítulo podrían ocasionar algún impacto social, nuestra aproximación pasa por entender cómo operan efectivamente los agentes intermediarios entre los datos que producen las plataformas y un cliente final que espera obtener algo de tales datos. Sin embargo, este último es un eslabón de la cadena que no podemos dejar fuera del análisis. En esta sección evaluaremos brevemente qué podemos aducir sobre los consumidores finales de las mercancías de las AMP.

Como hemos mostrado al referirnos al caso del informe Big Data, en las semanas posteriores al estallido social de octubre de 2019, el Gobierno de Chile parecía estar ávido de encontrar un proveedor de monitoreos de redes sociales. Pero no cualquier monitoreo, sino uno que permitiera comprobar una tesis en particular. Y pese a las advertencias de los académicos a los que contactaron, no tardaron en dar justo con lo que estaban buscando encarnado en la empresa española Alto Data Analytics. No es difícil encontrar en esto un paralelo con los hallazgos de la ya citada Helen Kennedy.

En su análisis, Kennedy descubre una percepción que ella describe como generalizada entre sus informantes al referirse a la actitud que tienen los clientes de las AMP: autosometimiento a la ignorancia deliberada. Los consumidores de las mercancías analíticas serían indiferentes a la calidad de los datos y a los límites de las herramientas. Con frecuencia, según Kennedy, las y los trabajadores de estas compañías actúan en complicidad con sus clientes en la producción de resultados referencialmente imprecisos.

Lo que queda claro frente a estos antecedentes es que parece haber un problema en ambos extremos de la cadena de valor de los datos del sistema de las plataformas. Por un lado tenemos a

una industria que impone ideas sobre lo que es relevante y comercia ilusiones de certidumbre ontológica basadas en las premisas del sublime digital. Por el otro extremo, aparece una profusión de clientes dispuestos a suspender su incredulidad con el objetivo de obtener números para ilustrar PowerPoints o anunciar que no hay una crisis política, sino un complot chavista. En estos casos pareciera ser que lo que está a la venta son mercancías hechas a la medida que les permitirían a los clientes sustentar en datos y grandes datos cualquier idea preconcebida que tengan y sea útil para sus fines políticos o corporativos, dotando a sus voluntades de una apariencia de cientificidad. Bajo este prima, las distorsiones que hemos descrito como parte integral del sistema de las plataformas, así como las peculiaridades atestiguadas en la cultura empresarial de las AMP, parecen ser atributos comerciales y no defectos del sector.

Respecto a esto, es adecuado citar nuevamente a Eileen Meehan. Al referirse a cómo las cadenas de televisión emplean los ratings para sus fines corporativos, indica:

Las cadenas de televisión dan soporte a técnicas de medición inadecuadas precisamente debido a esas deficiencias — porque los ratings miden el tamaño de la audiencia y no el interés, el disfrute o el aprendizaje. Los ratings son instrumentos pobres, entonces, porque las cadenas de televisión consideran útiles esas insuficiencias. (Meehan, 1984, p. 219)

Charles Wright Mills, cuyo trabajo se desarrolló en parte importante bajo el alero del Departamento de Investigación Social Aplicada de la Universidad de Columbia, una de las cunas de la *Mass Communication Research*, pero que fue un tenaz crítico de la postura "apolítica" de la sociología estadounidense de su época, propuso que una de las características que distingue a la élite contemporánea de otras castas gobernantes de la historia es cómo ha conseguido concentrar en sus manos el control sobre la información y el rol que han jugado los medios de comunicación de masas en la mantención de su poder. Otro aspecto significativo de esta élite —lo que recoge los planteamientos de Marx y Engels en *La Ideología Alemana*— sería que sus miembros reconocen su posición social y desarrollan acciones colectivas dirigidas a la conservación de dicha posición y de dicho orden (Mills, 2013). En otras palabras, dicho bloque es uno que se

distingue a sí mismo como clase, y como clase no se engaña respecto a la necesidad de tener un papel activo en sostener su posición dominante sobre las otras clases sociales.

En un principio, al alterar la proporción entre hablantes y oyentes, así como las normas informales que regulan quién puede comunicar, cuándo, qué y durante cuánto tiempo, las plataformas sociales de Internet pusieron en entredicho la estabilidad de los viejos sistemas de comunicación de masas. Con esto, y siguiendo a Wright Mills, la élite se enfrentó a una notoria amenaza a su hegemonía producto de la emergencia de una comunidad de públicos y contrapúblicos. Esto produjo importantes flancos de incertidumbre sobre elementos previamente estables, como el control de la élite sobre la autonomía con la que las personas se forman opiniones y la relación entre la formación de opiniones y su realización en acción social, uno de cuyos ejemplos más llamativos y epocales fue el rol de Twitter en la Primavera Árabe (Tufekci, 2017).

¿De qué formas esta élite del poder se enfrenta a espacios que por su propia naturaleza ponen de manifiesto el orden establecido y entregan la posibilidad que se entablen litigios en torno a este, como son las plataformas sociales de Internet?

Una de las vías de restitución del orden mediático, particularmente para las élites locales, podría estar personificada en el sector de las AMP. Las empresas que ofrecen productos analíticos basados en los datos sociales masivos de las plataformas, cuya promesa implícita y muchas veces explícita es tener la facultad de proporcionar a partir de ellos “una comprensión inigualable del mundo” (Alto Analytics, s.f.), se ubican en una posición ideal para entablar acciones encaminadas a la reducción de incertezas hegemónicas y el restablecimiento del control en favor de una o varias élites.

El poder tiene interés en decidir qué cuenta como conocimiento, cómo se difunde, quién lo tiene y por qué se usa (Sadowski, 2020). Una industria que produce conocimientos y hechos a la medida de las necesidades de sus clientes, cuya racionalidad se sustenta en conceptualizaciones que son hegemónicas respecto a la superioridad del *big data* sobre cualquier otra manera de dar

sentido al mundo, aparece como estratégica para retornarle a determinadas élites el control sobre la información y otorgarles, también, seguridad ontológica.

### **Colonialismo de datos y continuidad colonial**

Para no tomar partido por ninguno de los capitalismoos que actualmente se encuentran en disputa como etiquetas para nuestro tiempo (capitalismo digital, informacional, de vigilancia, de plataformas, etc.), aquí hemos optado por hablar sencillamente de un capitalismo orientado hacia lo digital. En este sentido, coincidimos con la continuidad que plantea Fuchs (Fuchs, 2012a) y, de forma más decidida, Couldry y Mejias (2020), cuando señalan que en sus aspectos fundamentales el capitalismo contemporáneo sigue siendo el mismo que ha sido siempre. Pero más aún, coincidimos con estos últimos autores en la necesidad de poner foco en una continuidad aún más profunda para entender los fenómenos económicos que dan forma a lo que hemos llamado el sistema de las plataformas y sus consecuencias sociales. Esta perspectiva implica situar al capitalismo como un proceso histórico entramado con el colonialismo.

El colonialismo de datos se distingue del colonialismo histórico, en la concepción de Couldry y Mejias, en varios aspectos. Por un lado y principalmente, mientras el colonialismo que caracterizó las políticas imperiales europeas entre los siglos XVI y XX conformaba un tipo de sociedad y economía enfocado en la apropiación de los recursos naturales y del trabajo humano transmutado en mercancía, el colonialismo de datos transforma la experiencia humana digitalizada en un insumo para el capital. Por otra parte, respondiendo a ciertas críticas sobre su modelo, estos autores coinciden en que el colonialismo de datos, a diferencia del colonialismo histórico, no se caracteriza por ejecutar su poder mediante la violencia física. Couldry y Mejias (2021) señalan que esos métodos han sido reemplazados por otros, incluyendo formas de violencia simbólica. Dado que la “violencia física extrema” es un rasgo definitorio del colonialismo histórico según estos mismos autores (Couldry y Mejias, 2020), la capitulación respecto a este punto rompe parcialmente con la idea de continuidad del colonialismo que está en el fondo de la propuesta de los investigadores.

No obstante, la idea misma de que la violencia física no es parte del repertorio de manifestaciones de poder de, por ejemplo, la expansión de las plataformas, resulta disputable. Desde luego, la violencia simbólica y la violencia económica están presentes en estos movimientos, pero acaso, ¿hay una completa ausencia de violencia física cuando el imperativo por incorporarse a las relaciones de datos que configuran plataformas como LinkedIn o Trabajando.com puede significar la diferencia entre obtener o no un medio de sustento en una sociedad altamente desigual? Para casos como este, el concepto de “violencia simbólica” no alcanza a describir el nivel de violencia material que hay en juego. Qué decir de las innovaciones en materia de explotación que han emergido para simular la experiencia humana digitalizada y alimentar de datos al sistema de las plataformas, como las granjas de bots situadas en “economías en desarrollo”.

Asimismo, no toda la violencia colonial histórica implicaba cadenas y látigos como punto de partida: mucha de ella suponía someterse “voluntariamente” o incluso activamente, porque esa era la única forma de subsistencia disponible en los contextos coloniales. Como ha mostrado Jean-François Bayart, la agencia de los colonizados abarca un rango más amplio que la mera subyugación pasiva (Bayart, 2011).

Por todo ello, a nuestro entender es crucial no reducir todas las formas de violencia que resultan de estas distribuciones históricas de poder a sus dimensiones simbólicas y que la colonialidad del poder a través de los datos (Ricaurte, 2019) implica ensamblajes sociotécnicos complejos en que la violencia adquiere todos sus ropajes, incluyendo elementos emocionales, físicos y materiales, de autoexplotación, de género y clase, entre otros.

Desde esta perspectiva, nuestro apego a la idea de la continuidad colonial es radical. El colonialismo ha sido y sigue siendo importante porque no es solo un sistema genocida y de apropiación económica, sino que se caracteriza por un esfuerzo constante de imposición ideológica cuya base primera es la destrucción, y no la superación, de los sustratos culturales e ideológicos previamente asentados. Se sustenta en un epistemicidio, en el lenguaje de Boaventura de Sousa Santos, en “la destrucción de los sistemas de conocimiento, formas de conocimiento,



auto- percepción y modos de ser en el mundo” (de Sousa Santos, 2018). Es una forma social que disponibiliza los recursos para su extracción no solo a través de la fuerza bruta, sino que también a través de la comunicación bruta, reconfigurando los flujos de datos e imponiendo nuevas teorías sobre lo relevante.

Contra lo sostenido por Couldry y Mejias (2020), sostenemos que la apropiación colonial de datos no es un fenómeno absolutamente nuevo, emergente y atado al capitalismo contemporáneo de orientación digital, sino que tiene raíces que se hunden al menos hasta el inicio de la modernidad.

Ya señalamos, por ejemplo, la relevancia de los datos para el registro y control imperial español bajo el reinado de Felipe II, cuyo lema, según afirma Peters (2017), era *Quod non est in actis non est in mundo* (lo que no está en las actas, no está en el mundo). Pero una evidencia más directa de la centralidad que tenía el control de la información y los datos de la experiencia humana para la colonia española en el siglo XVI es la rápida expansión del sistema de las Audiencias. La Real Audiencia era la institución española en América encargado de administrar la justicia real, pero también de rellenar los vacíos normativos, asesorar a los virreyes, preocuparse del “bienestar de los indios”, determinar qué libros podían circular y, en última instancia, modelar el mundo mediante ordenanzas y mantener un registro de todo aquello en un lenguaje incomprensible para los conquistados y gran parte de los conquistadores. En el papel era un órgano judicial, pero en la práctica tenían un importante rol político (Sánchez Bella, Hera y Díaz Rementería, 1992). Su relevancia para la colonia queda patente en la velocidad de su instauración: la primera Real Audiencia en América se fundó en Santo Domingo en 1511; 56 años después ya había una en el último rincón del continente, Concepción.

Mediante esta institución, entre otros mecanismos (ya hemos hablado más arriba de la Casa de Contratación), y a fuerza de poseer los medios sociotécnicos para contar y dar cuenta de las vidas humanas bajo su dominio (es decir, imponer sobre los datos categorías sobre lo relevante o metadatos), el Imperio español no buscaba describir a los sujetos americanos, sino que construirlos a partir de registros sesgados y convenientes. La idea que había del indio para la

Corona estaba sustentada en unos datos ventajosamente elaborados para los fines particulares de la colonia: su explotación. El indio era conceptualizado en la ley como un “incapaz relativo” y estaba “liberado de la presunción del conocimiento de la ley” (Dougnac, 1994). Así, en lo que respecta a la Corona, lo que poseía de los indios no era sino a sus dobles de datos; pero las decisiones que tomaban sobre dichos dobles, segmentados y empaquetados adecuadamente, afectaban a las personas reales.

### **Consecuencias de este enfoque para nuestra aproximación**

La radicalidad con que acogemos aquí la idea de la continuidad colonial, que hoy deriva en un colonialismo de datos, tiene por consecuencia lógica un distanciamiento relativo respecto al contexto inmediato de nuestro objeto de estudio en pro de entender su posición dentro de un sistema de carácter global, el sistema de las plataformas, y sus relaciones sociales y económicas con un capitalismo contemporáneo enraizado en el escenario mayor del colonialismo histórico. Por ello, y atendiendo al carácter exploratorio y descriptivo de este estudio, hemos puesto énfasis en analizar en detalle y de forma crítica los conceptos centrales, originados en el llamado norte global, que han guiado los avances sociotécnicos en torno a las plataformas para articular sobre esa base un marco teórico con fundamentos de escala global.

Si bien, siguiendo a Milan y Treré (2019), este estudio busca eludir el universalismo digital, entendido como la tendencia de asimilar la diversidad de los desarrollos tecnológicos regionales a los principios de Silicon Valley, como advierten Mejias y Couldry (2021) no debe perderse de vista el nivel global. Por ello, entendemos que es indispensable un cierto grado de asimilación crítica: aunque la resistencia contra el colonialismo de datos es necesariamente local y debe desarrollarse desde las condiciones locales, no es posible ignorar las prácticas y discursos hegemónicos que dan forma a dicho colonialismo.

Con todo, el punto de vista que adopta esta investigación es uno que necesariamente se sitúa en un contexto regional específico. Su objeto de estudio son empresas con sede en Chile, que producen sus mercancías con trabajo local y cuyos clientes son fundamentalmente una élite

también local. El mercado en el que participan las AMP está marcado por modulaciones contingentes de carácter regional, como lo atestigua el ejemplo que hemos descrito en la introducción (el informe Big Data), cuyos detalles tienen un componente colonial (no es trivial que la AMP encargada de dicho informe sea de origen español). Así, las consecuencias de la meridionalidad de estos fenómenos son relevantes tanto para nuestro encuadre teórico como para el análisis que desarrollaremos.

“La estructura social es global, pero en su mayor parte la experiencia humana es local, tanto en sentido territorial como cultural”, indica Castells (2013), aludiendo a una profunda fragmentación producto de “la doble lógica de la inclusión o exclusión en las redes globales que estructuran la producción, el consumo, la comunicación y el poder”, fragmentación que sería una característica estructural de la sociedad contemporánea. Desde nuestra óptica, es en los intersticios de esta fragmentación (y no en cualquiera de sus fragmentos) donde se debe ubicar la mirada decolonial y situada en el ámbito del capitalismo orientado hacia lo digital y el poder de las plataformas.

Por tanto, cualquier exploración crítica sobre estos fenómenos debe estar enmarcada globalmente y, al mismo tiempo, cuestionar los constructos heredados a la luz de los descubrimientos vernaculares y desarrollar categorías propias que puedan dar cuenta de la fragmentación estructural que ofrece el sistema. Esto, a su vez, impone un compromiso con la producción de datos empíricos locales que posibiliten comenzar a desentrañar la amplitud y diversidad de las apropiaciones particulares e idiosincráticas de la tecnología en los contextos particulares.

Esta mirada “desde los intersticios” permitirá evaluar si la industria del monitoreo de plataformas sociales en Chile contribuye o no a sostener y profundizar no solo el desequilibrio de poder entre aquellos que controlan el sistema de las plataformas (a nivel global) y los usuarios ordinarios (Andrejevic, 2014), sino que a entender la posición dentro de un sistema global de aquellos que controlan los datos a un nivel local (el sector AMP y sus clientes), qué clase particular de control tienen sobre los datos, cómo participan de un contexto social caracterizado

por desigualdades de todo orden (geográficas, económicas, raciales, sexogenéricas, etc.) y qué consecuencias tienen estas configuraciones para la sociedad. Dichas observaciones pueden otorgarnos tanto nuevas herramientas analíticas como sugerirnos vías de resistencia.

## OBJETIVOS DE LA INVESTIGACIÓN

Hasta aquí, hemos procurado delinear el contexto histórico, social y técnico que envuelve el fenómeno de las AMP, así como aproximarnos a las cualidades conocidas, de acuerdo a la literatura disponible y a un caso socialmente relevante, de las empresas del sector, sus operaciones y sus mercancías. No obstante, hay un enorme espacio que nos resulta desconocido, especialmente desde la perspectiva teórica que hemos adoptado, debido a que hasta ahora no se ha desarrollado un análisis específicamente dirigido a describir el sector. La ausencia de una conceptualización socializada explícita del sector de las AMP no solo tiene un efecto evidente en la opacidad de su rol social y en la comprensión y significación que diversos actores tienen de esta clase de aproximaciones relacionadas con el *big data* y las tecnologías de la información, sino que impide que los investigadores e investigadoras más próximos a los problemas de los medios digitales contemporáneos se planteen las preguntas de investigación adecuadas que el ámbito amerita.

La motivación de este estudio es precisamente esta: ofrecer una primera aproximación exploratorio-descriptiva al sector de las agencias de monitoreo de plataformas sociales en Chile, que permita abordar posteriormente a dicho ámbito como un problema de estudio en todas sus dimensiones. Para reductos de la economía contemporánea que se ven tan constreñidos en su evolución por factores de orden global, por el comportamiento impredecible y en constante ebullición de algunas de las corporaciones transnacionales más grandes, ya no del mundo, sino de la historia de la humanidad, y que forman parte de un sistema caracterizado por reajustes y transfiguraciones constantes, el esfuerzo descriptivo y topológico de una investigación como esta puede aparecer como algo trivial, incluso inconducente. Lo que describamos ahora como elementos centrales de una industria o de un modelo de negocio, puede dar un giro en 180° en pocos meses. Las mutaciones y los cambios de foco de las empresas que se han desarrollado como rémoras en torno a los gigantes de Internet, y que surgen y desaparecen incesantemente, dan buena cuenta de estos procesos.

Desde que iniciamos esta investigación, hemos sido testigos de algunos de estos procesos, de cómo el foco de algunas empresas que habíamos identificado como perteneciente al sector AMP fue cambiando (de acuerdo a las dramáticas vicisitudes del mercado global y nacional desde 2019) o de cómo otras empresas fueron incorporando un lenguaje más apropiado para las modas tecnológicas que iban surgiendo (la preeminencia de la inteligencia artificial en todas las conversaciones ha sido significativa, especialmente desde 2023). En ocasiones, la oferta de aproximaciones en “tiempo real” ha transmutado a propuestas más humildes, acorde a los cambios que ha sufrido el sistema de las plataformas en los últimos años y meses, especialmente en lo que respecta a X/Twitter.

No obstante, aunque las AMP quizá no se encuentran en su momento más descollante debido a estos y quizá otros factores que de momento no podemos vislumbrar, podemos apreciar que el sector aún persiste y, dentro del fenómeno de los negocios asociados al sistema de las plataformas que es global, sigue siendo una de las pocas áreas comerciales que manifiestan una conducta fundamentalmente local con ramificaciones relevantes para la élite de poder local. Se trata además de un sector relativamente longevo: como hemos visto en nuestros antecedentes, podemos identificar su presencia en el norte global al menos desde la década de 2010 y, aunque sufra nuevas y sustanciales transformaciones a futuro, es posible que muchos de los elementos que lo envuelven sigan caracterizando a sus posteriores etapas de desarrollo. Por ello creemos que es fundamental dar cuenta de sus características de una forma más precisa que lo que se ha hecho hasta ahora.

Asimismo, como hemos mostrado en las secciones previas, existen múltiples aspectos problemáticos asociados al registro y mercantilización de los datos provenientes de las plataformas publicitarias. La indeterminada incidencia que estas mercancías analíticas podrían estar teniendo en la sociedad chilena exhorta a aproximarnos a este fenómeno, entender a las empresas dedicadas al monitoreo de las plataformas sociales de Internet e interrogarlas sobre las promesas explícitas e implícitas que caracterizan a sus productos.

Para abordar estas problemáticas, la presente investigación se plantea los objetivos detallados a continuación.

## **Objetivo general**

Caracterizar el sector de las agencias de monitoreo de plataformas sociales (AMP) en Chile a través de un análisis a su oferta de productos, servicios y mercancías.

## **Objetivos específicos**

- 1) Esbozar un mapa preliminar del sector de las agencias de monitoreo de plataformas (AMP) en Chile.
- 2) Explorar y evaluar desde una perspectiva crítica la oferta de mercancías analíticas que caracteriza a las agencias de monitoreo de plataformas (AMP) en Chile.
- 3) Determinar los atributos más relevantes de las empresas que forman parte del sector de las agencias de monitoreo de plataformas (AMP) en Chile.





## DISEÑO METODOLÓGICO DE LA INVESTIGACIÓN

Para avanzar hacia una comprensión en profundidad del sector de las AMP en Chile, se requiere en primer lugar proveernos de una conceptualización más acabada de este, con base empírica y con un pie en sus particularidades locales. Este será el norte del diseño metodológico que expondremos en el presente capítulo y que persigue constituir un aporte en el contexto de los estudios en comunicación política en nuestra región.

### **Enfoque de la investigación**

Para alcanzar los objetivos de esta investigación, esta propuesta metodológica considera la realización de un estudio cualitativo utilizando un enfoque exploratorio-descriptivo.

Nuestra aproximación es exploratoria, en tanto nos interesa dar cuenta de un fenómeno que es poco conocido y al que hasta aquí solo hemos podido acercarnos teóricamente basándonos en experiencias y conocimientos generados en las áreas disciplinares abocadas a esta familia de problemáticas en países tecnológicamente desarrollados, en contextos análogos, pero diferentes al nuestro. En lo que respecta a Latinoamérica y en particular a Chile, este sector no ha sido objeto de indagatorias metódicas desde la Comunicación Política ni desde las Ciencias Sociales en general.

Asimismo, nuestra aproximación es descriptiva en tanto propende a hacer un acopio de documentación que posibilite dotar a ulteriores análisis de una sustancia empírica mayor a la que hasta el momento existe sobre este objeto de estudio. Dado que nuestro objetivo general es realizar una caracterización de este sector con foco en la producción y distribución de sus mercancías, es esencial contar con datos que nos permitan establecer con mayor claridad y profundidad tanto su oferta de productos y servicios, como entender las motivaciones o expectativas comerciales o simbólicas de sus clientes.

Desde estas premisas, este estudio nos permitirá poner en relación el fenómeno de las AMP en Chile con los antecedentes y las inferencias que hemos abordado en el marco teórico, lo que a su vez posibilitará en futuras investigaciones generar hipótesis tendientes a avanzar en el conocimiento de este sector en Chile y su papel en la comunicación política y social.

## **Encuadre y procedimientos**

Como hemos hecho hincapié en los antecedentes de este estudio, entre las problemáticas emergentes del sistema de las plataformas destacan las que derivan de la opacidad de sus procedimientos técnicos y de gobernanza y de las “vías implícitas” que lo caracterizan como medio comunicacional (Armano, Briziarelli, Flores y Risi, 2022), a través de las cuales un determinado bloque ideológico ha expandido su hegemonía. De acuerdo a diversos autores y autoras, estas vías implícitas fomentan el deseo por los números y sostienen lo que se ha llamado sublime digital y algorítmico, en que los datos se constituyen como sustitutos de la verdad y vehículos de persuasión que otorgan a ciertos sujetos y grupos de poder la capacidad de imponer al resto sus teorías sobre lo relevante. En última instancia, estos aspectos marcan el carácter de las relaciones de datos que dan validez a la extracción de datos y su mercantilización y entregan estabilidad al colonialismo de datos (Couldry y Mejias, 2020).

En el contexto del sistema de las plataformas observamos al menos dos procesos de mercantilización relacionados con estas dinámicas: por un lado, se despliega una industria que reproduce y reformula algunas de las características de los medios de comunicación de masas del siglo XX, convirtiendo la atención de las audiencias digitales en mercancías; por el otro, a través de la minería de datos y la labor del sector AMP, podríamos estar en ante procedimientos que devuelven a las élites locales la seguridad ontológica amenazada por el desorden comunicacional provocado por la irrupción de las propias plataformas sociales.

Desde nuestra perspectiva teórica, estudiar el sector de las AMP implica explorar las cualidades de sus mercancías, lo que a su vez supone entender sus aspectos comunicacionales y persuasivos. A partir de nuestro ejemplo paradigmático, el informe Big Data del gobierno de Sebastián Piñera, y la literatura disponible, establecimos que el acceso a las mercancías analíticas que elaboran las AMP tiene rasgos de ser en parte importante, sino exclusivamente, un consumo de carácter simbólico, que otorga a su poseedor el poder de la razón; una razón mediada por los números y datos, elaborada a la medida del consumidor y que podría constituir una pieza central en el restablecimiento del poder de ciertas élites locales.

Para explorar esta posibilidad y atendiendo al carácter de las mercancías analíticas en tanto construcciones discursivas, hemos optado por encuadrar esta parte de la investigación desde la perspectiva del análisis crítico del discurso (ACD). Esta perspectiva nos permitirá comprender en qué consisten y cómo se da forma a los discursos que dan cuenta y que contienen las mercancías de las AMP y profundizar en el entendimiento de su posible utilidad instrumental.

### **Análisis crítico del discurso**

El análisis del discurso hace referencia a la lectura detallada del uso “real” o “efectivo” del lenguaje con el propósito de diseccionar sus estructuras y establecer sus significados (Waring, 2017). El análisis crítico del discurso (ACD), por su parte, se enmarca dentro de la teoría crítica de la sociedad y, mediante una lectura atenta a los usos efectivos del lenguaje, estudia las formas en que el poder y la desigualdad se manifiestan en el texto y la conversación en contextos sociales y políticos (van Dijk, 2001). Para llevar a cabo esta agenda, el ACD no solo describe las estructuras de los discursos, sino que procura explicarlas en términos de sus propiedades de interacción y su lugar en las estructuras sociales. El ACD supone una relación dialéctica entre los eventos discursivos particulares y las situaciones, instituciones y estructuras sociales en que se enmarcan (Wodak, 2014). Los eventos discursivos, desde esta perspectiva, son moldeados pero también dan forma a dichos fenómenos sociales.

El discurso encarna la ideología y puede ayudar a producir, reproducir y resistir las relaciones de poder desiguales a través de las maneras en que representa las cosas y posiciona a las personas. Su análisis crítico, que involucra la búsqueda de las relaciones entre la producción de significados mediante signos, la estructuración de las prácticas sociales y las estrategias de los agentes sociales que se manifiestan en los textos, requiere de una investigación transdisciplinaria en que se seleccionan métodos particulares de acuerdo a cómo se construye el objeto de estudio (Fairclough, 2013). El foco del ACD son los problemas sociales y no los paradigmas académicos, por ello los procedimientos y técnicas para su aplicación deben seleccionarse en función de su pertinencia para el propósito de cada investigación (van Dijk, 2003).

En este caso, como marco analítico general para el tratamiento de nuestros datos y su análisis, hemos empleado y adaptado la propuesta desarrollada por Dianna Mullet (2018), en que la investigadora sistematiza las características y procesos comunes dentro de los enfoques de ACD. Dicho marco de trabajo describe un conjunto de pasos no rígidos y fomenta la selección de los métodos que mejor se ajusten a los problemas explorados en cada investigación.

Además, hemos optado por un enfoque basado en corpus. Un corpus es un gran conjunto de datos de lenguaje natural, de miles o incluso millones de palabras, almacenado digitalmente para ser procesados de diversas formas por programas computacionales con el fin de descubrir patrones lingüísticos que contribuyan a entender cómo se utiliza el lenguaje en la construcción de discursos (Baker, 2006). Los y las investigadores que desarrollan ACD usan los corpus para intentar vincular patrones recurrentes en los textos con cualidades sociolingüísticas del entorno y viceversa (Flowerdew, 2013).

Una de las ventajas de efectuar un análisis de discurso con foco en un corpus es reducir en parte los sesgos cognitivos que pueden exacerbarse frente a la selección de ciertos textos. Un corpus consistente hace menos fácil ser selectivo respecto a los discursos que sustentan nuestros sesgos. Es además una forma útil para aproximarnos a un discurso debido al efecto incremental: un solo texto o incluso una sola frase puede sugerir la existencia de un discurso, pero es difícil a partir de ahí decir si dicho discurso es típico o no en un contexto sociocultural. Mediante la

recolección y análisis de numerosos ejemplos de construcciones discursivas, en cambio, podemos ser capaces de apreciar un efecto acumulativo y la aparición de patrones repetitivos que nos informen de lo habitual que es un discurso particular (Baker, 2006).

### **Selección del discurso de interés**

Los procesos que hemos descrito a lo largo de este trabajo, ponen de manifiesto la necesidad de entender los discursos mediante los cuales las AMP en Chile ofertan sus mercancías. Estos discursos modelan y son modelados por las relaciones sociales y políticas en que se enmarca el sector, por lo tanto, constituyen una vía adecuada para explorar y describir a estas empresas y entender si las perspectivas ideológicas que hemos caracterizado en nuestro marco teórico se reproducen en un contexto nacional como el chileno y en qué grado varían.

Al poner el foco en estos discursos, las siguientes preguntas guiarán el análisis:

**Pregunta de investigación 1:** ¿Qué componentes ideológicos caracterizan a la oferta de mercancías o servicios que ofrece el sector de las AMP en Chile?

**Pregunta de investigación 2:** ¿Cómo se hace cargo el sector de las AMP chilenas de las diversas problemáticas relacionadas con el sistema de las plataformas y los procedimientos de la minería de datos?

### **Localización y preparación de las fuentes de los datos**

Una vía de acceso comunicacional relevante para entender los eventos discursivos que caracterizan y diferencian a un sector comercial es su presencia en Internet, en sus sitios web, perfiles en redes sociales, podcast, *reels* promocionales y otros instrumentos. La *World Wide Web* ofrece la potencialidad de entregar información a una audiencia más amplia, distribuida a nivel global y no solo local o nacional, en un marco de tiempo acotado, de forma regular y a un menor costo que otros medios. Los sitios web corporativos, en particular, son superiores a los medios de

comunicación tradicionales en varios aspectos (Pollach, 2005). Permiten transmitir una cantidad prácticamente ilimitada de información a todas las audiencias potenciales, desde empleados hasta clientes, inversores y periodistas. Se trata de un medio más accesible que los tradicionales *brochures* e informes anuales de las compañías, pero que a la vez debe interpelar a una variedad más amplia de *stakeholders* (Williams y Pei, 1999), incluyendo clientes potenciales que pueden incluso no haber sido identificados aún. Esto supone retos a la comunicación para las empresas, que deben producir mensajes menos técnicos, que deben suponer menos conocimientos por parte de la audiencia, pero que aún así deben ser persuasivos.

Los sitios web corporativos son esencialmente informativos y persuasivos y, a diferencia de las comunicaciones de las empresas en plataformas sociales, la información que entregan no es fugaz ni contingente, sino que busca ofrecer una representación suficientemente clara, seductora y permanente de cara a los clientes. En sus sitios web, además, las compañías generalmente no se ven en la necesidad de lidiar con los imperativos particulares de los algoritmos de las plataformas sociales y las distorsiones que pueden provocar.

Así, los mensajes que las empresas envían a través de los sitios no son filtrados por intermediarios, como ocurre con otros medios, incluyendo las plataformas sociales: en los sitios web, el control está exclusivamente en manos de las propias empresas. Pero a su vez, dado que las audiencias de los sitios web tienen mucho más control sobre lo que quieren ver que en los medios tradicionales o que en las plataformas sociales, se trata de un medio que por necesidad propende a la eficiencia comunicacional.

Estas cualidades permiten inferir que un buen sitio web corporativo proveerá al receptor de una caracterización eficaz de cómo una empresa pretende mostrarse, a través de un medio mayormente unidireccional, ante una audiencia que no puede incidir de forma equilibrada en la modelación de dichos mensajes.

Para el abordaje documental y la producción de datos para esta investigación hemos optado por enfocarnos a los sitios web corporativos de las AMP chilenas. Dado que la

mercantilización de un bien está asociada a la mercantilización del lenguaje usado para referirse a este (Herrero, 2016), los discursos promocionales de los sitios web tienen la finalidad de vender una representación simbólica de un producto y de imponer representaciones mentales de los valores en juego de la forma más eficaz posible, por lo que los textos presentes en dichos espacio aparecen, de manera preliminar, como particularmente representativos de los discursos que singularizan al sector de las AMP en Chile.

Para acceder a estos textos, en primer lugar se realizó un mapeo de las AMP chilena con presencia en Internet. Se sistematizó una amplia muestra de organizaciones candidatas (34), sobre las que se hizo una limpieza para descartar aquellas que no correspondieran a nuestra definición de AMP o cuya adscripción no pudiera establecerse con toda claridad. En la siguiente sección se detallan los criterios usados para esto.

## **Criterios para delimitar el sector AMP**

Las agencias de monitoreo de plataformas sociales, tal como las hemos caracterizado en los capítulos precedentes, son empresas intermediarias entre las plataformas sociales y terceros clientes de pago, que realizan a solicitud de estos, minería de datos sobre *social big data*, que acceden a dichos datos principalmente —aunque no exclusivamente— a través de las interfaces de programación de aplicaciones (API) de las plataformas, y que mediante la búsqueda, recopilación, procesamiento, análisis, interpretación y almacenamiento sistemático de contenidos generados por usuarios y usuarias (UGC) producen un tipo de producto infocomunicacional que aquí hemos denominado mercancía analítica, frecuentemente consistente en informes y presentaciones en diversos soportes con un fuerte énfasis en la visualización de datos.

Como se puede apreciar, se trata de una industria bastante específica y discreta, pero no necesariamente fácil de reconocer dentro de un vasto panorama de consultoras y agencias comunicacionales e informacionales que ofrecen una amplia gama de servicios a clientes semejantes, usando descriptores similares (p. ej. “*stakeholder*”, “*clipping*”, “gestión de crisis”,

“redes sociales”, “monitoreo”, etc.), pero que, sin embargo, no desarrollan minería de datos de plataformas sino que otro tipo de aproximaciones (como análisis de prensa, etnografía digital, lobby, asesorías, *social media management*, *community management*, etc.). Asimismo, en el ámbito de la ingeniería y la informática, en la actualidad hay muchas empresas que comercializan productos basados en datos o *big data* para diferentes sectores, como instituciones de salud o diversos procesos industriales, y pese a utilizar un lenguaje corporativo que es, en muchos aspectos, similar al de las AMP, sus operaciones, prácticas y su negocio principal no tiene relación con estas.

Por todo esto, tras familiarizarnos con la literatura disponible (incluyendo trabajos académicos y técnicos, sitios web, prensa y redes sociales) y habiendo desarrollado una aproximación preliminar a un amplio rango de organizaciones, para determinar la adscripción de una empresa particular al sector de las AMP hemos establecido los siguientes requisitos:

1. Ser una empresa que comercializa mercancías y servicios basados en minería de datos. Este concepto, pese a ser útil como descriptor más o menos preciso del campo según la literatura consultada, no está siempre presente (de hecho, no es un concepto común en el sector, por lo que el tercer criterio será esencial).
2. Ser una empresa que comercializa sus productos o servicios de forma pública, es decir, estar abierta a todo tipo de clientes (a los que puedan pagarlo y aprecien su utilidad), en oposición a estar orientada a una única industria (por ejemplo, solo a análisis de datos para la industria del agro) o sector (por ejemplo, solo ofrecer servicios al gobierno).
3. Utilizar en sus comunicaciones al menos uno de los ocho conceptos generales o sus variantes para describir sus servicios, productos o soluciones (Tabla 1).



*Tabla 1. Conceptos generales y variantes habituales característicos del sector AMP.  
Elaboración propia.*

<b>Concepto general</b>	<b>Variantes habituales</b>
Análisis de redes sociales	Análisis de plataformas, Análisis social media, Social media analysis
Analítica en redes sociales	Analítica de plataformas, Social media analytics
Escucha social	Escucha de redes sociales, Social listening
Inteligencia social	Social media intelligence
Informes de redes sociales	Social media reporting
Métricas de redes sociales	Social media measurement
Monitoreo de redes sociales	Monitoreo social media, Monitoreo de plataformas, Social media monitoring
Minería de opiniones	Opinion mining, Análisis de sentimiento, Sentiment analysis

Pese a que trataremos con compañías chilenas o que tienen operaciones en Chile, en el criterio 3 se incluyen términos en inglés y español, ya que hemos observado que se trata de una industria donde son frecuentes los anglicismos. Todos estos calificativos refieren comúnmente a enfoques, variedades o nombres comerciales para procedimientos de minería de datos de plataformas sociales.

También se incluyeron los términos “minería de opiniones” y “análisis de sentimiento”, frecuentemente considerados equivalentes. Estos conceptos refieren a un procedimiento particular y más acotado dentro de la minería de datos que persigue dar sentido a la subjetividad expresada tras un texto mediante diferentes modelos de procesamiento de lenguaje natural (PLN) para identificar su “tono emocional” (positivo, negativo o neutral) y clasificar las opiniones en línea (Barney, s. f.).

Es importante señalar que los términos citados no agotan las posibilidades neológicas del sector, pero son los más habituales de acuerdo al mapeo que hemos llevado a cabo entre 2022 y 2023 (el dinamismo de Internet puede hacer que esto cambie con el tiempo). Todos estos conceptos suelen ser definidos de forma muy similar en distintas fuentes internacionales no académicas relacionadas con el mercado de la minería de datos de redes sociales. Algunos ejemplos de estas definiciones, extraídos de sitios web de empresas internacionales que ofrecen software de gestión de redes sociales, son los siguientes:

El análisis de las redes sociales es el proceso de recopilar los datos más valiosos de sus canales de redes sociales y sacar conclusiones procesables. (Pinkowska, 2023)

La analítica en redes sociales es el proceso de recoger y analizar la información obtenida de tus canales sociales con el objetivo de realizar acciones en base a los resultados. (Metricool, 2021)

La escucha social es la práctica de monitorear los canales de redes sociales para menciones de su marca, marcas de competidores y palabras clave relacionadas. (...) Esto le dará información valiosa. (Newberry y Macready, 2022)

La inteligencia de las redes sociales es el proceso de recopilación y análisis de datos de plataformas de redes sociales para obtener ideas y tomar decisiones informadas. (Spiryn, 2023)

El monitoreo de las redes sociales es el proceso de identificación y determinación de lo que se dice sobre una marca, individuo o producto a través de diferentes canales sociales y en línea. (BigCommerce, s. f.)

La métrica de las redes sociales (social media measurement) es el proceso de analizar y evaluar de cerca los resultados de sus actividades en las redes sociales. (Gharai, 2023)

Desde luego, la profusión de términos tiene por objetivo, además de permitir distinguirse de la competencia, establecer algunas diferencias entre las aproximaciones. No obstante, estas diferencias son poco sustantivas y detrás de todos estos procedimientos se encuentran prácticamente las mismas herramientas, técnicas, fundamentos, ideologías y profesionales.

Por otra parte, en el mercado existen grandes consultoras que ofrecen diversas asesorías y monitoreos (sobre el contexto político, social, legislativo, periodístico, etc.), que dependiendo de las necesidades particulares de sus clientes, en ocasiones pueden adquirir, de terceras empresas, mercancías analíticas basadas en datos de plataformas sociales. Estas compañías pueden mostrar en sus sitios corporativos claves generales sobre lo que ofertan, incluyendo aquello que tercerizan hacia el sector AMP (“Realización de estudios ad-hoc”; “Informes de contexto político a nivel nacional”, “Monitoreo de medios”<sup>1</sup>). No obstante, en estos casos, el no explicitar de forma clara su relación con las redes sociales, *social media* o plataformas sociales de Internet, también se consideró como criterio de exclusión de nuestro canon.

Con base en estas pautas, pudimos identificar 18 empresas para el sector AMP en Chile con presencia en Internet a través de sitios corporativos (Tabla 2). Sobre los discursos (textos) plasmados en estos sitios web se desarrollaron los procedimientos de producción de un corpus que detallaremos en la siguiente sección.

---

<sup>1</sup> Estos son algunos de los conceptos que aparecen en el sitio web de la consultora Tironi y Asociados, <https://tironi.cl/>.

Tabla 2. Mapeo preliminar del sector AMP en Chile.

<b>Empresa</b>	<b>Sitio web</b>
Analitic	www.analitic.cl
Artool	www.artool.cl
Azerta	www.azerta.cl
Brandmetric	www.brandmetric.com
Conecta Media	conectamedia.cl
CyberCenter	www.cybercenter.cl
Edgy	www.edgy.cl
Extend	www.extend.cl
Feedback	www.feedback.cl
Global News Group	www.globalnewsgroup.com
Imaginación	www.imaginacion.cl
Jelly	www.jelly.cl
LFI	www.lfi.la
LitoralPress	www.litoralpress.cl/sitio/mediossociales.html
Media Interactive	www.mediainteractive.cl
Nexos	www.nexos.cl
Rompecabeza	www.rompecabeza.cl
Soul	www2.souldigital.cl

## **Procedimientos para la producción del corpus de análisis**

Una vez definido el canon de sitios web corporativos para el análisis, la producción del corpus textual se desarrolló de acuerdo a las siguientes consideraciones.

- 1) **Descarga del contenido textual de los sitios web seleccionados.** Los 18 sitios listados se presentaban en una variedad de arquitecturas y plataformas, desde algunos muy básicos montados sobre Wordpress con la intención aparente de ser meras páginas de contacto, a otros desarrollos algo más complejos en términos de *front-end* (capa de visualización), con diversas combinaciones de HTML, CSS y JavaScript. Considerando esta diversidad, para asegurar la recopilación de la mayor parte de la información disponible, se descartó la opción de recopilar los datos mediante una herramienta automatizada de *web-scraping*<sup>2</sup>. Además, ninguno de los sitio ofrecía una Web API (interfaz de programación de aplicaciones para la web), por lo que esa tampoco era una vía posible para acceder a los datos.

En cambio, dado que el volumen de sitios lo permitía, se realizó la captura “manual” de todo el texto de todas las páginas dentro de cada sitio utilizando el software de gestión de documentos DEVONthink 3. Cada página o módulo se guardó en formato de texto plano o pdf, de acuerdo a la opción que nos permitiera rescatar la mayor cantidad de información de acuerdo a la arquitectura de cada sitio web.

- 2) **Selección, limpieza y consolidación del corpus.** De los sitios web, lo que se incluyó en la recopilación fueron todas las páginas que estuvieran vinculadas a su *landing page* o su página principal, salvo por los vínculos que dirigieran a blogs o muros de noticias de las empresas (que suelen tocar temas de variada naturaleza, muchas veces coyunturales y fuera de los aspectos propios del negocio, y que requerirían un análisis particularizado), webinars, podcast y perfiles en plataformas sociales (cuyos contenidos también

---

<sup>2</sup> Web-scraping es el proceso de usar bots para extraer contenido y datos de un sitio web a partir del análisis del código HTML subyacente, para registrar así los datos almacenados en una base de datos.

merecerían otras aproximaciones)<sup>3</sup>. Asimismo, se excluyeron del análisis sitios conectados al *landing* pero que estuvieran situados en un dominio diferente (habitualmente correspondiente a empresas asociadas o a oficinas internacionales de la empresa en observación), salvo que fuera clara su vinculación con los aspectos centrales de este estudio (como en el caso de empresas que usan un dominio distinto para cada uno de sus departamentos, pese a formar en conjunto una entidad única).

Además, se consideró pertinente incluir en el corpus de texto los videos o *reels* promocionales que las empresas situaran en sus páginas principales aludiendo a su oferta de productos y servicios. Cuando los sitios ofrecían este tipo de material, se capturó el audio de dichos videos y se transcribió utilizando el programa Whisper Transcription para MacOS, que funciona sobre Whisper, modelo de aprendizaje profundo entrenado por la empresa OpenAI (Radford et. al., 2023). Las transcripciones resultantes fueron revisadas para corroborar su fidelidad y corregir los posibles errores.

El registro de los sitios corporativos se llevó a cabo entre noviembre de 2023 y enero de 2024.

- 3) **Corpus final.** El resultado fue un corpus de 87.781 palabras y 616.781 caracteres. El número de páginas capturadas en cada sitio corporativo fue de 1 a 20, con un promedio de 8 páginas por sitio, lo que dependía de la envergadura y complejidad de estos. Este aspecto también resulta un indicador de la cantidad y variedad de servicios (relacionados

---

<sup>3</sup> Una de las expresiones más repetidas en el contexto de la minería de datos comercial es que para triunfar y crecer, una empresa necesita poder adquirir, retener, satisfacer e involucrar a sus clientes de manera efectiva a través de las plataformas sociales. Curiosamente, las AMP chilenas tienen una presencia modesta en las plataformas que se dedican a investigar. Por ejemplo, @bmetric, el perfil de Brandmetric en X/Twitter, tiene solo 2157 seguidores y durante 2021 solo publicó 11 veces (5 solo en febrero). @Analytic, mientras tanto, tenía a noviembre de 2023 solo 3098 seguidores en X/Twitter y 435 seguidores en Instagram. No son casos aislados.

con minería de datos, pero también con otros ámbitos comunicacionales, como marketing, diseño, clipping de prensa, relaciones públicas y lobbying, entre otros) que cada empresa ofrece.

El corpus de análisis quedó constituido según se detalla en la siguiente tabla, donde también se indica el número de páginas y el volumen de texto capturado en cada caso.

*Tabla 3. Detalles del corpus de análisis.*

<b>Sitio web</b>	<b>N.º de páginas</b>	<b>N.º de reels</b>	<b>N.º de palabras</b>	<b>N.º de caracteres</b>
Analitic	13	0	10021	66917
Artool	4	1	3321	23408
Azerta	4	0	9522	65813
Brandmetric	7	0	2883	19935
Conecta Media	10	1	3385	23039
CyberCenter	10	0	1389	10097
Edgy	4	0	1794	12105
Extend	11	0	5086	34979
Feedback	13	0	4956	36791
Global News Group	3	0	1528	10454
Imaginación	8	0	7772	53681
Jelly	3	1	770	5450
LFI	2	0	991	6911
LitoralPress	20	0	11905	90369
MediaInteractive	1	0	1237	8455
Nexos	4	0	6282	45994
Rompecabeza	20	0	14519	99652
Soul	1	0	390	2731

## Análisis del corpus

El análisis desarrollado aquí tiene por fin identificar las expresiones mediante las cuales las AMP ofertan sus productos y, a través del análisis crítico de esos discursos, caracterizar los contextos semánticos en las que se sitúan para clasificar las diferentes promesas y expectativas que están en juego en el sector en Chile.

Para hacerlo, hemos utilizado la técnica de concordancia. Esta consiste en listar todas las ocurrencias de un término particular dentro de un corpus, presentándolas dentro del contexto en el que ocurren, es decir, las palabras que rodean a nuestra palabra clave o *keyword* (Baker, 2006), que contribuyen a dar sentido a lo expresado.

La identificación de estas concordancias es útil para mostrar el rango de los discursos, revelando usos similares o desplazamientos de interés en cuanto a las preferencias semánticas: la relación que hay, no entre palabras individuales, sino entre un lema (la unidad semántica investigada) y un grupo de palabras relacionadas semánticamente (su co-texto). El análisis de concordancias opera bajo el supuesto de que las palabras frecuentemente asociadas, las expresiones comunes y las figuras retóricas y tropos habituales, entre otras unidades discursivas, constituyen unidades conceptuales relevantes para el examen de los conocimientos hegemónicos o típicos en una sociedad o colectivo (Pardo, 2007). Estas observaciones permiten clasificar las expresiones de acuerdo a temas, en categorías temáticas de discurso, que facilitan su análisis crítico.

Con base en dichas categorías temáticas analizaremos las conexiones internas y externas, interdiscursivas, de los textos. Ello nos permitirá examinar, asimismo, las relaciones que controlan la producción de esos textos, cómo los textos afectan y estructuran las prácticas sociales, qué intenciones o preceptos se desprenden de ellos, qué representaciones del contexto social y sus actores están implícitas o explícitas y cuál es la posicionalidad de los hablantes que proponen estos discursos.



## **Etapas del análisis de concordancia**

El análisis de concordancia sobre el corpus de sitios web corporativos de las AMP se llevó a cabo siguiendo, en líneas generales, los pasos sugeridos por Baker (2006, p. 92) .

Tras familiarizarnos con el corpus constituido de la manera señalada más arriba, lo incorporamos para su tratamiento al software AntConc (versión 4.2.4 para MacOS), programa desarrollado por Laurence Anthony (Anthony, 2005). AntConc es una herramienta de análisis de corpus gratuita, multiplataforma y multipropósito, cuya principal característica es un poderoso concordanciador (programa especializado en construir concordancias de forma automática). Las búsquedas de concordancia o *keyword in context* (KWIC o palabra clave en contexto) y el uso de la herramienta n-gram (dentro del mismo software) para rastrear los contextos de secuencias de palabras (en el caso de conceptos de más de una palabra, p. ej. “redes sociales”) que posibilitaron la identificación de expresiones frecuentes y patrones de usos lingüísticos compartidos por las AMP chilenas, se llevaron a cabo de acuerdo a dos etapas.

Primero, usando la lista de expresiones más usuales en la literatura consultada y en nuestro mapeo preliminar, identificamos 14 palabras y conceptos claves (junto a una serie amplia de variaciones) que usamos como términos de búsqueda para abordar cómo se usan efectivamente en el contexto de la oferta comercial de estas empresas (Tabla 4).

Tabla 4. Términos clave para el análisis de concordancia relacionado con las mercancías y servicios del sector AMP.

<b>Término clave</b>	<b>Formas alternativas</b>
Social media	Medios sociales
Redes sociales	Red social; Social network
Análisis	Analysis
Plataforma	Platform
Monitor	Monitorear, Monitoreo, Monitorea, Monitoreado, Monitorean, Monitoreamos, etc.; Monitoring
Analítica	Analytics
Escucha	Escuchamos, Escuchas, Escuchar, Escuchando, etc.; Listen, Listening
Opinión	Opiniones, Opine, Opina, Opinan; Opinion
Sentimiento	Sentimientos; Sentiment
Informe	Informes, Reporte, Reportería, etc; Report, Reporting
Inteligencia	Intelligence
Medidas	Medición, Mediciones, etc.; Measure, Measurement
Minería	Mining
Métrica	Metrics

Segundo, dado que tenemos un interés particular en entender cómo se enfrenta el sector AMP con las problemáticas propias del sistema de las plataformas que hemos delineado en capítulos previos, elaboramos un compendio de palabras clave significativas relacionadas con estos desafíos (Tabla 5).

Para desarrollar esta lista de *keyword*, analizamos tres publicaciones relevantes. En primer lugar, como referencia general para nuestra búsqueda, utilizamos el sobresaliente trabajo de Kennedy (2016) que ya hemos citado, debido a que recurre a fuentes primarias. En dicho estudio,

la autora manifiesta de forma explícita su intención de conocer qué problemas sociotécnicos y éticos afrontan los trabajadores que ejecutan los análisis para estas empresas. Basándonos en sus conclusiones, delimitamos los principales conceptos que hicieran referencia a estos aspectos. Adicionalmente, tomamos en consideración dos publicaciones que desarrollan revisiones sistemáticas de la literatura en torno al análisis de *social media*.

A diferencia de Kennedy y otros autores, Stieglitz et al. (2018) se enfocan en los desafíos y dificultades que enfrenta las investigaciones de las plataformas sociales en las etapas previas al análisis propiamente tal, es decir, durante el descubrimiento, recopilación y preparación de los datos. Para esto, llevan a cabo un examen de la literatura con el fin de identificar qué problemas han sido abordados, en qué proporción y qué soluciones se han propuesto. El trabajo reconoce cinco grandes desafíos, a partir de los cuales extrajimos para nuestros fines una serie de conceptos clave.

Kapoor et al. (2018), por otra parte, realizan una exploración que, a través de un modelamiento algorítmico y estadístico, identifica los 100 temas más relevantes de la literatura en el ámbito de los estudios relacionados con redes sociales y asocia a cada tema un conjunto de *keywords*. Las fuentes consultadas en este caso (132 artículos en total, publicados entre 1997 y 2017) abarcan una amplia variedad de preocupaciones e intereses que ponen en relación aspectos conductuales, sociales, organizacionales, técnicos y de mercado. De los tópicos identificados en esta investigación, seleccionamos aquellos pertinentes a los asuntos problemáticos propios de las diferentes etapas de la minería de datos, de acuerdo a los encuadres que emergieron de la revisión de los trabajos de Kennedy (2016) y Stieglitz et al. (2018) y del conjunto de palabras clave asociadas a dichos temas, seleccionamos aquellas que resultaran menos genéricas.

La finalidad de este conjunto de *keywords* fue identificar si estaban o no presentes y de qué manera en los discursos relacionados con la oferta de las mercancías analíticas de las AMP en Chile, como un indicio del conocimiento, preocupación o claridad que estas empresas tienen o muestran —al menos de cara a sus clientela al momento de ofertar sus “soluciones”— respecto a los desafíos que el campo presenta de acuerdo a una variedad de trabajos académicos.

Tabla 5. Términos clave para el análisis de concordancia relacionado con los desafíos de la minería de datos.

Ámbito	Términos clave
Explicitación de barreras y dificultades	Barrier/Barrera Challenge/Desafío Complexity/Complejidad Difficulties/Dificultades Limitations/Limitaciones Obstacle/Obstáculo Problem/Problema Weakness/Debilidad
Exposición de aspectos metodológicos	Cleaning/Limpieza Control/Control Discovery/Descubrimiento Effectiveness/Efectividad Evidence/Evidencia Filtering/Filtrado Method/Método Practices/Prácticas Preparation/Preparación Sources/Fuentes Steps/Pasos Technique/Técnica
Relacionado con cualidad y calidad de los datos	Accuracy/Exactitud Ambiguous/Ambiguo Authenticity/Autenticidad Credibility /Credibilidad Data errors/Errores de datos Data integrity/Integridad de datos Flawed/Defectuoso Incomplete/Incompleto

	Messy/Desordenado Quality/Calidad Raw/Crudo/a, en bruto Relevant/Relevante Representativeness/Representatividad Structured/Estructurado Semi-structured/Semi-estructurado Trust/Confianza Truthfulness/Veracidad Unstructured/No estructurado Useful/Útil Volume/Volumen Volatile/Volátil
Relacionado con los valores de la investigación	Codes of practice/Códigos de práctica Consent/Consentimiento Disclosure/Divulgación Ethics/Ética Guidelines/Directrices/Pautas/Lineamientos Morally/Moralmente Public/Público/as Responsibility/Responsabilidad Risks/Riesgos Rules/Reglas Security/Seguridad Transparency/Transparencia

## Triangulación

Persiguiendo un mayor rigor en el análisis de los discursos y poner a prueba la calidad de los datos recopilados en nuestro corpus, adicionalmente recopilados la totalidad de los textos disponibles públicamente en tres sitios web corporativos de AMP extranjeras. Las tres compañías seleccionadas (Tabla 6) se ajustaron a los mismos criterios de delimitación del sector AMP

informados previamente en este capítulo, fueron objeto de la producción de un corpus siguiendo los mismos procedimientos y su análisis se llevó a cabo a través de las mismas herramientas y técnicas. El cuerpo de expresiones resultantes fue puesto en relación con las mismas categorías temáticas exploradas en nuestro análisis crítico del discurso de los sitios web corporativos chilenos (como detallaremos en el siguiente capítulo), para de esa forma valorar los discursos identificados en el caso nacional a la luz de lo que muestran otros contextos sociales, políticos e históricos.

*Tabla 6. Detalles del contenido del corpus de sitios web corporativos internacionales.*

<b>Sitio web</b>	<b>N.º de páginas</b>	<b>N.º de reels</b>	<b>N.º de palabras</b>	<b>N.º de caracteres</b>	<b>País</b>
Alto Data Analytics	11	0	5780	40075	España
BrandHook	5	0	2229	14626	UK
Pulsar	14	0	9906	63202	UK

## RESULTADOS

El detallado análisis de la literatura disponible del que hemos dado cuenta en los apartados previos de esta investigación, así como la familiarización con el mercado de las agencias de monitoreo de plataformas sociales a través del escrutinio de sus sitios corporativos, entrevistas y reportajes disponibles, sus comunicaciones a través de perfiles en redes sociales, entre otros, nos permitió en primer lugar determinar las principales características de las empresas del sector AMP y desarrollar los criterios para delimitar el abanico de posibles candidatas en Chile.

Con ello y a través de la aplicación de nuestro diseño metodológico, con apego a nuestros objetivos específicos, desarrollamos un primer mapa del sector de las AMP en Chile que incluyó a 18 empresas que ofrecen servicios asociables a la minería de datos de plataformas sociales (Tabla 2). Como veremos, los siguientes pasos de nuestro análisis nos posibilitaron ajustar aún más el foco de este mapeo preliminar.

Es relevante hacer notar que la aplicación de estos criterios, sobre una lista de más de 30 agencias y consultoras, no implica que aquellas que hayan quedado fuera del análisis no realicen o hayan realizado nunca aproximaciones de minería de datos comercial como parte de sus servicios, sino que a partir de la información promocional de sus sitios corporativos no fue posible determinar que dichos servicios estuvieran dentro de su repertorio de ofertas. Asimismo, la inclusión en la lista final de este análisis tampoco garantiza que dichas empresas desarrollen o subcontraten minería de datos, sino que debido a la jerga que emplean al menos se sugiere, de cara a sus clientes, que ese es una de las herramientas o métodos que tienen a su disposición.

## **Análisis crítico de los discursos en torno a la oferta de las AMP chilenas**

Para el desarrollo de nuestro segundo objetivo específico, consistente en la exploración y evaluación crítica de la oferta de mercancías que caracterizan al sector AMP en Chile, nuestro marco metodológico se apoya en el análisis crítico del discurso (ACD), orientado de acuerdo a las etapas generales detalladas por Mullet (2018). Tras la selección de los discursos a analizar (la oferta de mercancías analíticas en los sitios corporativos de todas las empresas AMP con operaciones en Chile que logramos identificar), la localización y preparación de estas fuentes consistió en la descarga del contenido textual de los sitios de las 18 empresas que cumplían con los tres requisitos que hemos detallado previamente en el apartado “Criterios para delimitar el sector AMP”, lo que produjo un corpus de 87.781 palabras (Tabla 3). En paralelo, se desarrolló un segundo corpus que consideraba tres compañías extranjeras (Tabla 6) y cuyo tamaño fue de 17.915 palabras.

Un último paso de la preparación de los textos, previo a su codificación, se desarrolló mediante la técnica de concordancia, haciendo uso de dos herramientas (las ya mencionadas KWIC y n-gram), con las que cuenta el software AntConc, y 14 conceptos clave y sus variaciones (Tabla 4). Se tomó en cuenta para cada uno de estos términos de búsqueda la inclusión de sus formas plurales (p. ej. sentimiento y sentimientos; social y sociales, etc.), variaciones de las formas verbales (p. ej. monitorear, monitoreamos, monitorean, etc.) y su traducción usual en inglés (p. ej. análisis y analysis; minería y mining; etc.).

Este procedimiento nos permitió circunscribir el análisis a un número acotado de expresiones, frases y párrafos que fueran (a) representativos de discursos que efectivamente se estuvieran utilizando para ofertar comercialmente los productos de estas empresas (por lo que cada co-texto resultante debió ser evaluado individualmente) y (b) no demasiado dispares en términos de extensión, para que la comparación entre los diferentes segmentos o unidades



significantes resultara adecuada (siguiendo a Dey, 2005). Así, tras registrar todas las expresiones resultantes del análisis de concordancia, realizamos una limpieza de los resultados para reconstruir frases truncadas (en casos en que el tamaño del contexto con el que estábamos trabajando fuera menor que el de las expresiones significativas), acotar los segmentos a lo esencial cuando el contexto fuera mayor que lo relevante, eliminar las expresiones repetidas (algo habitual en sitios web que reiteran contenidos en diferentes secciones) o eliminar expresiones en que nuestras *keywords* se estuvieran empleando con usos distintos a los que eran pertinentes para nuestros fines. Adicionalmente, el proceso de análisis y limpieza permitió identificar algunas expresiones relevantes que no necesariamente incorporaban las *keywords* seleccionadas, pero que se encontraban dentro de sus contextos, más atrás o más adelante, y de esta forma fue posible incorporarlas al análisis.

Todo esto dio como producto final un total de 134 expresiones significativas identificadas dentro del corpus.

## **Problemas durante la preparación de los textos**

Entre los problemas identificados en esta etapa del análisis es que algunos de los términos clave inicialmente seleccionados resultaron ser muy genéricos y muchas de las frases identificadas apuntaban a aspectos no relacionados con el objeto de esta investigación. En otros casos, las *keyword* seleccionadas sobre la base de la literatura consultada podían no resultar habituales en los usos discursivos de las AMP en Chile (por ejemplo, “social media”).

Otro desafío fue que muchas de estas empresas realizan una amplia variedad de servicios de consultoría y muchos segmentos discursivos, al ser descontextualizados de su posición precisa dentro de los sitios web, no permitían establecer si estaban haciendo o no referencia al tipo de mercancías que aquí nos preocupa. Esto obligó a un análisis mucho más minucioso y a explorar los segmentos de texto no solo en relación con sus co-textos, sino que además con referencia a las arquitecturas particulares de los sitios corporativos.

Por otra parte, identificamos ocasionales faltas de ortografía en las fuentes, lo que implica la posibilidad de que las búsquedas KWIC con las palabras clave escritas correctamente no dieran siempre como resultado el 100 % de los usos efectivos de esos términos. Otro problema fue la repetición de las mismas frases o de frases muy similares en diferentes partes de muchos de los sitios, lo que incrementó bastante el volumen de expresiones recopiladas mediante el software. Asimismo, nuestras palabras y conceptos clave muchas veces se encontraban aglomerados formando parte de las mismas frases promocionales, por lo que dichas unidades significativas se repetían en la recopilación, incrementando artificialmente el número de registros.

Por todo esto, se realizó un exhaustivo proceso de limpieza de los resultados del procesamiento de los corpus. El análisis de concordancia y la subsecuente limpieza de los datos recolectados permitió, además, mejorar la resolución de nuestro mapeo del sector, ya que en el caso de tres sitios corporativos no se encontraron frases que, asociadas a los co-textos de los conceptos claves empleados, estuvieran con claridad relacionadas con la oferta de mercancías analíticas del tipo que hemos descrito para el sector de las AMP. Estas tres compañías (CyberCenter, Jelly y MediaInteractive), pese a utilizar dentro de los textos de sus sitios corporativos conceptos similares a los que emplean las empresas de minería de datos de redes sociales, no explicitaban de forma suficientemente clara en sus comunicaciones, con frecuencia extremadamente genéricas, su pertinencia al sector comercial de las AMP u ofrecían productos completamente orientados al marketing digital en un sentido restringido. Con ello, la parte del corpus que efectivamente produjo segmentos significativos para el análisis quedó acotada a los sitios web de 15 empresas.

## **Análisis**

A partir de este registro de 134 unidades significativas, procedimos a reducir los datos mediante su codificación abierta (Strauss, 2003) y categorización (Dey, 2005). El reconocimiento de las dimensiones más relevantes expresadas en este conjunto de expresiones nos condujo a la

identificación de una serie de categorías temáticas que se presentaban transversalmente en los discursos. Este método, que supone un acercamiento repetido y atento a los textos registrados para captar más aspectos y cualidades de estos, no agota, desde luego, todas las posibles dimensiones y categorías en las que podría ser organizado el corpus. Sin embargo, apoyados en los procedimientos descritos por Ian Dey (2005), constituye un paso importante para organizar el material y determinar unidades relevantes manejables para poder enfocarnos en su análisis crítico.

De esta forma, los 134 segmentos significativos fueron distribuidos en 11 categorías temáticas, cada una de las cuales se configura como una representación sociosemántica de las mercancías analíticas ofertadas en los sitios web estudiados. Desde luego, muchas de las frases identificadas podrían pertenecer a varios de estos grupos, así como podríamos ampliar la variedad de temáticas atendiendo a diferentes fundamentaciones. En este sentido, la interpretación aquí explorada en torno a ciertos énfasis discursivos es siempre aproximativa y el presente análisis crítico no agota la discusión. No obstante, para los fines analíticos de este trabajo y en atención a nuestros objetivos, hemos optado por las categorías que se detallan a continuación. En cada caso y en pos de la brevedad, no se transcribe la totalidad de los ejemplos disponibles para cada categoría, sino los más representativos o característicos para entender el rango de los discursos que estas incorporan. Las expresiones han sido citadas con extremo cuidado, procurando reproducirlas tal cual como aparecen en los sitios, incluyendo sus errores gramaticales, sintácticos, ortográficos y de puntuación.

En total, se citan 64 segmentos del corpus de AMP nacionales, los cuales han sido numerados correlativamente en formato decimal dentro de cada categoría (1.1, 1.2; 2.1, 2.2; etc.), con el fin de simplificar su referenciamiento en el texto. En el caso de las expresiones del corpus internacional, los ejemplos citados, 17 en total, han sido jerarquizados alfabéticamente dentro de cada categoría (1.a, 1.b; 2.a, 2.b; etc.). Las categorías están ordenadas de acuerdo a la frecuencia con la que se manifestaban en el corpus, de más frecuente a menos frecuente. La figura a continuación (Figura 2) ilustra las proporciones relativas de las distintas categorías de análisis.

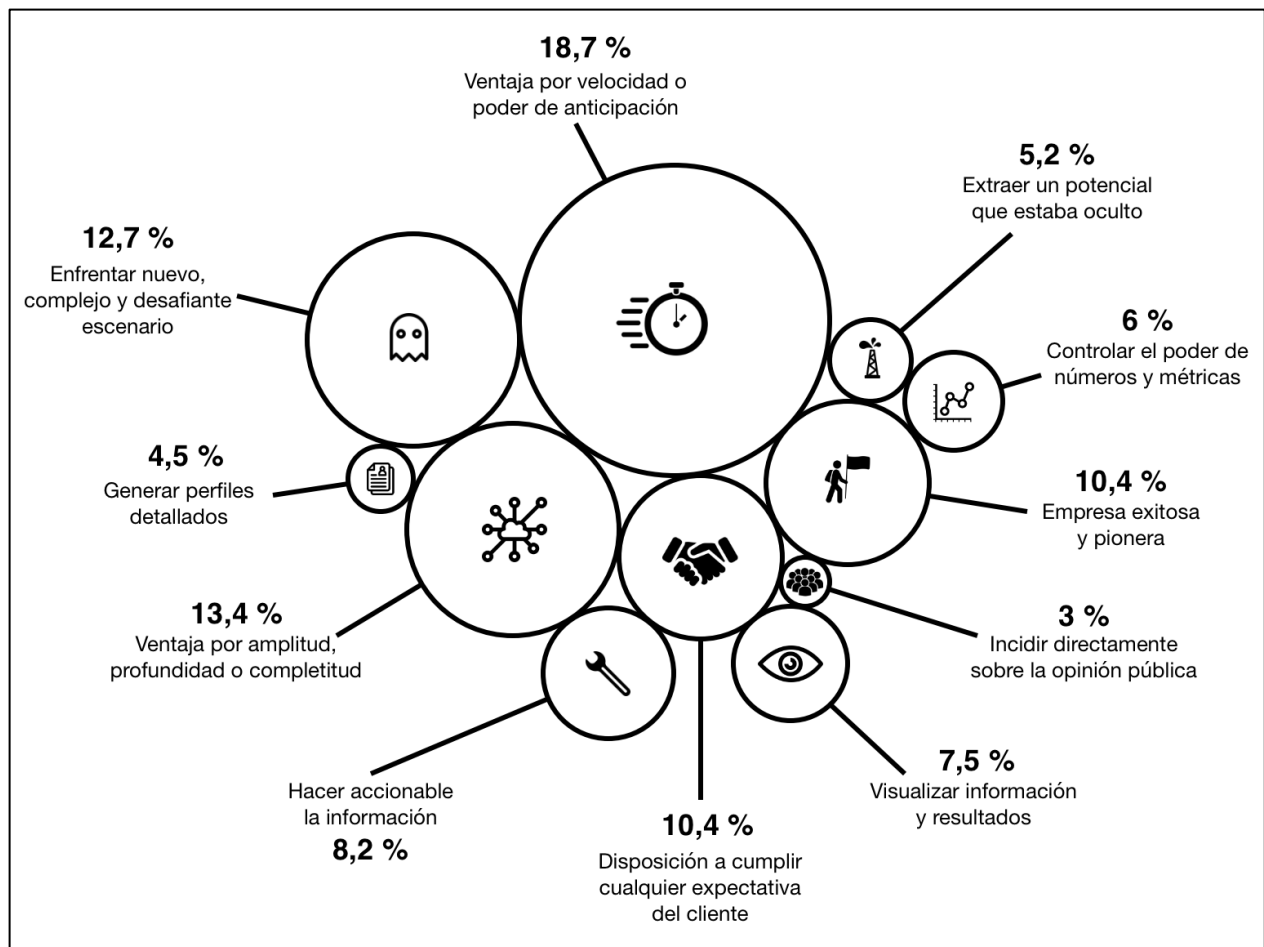


Figura 2. Proporciones relativas de las 11 categorías de análisis identificadas de acuerdo a su aparición en el corpus. Elaboración propia.

### Categoría 1: El producto concede una ventaja por su velocidad o poder predictivo

1.1	El servicio está enfocado en la detección oportuna de cualquier evento crítico, el cual pueda generar algún impacto en las operaciones de la compañía, a través de un monitoreo en tiempo real de redes sociales y medios digitales
1.2	Monitorea de manera planificada y en tiempo real lo que impacta a TU organización.
1.3	es una servicio personalizado de datos sociales categorizados por Medios de Comunicación, Influencers y Movimientos Sociales; adaptándose al dinamismo de las RRSS y la inmediatez del impacto ante una contingencia.
1.4	Más valor adquiere esa capacidad analítica cuando uno enfrenta situaciones de crisis donde necesita claridad en poco tiempo para reaccionar y tomar el control de la narrativa.

1.5	Monitoriza constantemente tu reputación en LinkedIn y obtén valiosas perspectivas sobre cómo tu empresa es percibida en la red profesional.
1.6	Cuidamos la reputación de nuestros clientes con herramientas sofisticadas de escucha social y equipos que trabajan 24/7 para dar a las áreas de comunicaciones y relaciones corporativas las alertas necesarias para definir planes de acción y contención.
1.7	escuchamos las redes sociales para conocer las preocupaciones de las comunidades en torno a los proyectos, lo que nos ha permitido anticipar tendencias y detectar riesgos reputacionales.
1.8	Seguimos los temas de conversación y su temperatura, anticipando potenciales crisis y aportando insights en tiempo real de cara a la toma de decisiones.
1.9	Todo con el objetivo de conocer el pulso de una campaña o predecir futuras conversaciones y crisis comunicacionales.
1.10	Monitoreamos y generamos conversaciones a través del social listening para anteponerse a los intereses y necesidades de las comunidades.

Este es el grupo de unidades significativas más abundante, concentrando más del 18 % de los segmentos identificados. Dentro de esta categoría se incluyen las menciones que al hablar de los productos y servicios de las AMP enfatizan la variable temporal. Esto considera una variedad de mensajes que van desde aquellos que se enfocan en la velocidad, la capacidad de poder registrar los fenómenos de forma simultánea a su ocurrencia o, inclusive, que sugieren un cierto poder de adelantamiento que bordea lo predictivo.

Lo más frecuente aquí es caracterizar las mercancías analíticas como objetos que otorgan la capacidad de observar en tiempo real (ejemplos 1.1 y 1.2), es decir, saber lo que ocurre en el mismo instante en que está ocurriendo, para adaptarse a la velocidad de las redes sociales (ejemplo 1.3) y actuar de forma inmediata (ejemplo 1.4). La rapidez es lo esencial en esta categoría, pero también se expresa una capacidad de abarcar la totalidad en la dimensión temporal cuando se habla, por ejemplo, de un monitoreo constante (ejemplos 1.5 y 1.6).

De forma más interesante aún, en algunos casos se les atribuye a las mercancías el poder de predecir acontecimientos. Si la rapidez es un atributo relativo cuando se trata de minería de datos de plataformas sociales, la capacidad de anticipación (ejemplos 1.7 y 1.8), de “predecir

futuras conversaciones” (ejemplo 1.9) o “anteponerse a los intereses” de las comunidades (ejemplo 1.10) es todavía mucho más discutible. No obstante, las expresiones orientadas a ofrecer tales competencias resultaron ser llamativamente frecuentes dentro de la oferta de las AMP.

En los sitios corporativos de las empresas no chilenas analizadas, estos énfasis también se encontraban representado de manera transversal, haciendo referencia tanto a la capacidad de observar en tiempo real y reaccionar con rapidez, como a la de adelantarse a los eventos, como en los siguientes ejemplos.

1.a	La tecnología y experiencia superior de Alto ayudan a las organizaciones a mantener una visibilidad nítida en cada punto de contacto digital para monitorear, evaluar y responder con confianza tomando decisiones informadas en tiempo real a través de Alertas, informes en tiempo real o Informes de Inteligencia de Amenazas.
1.b	Predice el potencial de una tendencia a través de datos de búsqueda, redes sociales y noticias con nuestro Virality Framework patentado, basado en una década de análisis sobre cómo se propaga el contenido en línea.

**Categoría 2: El producto concede una ventaja por la vastedad, profundidad o completitud de la información que abarca**

2.1	Brandmetric le ofrece una visión completa de cómo su marca está siendo percibida en línea.
2.2	Escuchamos todo lo que pasa en las redes sociales, diferenciando un tweet de un trend.
2.3	Monitorea todas las conversaciones sobre tus marcas, temas, actores o competencia, ya sea en redes sociales como Twitter, Facebook, Instagram, YouTube, Google+, blogs, foros y sitios online.
2.4	La potente tecnología de búsqueda del MonitorConecta tiene alcance global y te asegura estar encima de cada conversación pertinente.
2.5	nos apasiona ayudar a las empresas a obtener una comprensión profunda de su presencia en línea y a tomar decisiones informadas basadas en datos.

2.6	Ponemos a tu disposición una potente plataforma de escucha de redes sociales y canales digitales diseñada para dar respuesta a las preguntas ¿Qué?, ¿Cómo?, ¿Quiénes? y desde ¿Dónde?
-----	---

Dentro de esta categoría se incluyeron 18 descripciones en que los productos son dibujados como articuladores de un conocimiento hiperbólico en cuanto a lo que abarcan, con una capacidad tendiente al registro de la totalidad de lo posible o cuya cualidad es inédita en términos de profundidad y variedad. Las mercancías analíticas ofrecen, bajo este prisma, completitud, totalidad y variedad (ejemplos 2.1, 2.2 y 2.3), así como profundidad, potencia y la capacidad de responder cualquier tipo de pregunta (ejemplos 2.5 y 2.6).

En el caso de la amplitud y profundidad que presuntamente ofrecen sus mercancías analíticas, las AMP extranjeras resultaron ser aún más enfáticas, si se puede, respecto a los volúmenes y la variedad de datos que analizan. A diferencia de las AMP chilenas, en los casos foráneos la imagen de la extrema amplitud se representa de formulaciones que pretenden ser más concretas, muchas veces a través de expresiones numéricas, como en los siguientes casos:

2.a	Nuestra tecnología de clase mundial ofrece inteligencia sin límites que abarca miles de millones de puntos de datos y permite a nuestros clientes acceder a inteligencia sobre amenazas de todo el espectro digital.
2.b	Alto utiliza un análisis de redes avanzados para señalar anomalías en la esfera digital con una visión en 53 idiomas y 125 países.
2.c	Pulsar recopila datos sociales, de búsqueda, de medios de comunicación y web de 195 países, detectando 62 idiomas y comprendiendo el sentimiento en 24.

### **Categoría 3: El producto permite enfrentar a un nuevo, complejo y desafiante escenario**

3.1	NO SE QUEDE FUERA DE LA CONVERSACIÓN En las redes sociales surge información minuto a minuto
3.2	Como si eso fuera poco, medir y analizar los resultados de esta gestión se volvió una demanda ineludible para retroalimentar la toma decisiones y reportar resultados.

3.3	El mundo cambia rápidamente y nuestros clientes deben cumplir expectativas cada vez más exigentes. Nuestro desafío es ayudarlos a adelantarse a esos cambios.
3.4	Hoy, las redes sociales forman opinión, crean realidades y marcan tendencias. Realizamos monitoreo, seguimiento y gestión de los temas que afectan a la empresa o institución.
3.5	En un contexto en donde los datos emergen de casi todos los aspectos de la vida y de la empresa, en Feedback Research generamos, ordenamos, analizamos y visualizamos datos, para transformarlos en información clave en la toma de decisiones.
3.6	En un mundo cada vez más segmentado e hiperconectado, los canales digitales son medios críticos para gestionar el capital reputacional de las empresas y organizaciones
3.7	En la Era de la información, identificar qué, cuándo, dónde y quiénes están interactuando en nuestros ecosistemas es de suma relevancia para enfrentar la toma de decisión a la que nos vemos expuestos día a día.
3.8	Para ayudar a los negocios capturamos, procesamos y clasificamos Social Data con la única finalidad de resolver Tus preguntas de qué pasa allá afuera en los Ecosistemas Digitales.
3.9	Noticias que no paran y redes sociales voraces... son un desafío titánico para los profesionales de las comunicaciones.

Bajo esta categoría, la oferta pasa por transmitir un imperativo de urgencia frente a un mundo que se ha vuelto inabarcable e inaprensible por su tamaño, novedad o complejidad. Además de la idea de premura (ejemplos 3.1, 3.2 y 3.3), en esta categoría son frecuentes las expresiones de tono pedagógico sobre los cambios radicales e ineludibles que deben enfrentar los clientes, con formulaciones que incluso llegan a sonar indulgentes, dirigidas a personas que no manejan los códigos que dan forma a estas inéditas condiciones de la sociedad y los negocios (ejemplos 3.4, 3.5, 3.6 y 3.7). Aquí aparecen algunas figuras retóricas coloridas, como la dicotomía entre lo interno y lo externo (ejemplo 3.8), dibujando a lo externo (“allá afuera”) como lo incierto y peligroso, y la idea de la “voracidad” de las plataformas sociales y las dificultades de proporciones colosales, “titánicas”, que plantean (ejemplo 3.9).

El foco comunicacional en estos casos está orientado a potenciar la idea de que nos encontramos frente a una realidad inestable, caracterizada por una horizontalidad radical entre hablantes y oyentes, en que la seguridad ontológica de la élite se ha perdido o se encuentra en



declive. Las AMP se ofrecen mediante estos mensajes como las organizaciones con la capacidad de poner orden a la vastedad de los datos, a defender al cliente de la “voracidad” de las redes sociales y de sus usuarios antropofágicos, y “transformarlos en información clave en la toma de decisiones”. Esto apunta, así, a unas determinadas mercancías con el potencial de permitir la recuperación de una hegemonía en riesgo.

La idea de un nuevo espacio, una esfera digital en que las prescripciones comunicativas y de poder se han distorsionado respecto a una norma previa (ejemplo 3.a), también figura con frecuencia en las AMP extranjeras de nuestra muestra.

3.a	La esfera digital está cada vez más abarrotada y llena de una gama diversa de actores que compiten por ser escuchados y llegar a las audiencias clave.
-----	--

#### **Categoría 4: El producto es una empresa exitosa y pionera en el ámbito tecnológico**

4.1	Con 10 años de experiencia en el desarrollo de herramientas de escucha y análisis de redes sociales, y trabajado con múltiples industrias tanto a nivel nacional como internacional, sabemos que somos el mejor complemento para otorgar un servicio de excelencia a todos tus clientes.
4.2	Los últimos 25 años hemos buscado acompañar las grandes transformaciones de la comunicación en la vida digital con el foco en ayudar a los profesionales de la industria de todo el mundo
4.3	Cuentan con un equipo multidisciplinario con expertise y están disponibles siempre que lo necesitamos, 24/7
4.4	contar con un equipo multidisciplinario compuesto por profesionales expertos en sus respectivas áreas de conocimiento nos convirtió en la primera empresa chilena que lanza al mercado una plataforma de monitoreo y análisis automático de redes sociales en Chile y ser distinguidos por Conicyt, Corfo, Prochile y la Cámara de Comercio de Santiago.

Otro tipo de unidad significativa muy habitual en los sitios corporativos investigados expresa la idea de que la mercancía por excelencia de las AMP es la empresa misma, antes que

sus productos. Esta dimensión está muy relacionada, como veremos, con la siguiente categoría: en ambos casos lo que se oferta, ante todo, es el carácter de la relación entre agencias y clientes.

Dentro de esta dimensión se da énfasis a la historia de la empresa, el tiempo que llevan en el mercado (ejemplos 4.1 y 4.2) y otras características distintivas, como contar con equipos multidisciplinarios, ser reconocidos o ser pioneros tecnológicos (ejemplos 4.3 y 4.4). Es destacable que muchas de las empresas mencionan en algún espacio importante de sus sitios corporativos la cantidad de años desde su fundación (de los 18 sitios del corpus, 13 indican la longevidad de su empresa), lo que contrasta con el poco relieve que se manifiesta, en términos generales, respecto a aspectos tales como los detalles técnicos de sus aproximaciones y la tecnología o metodologías que emplean.

El mensaje para los clientes es que al acceder a sus productos lo que se obtiene es una relación con equipos experimentados (aunque rara vez se aclare en qué lo son), que entregan confianza (el aval es el tiempo que llevan en el mercado) y permanentemente disponibles atender las necesidades del cliente. Esta categoría de la oferta también se presentó en los ejemplos extranjeros, aunque en casos como los ejemplificados a continuación, los detalles de la *expertise* se explicitan con mayor claridad.

4.a	El equipo de datos de Alto Insights cuenta con una variedad de antecedentes y experiencia incluyendo sociología, diseño, matemáticas, ciencia de datos, periodismo, negocios o ciberinteligencia.
4.b	Pulsar fue creado a principios de la década de 2010 dentro de la práctica de insights sociales en la consultora de innovación Face. Cuando algunos clientes de Face comenzaron a pedir herramientas para comprender las conversaciones que tenían lugar en las redes sociales, yendo más allá de simples análisis para desbloquear información sobre sus audiencias en línea, nació Pulsar.

## **Categoría 5: El producto es una empresa dispuesta para cumplir cualquier expectativa de sus clientes**

5.1	Nuestro equipo está preparado para ayudarte a ti y a tu organización a sacar el máximo provecho a la plataforma a través de un acompañamiento continuo y soporte personalizado.
5.2	Nuestra tecnología convierte tus análisis cualitativos en analítica, gracias a los procesos de Machine Learning y Social Listening, los cuales son 100% personalizados a las necesidades de tu negocio.
5.3	Solo tienes que decirnos que escuchar, del resto nos encargamos nosotros.
5.4	tenemos un sello que nos caracteriza: la agenda del cliente es nuestra agenda.
5.5	Discutimos poco, medimos mucho
5.6	entendemos que cada enfoque estratégico es único, y por lo tanto las métricas e indicadores deben estar pensados a la medida.
5.7	Nuestros servicios personalizados son un traje a la medida para cada cliente. Para ello contamos con equipos multidisciplinarios, encabezados por un socio en cada cuenta.

Esta categoría (aproximadamente un 10,5 % del total de unidades recopiladas) guarda evidentes lazos con la categoría anterior. Aquello que se comercia a través de estas expresiones es, ante todo, la relación entre la agencia y el cliente, pero aquí se manifiesta un énfasis más explícito: el foco no es destacar características de la empresa o sus productos, sino las cualidades de la relación que se teje entre ambas partes.

Un aspecto de esta relación consiste en, por supuesto, cumplir con las expectativas de los mandantes y ofrecer servicios “personalizados” y de forma “continua” (ejemplos 5.1 y 5.2), pero los casos aquí listados van, también, un poco más allá. El control respecto a los atributos de los resultados entregados por las AMP es dejado por completo a la voluntad de los clientes, en un giro que es opuesto a las aseveraciones del sector respecto a la producción de conocimiento basado en abundantes datos (categoría 2). Aquí priman las orientaciones y agendas del cliente, las cuales se “discuten poco” (ejemplos 5.3, 5.4 y 5.5). Cada cliente requiere resultados y métricas particulares, por lo que pese a que estas se fundan supuestamente en observaciones empíricas realizadas con instrumentos estandarizados sobre sistemas que, en lo esencial, están fuera del control de las AMP, las empresas ofrecen “trajes a la medida”. Esto implica relaciones plásticas y

de alta confianza con los clientes, lo que puede explicar en parte la confidencialidad que se desprende de los escasos detalles técnicos que aportan los sitios corporativos.

Este tipo de expresiones también figuran en las AMP extranjeras, como vemos en los siguientes ejemplos.

5.a	Programas de investigación confeccionados a la medida de las necesidades de su negocio, aprovechando la experiencia de nuestro equipo interno para diseñar y entregar informes de insights y orientación estratégica.
5.b	Los análisis y reportes se llevan a cabo para cumplir con sus objetivos estratégicos específicos como máxima prioridad por nuestro equipo de expertos científicos de datos.

**Categoría 6: El producto permite hacer accionable la información**

6.1	La principal fortaleza de Artool es que no te entregan informes planos, sino que hacen accionable la data, en base a un análisis profundo y realmente valioso en cuanto a insights.
6.2	Transformamos conversaciones en canales digitales sobre tu marca, mercado o intereses, en insight accionables para tomar la decisión correcta en el momento exacto.
6.3	Analistas especializados en tu cobertura crean informes que te entregan conocimiento accionable, muestran tu crecimiento y hacen evidente el impacto en el éxito de tu empresa.

El concepto de “conocimiento accionable” tiene un lugar destacado en el mercado y la literatura internacional relacionada con la minería de datos comercial. Su uso también apareció de forma transversal en el corpus que hemos analizado. Este concepto apunta a la elaboración de un tipo de información que tiene la potencialidad de ser puesta en práctica y permitir la toma de decisiones gerenciales. En el ámbito de la minería de datos desarrollada por las AMP, el concepto está muy vinculado con la siguiente categoría de expresiones, relativa a una oferta que prioriza la visualización de los datos, ya que esta información se vuelve “accionable” al ser presentada de una manera comprensible y evidente para los clientes (ejemplo 6.3).

Pese a que la palabra misma es un préstamo reciente del idioma inglés, en el corpus internacional solo una AMP usaba este lenguaje (Alto Data Analytics), aunque de manera frecuente y en los mismos términos ya señalados.

**Categoría 7: El producto ofrece la capacidad de visualizar la información y mostrar resultados**

7.1	Ponemos a tu disposición una potente herramienta de escucha en la cual podrán visualizar en todo momento, tendencias, percepción de usuarios, temas de conversación, influenciadores, entre otras.
7.2	A través del "social listening" vas a poder visualizar la opinión del público e influencers con respecto a tu negocio.
7.3	Nuestra plataforma de Social Monitor está pensada en entregar una mirada rápida y de fácil acceso a nuestros clientes
7.4	Nuestras plataformas permiten extraer analítica, gráficos, datos y noticias específicas en cualquier momento para compartir resultados con el directorio u otras miembros de la organización.
7.5	Demuestra con datos el éxito de tu gestión Explicar el real valor de cada aparición en los medios, crear benchmarks e indicadores y mostrar el aporte global al negocio puede ser una tarea abrumadora.

En esta categoría, que conformó cerca de un 7,5 % del total, hemos considerado aquellas expresiones en que el aspecto visual de las mercancías analíticas aparecía como lo más destacado. Sin embargo, es relevante señalar que con mayor o menor énfasis, la idea de que las AMP contribuyen a “mostrar”, “graficar” o “visualizar” datos resultó bastante ubicua y varias de las unidades significativas que hemos situado en otras categorías también hacen referencia en algún grado a una oferta con esas características.

Las referencias a la capacidad de visualizar apelan a un uso “interno” de los productos de las AMP, que posibilitan a quienes las contratan en primer lugar alcanzar por sí mismos una comprensión de aquel conocimiento que están comprando. Por ello, en esta categoría se incluyen también frases que prometen entregarle a miembros de las organizaciones el poder de “mostrar” o

“demostrar” internamente resultados y cumplir expectativas alojadas en el seno de sus empresas (ejemplos 7.4 y 7.5).

Como han señalado múltiples autores, la visualización de datos es en la actualidad central en la comunicación política. Su poder persuasivo radica en la idea de que estas representaciones visuales permite que los espectadores experimenten por sí mismo el descubrimiento de patrones (Coopmans, 2014), como si se estuviera experimentando de primera mano un fenómeno. Dado que visualizar no es lo mismo que entender, el reemplazo de estos términos en frases como “vas a poder visualizar la opinión del público” (ejemplo 7.2), en lugar de entenderla, es muy interesante. Habla con fuerza del poder persuasivo de lo visual en estos contextos y de un consenso en cuanto a su capacidad instrumental.

Como señalan Kennedy y Hill ( 2017) y Kennedy y Engebretsen (2020) , nuestro compromiso emocional con las visualizaciones de datos está estrechamente relacionado con sus aspectos estéticos. Un material gráfico estéticamente agradable pueden generar la ilusión de haber alcanzado la comprensión sobre algo (Laaksonen y Pääkkönen, 2020). En este sentido, una forma de ofertar las mercancías visuales de las AMP se sustenta en discursos que nos transmiten determinadas características estéticas de estas, como dinamismo y sencillez (ejemplo 7. 3). Estos aspectos de la visualización de datos son abordados de forma mucho más explícita en algunos de ejemplos extranjeros:

7.a	Los entregables incluyen informes visuales exclusivos de alto impacto confeccionados a la medida de sus preguntas estratégicas más importantes.
7.b	Monitoree el impacto, intensidad y matices de cada opinión en línea con nuestra IA de Visibilidad y Emoción, y vea cómo diferentes influencers, partes interesadas y personas están afectando la reputación de su marca, en noticias y redes sociales.
7.c	Sumérjase en tiempo real desde la visualización de miles de millones de puntos de datos hasta el nivel granular de conexiones individuales. El seguimiento de manos y movimientos permite una interactividad única para que pueda explorar libremente datos e insights desde su propio punto de vista y grabar videos de realidad virtual excepcionales con su análisis personalizado.

## **Categoría 8: El producto otorga control sobre el poder o atractivo de los números y métricas**

8.1	Ponemos a tu disposición todas las métricas necesarias para poder evaluar el impacto generado por tus comunicaciones y la recepción de tus seguidores en redes sociales.
8.2	Creemos que la creatividad debe ir de la mano con la medición, por eso los números son protagonistas en todo lo que hacemos.
8.3	Vas a poder apoyarte en datos de valor para tomar mejores decisiones.
8.4	Creemos que las comunicaciones pueden y deben ser monitoreadas, planificadas y medidas utilizando el método científico y estamos aquí para ayudarlos a concretar esta visión
8.5	Nos gusta que las métricas nos hablen para definir objetivos y metas, que alcanzamos con creatividad, diseño e innovación.
8.6	Tenemos una obsesión por las métricas por lo que todas nuestras hipótesis las medimos y testeamos con el fin de entregar un buen argumento.

Esta categoría es la que mejor expresa el sublime digital, la percepción sobre la superioridad intrínseca del conocimiento basado en la cuantificación creciente de todo, que considera a los números y las métricas como intérpretes exclusivos de la realidad, por sobre las subjetividades, experiencias e interpretaciones humanas. Los discursos que emergen tras las expresiones dentro de esta categoría son análogos a valoraciones sobre las mercancías analíticas de las AMP como la realizada por el exministro Gonzalo Blumel cuando presumía de estar en posesión de una “información extraordinariamente sofisticada, a partir de análisis con tecnología de *big data*”.

Clasificamos cerca de un 6 % del total de las frases analizadas como mensajes que ponen en relieve la entrega de números, estadísticas, métricas o elementos cuantitativos. Pese a que para autores como Dijck, Poell y de Waal (2018) la cuantificación creciente de la audiencia y la producción de métricas son parte fundamental de la economía de las plataformas, si excluimos las menciones al concepto de “dato”, generalmente presentado con vaguedad en el corpus analizado, los aspectos relativos a las mediciones y los números en los discursos asociados a los productos del sector AMP son poco preeminentes. Esto no quiere decir, por supuesto, que la conversión de la experiencia humana en métricas no sea relevante o central, técnicamente hablando, para el

sistema de las plataformas. Pero al menos, en lo que respecta a los discursos de las AMP consideradas aquí, ese “deseo por los números” (sean estos o no precisos) que destaca Kennedy (Kennedy, 2016) no es el principal “gancho”.

En el corpus de empresas extranjeras, esta categoría tampoco se situó como un tema destacado en nuestro análisis.

**Categoría 9: El producto posibilita el uso de información previamente desperdiciada o extrae un potencial que estaba oculto**

9.1	Extraemos que dicen las personas acerca de tu marca, empresa y competencia.
9.2	Te invitamos a sacar partido del potencial que nos entregan los canales digitales y redes sociales a través de nuestro servicio de reportería personalizada.
9.3	nos permite identificar nuevas tendencias asociadas a los consumidores, evaluar el impacto de las acciones de tus competidores y en definitiva levantar nuevas oportunidades de negocio que puedan ser explotadas.
9.4	¿cuál es su verdad? Lo macro nos lleva a lo micro y lo micro a lo macro. El mundo digital nos da herramientas para obtener información de usuarios a gran escala, y otras veces solo un relato individual revela insights grupales que ningún sistema detecta.
9.5	¿Cómo podemos apoyar a las organizaciones a tomar mejores decisiones basándonos en información que no están utilizando? "Aplicando técnicas de análisis automático de texto para entender las conversaciones en redes sociales y aprovechar oportunidades de negocio que puedan tener impacto en los ingresos y la satisfacción de los clientes”.

Esta categoría está vinculada con la tercera categoría de este análisis, que establece y advierte de la existencia de una nueva y desafiante realidad de difícil acceso. Aquí, el énfasis está puesto en la capacidad de las AMP y sus productos de acceder de forma privilegiada a un potencial vedado para otros actores. Estas expresiones explicitan las cualidades de las mercancías analíticas de, supuestamente, revelar lo que hasta ahora ha estado oculto, de dar acceso a lo que de otro modo (o para otros) sería inaccesible. Los términos empleados para describir este potencial son “identificar”, “aprovechar”, “sacar partido”, “extraer” o “explotar”, acciones que



con diferente grado de agresividad permiten refinar una materia prima que aparece como confusa y heterogénea, para obtener valor de ella.

En el caso de las empresas extranjeras, si bien sus expresiones podían sugerir la idea de un potencial desperdiciado u oculto (“mirar donde nadie más lo ha hecho”; conocer aquello que los usuarios “realmente” piensan y sienten), en general no incorporan analogías extractivistas como las detectadas en el corpus chileno.

**Categoría 10: El producto permite generar perfiles detallados de las personas**

10.1	Vemos en tiempo real la percepción (sentimiento) de tu audiencia de interés, cuáles son sus temas de conversación (clasificación de temáticas), quienes estén participando en la conversación (influenciadores y tipos de usuarios), cómo interactúan entre ellos (redes) y en qué lugar se encuentran.
10.2	El contratar esta cobertura permite saber qué se habla de tu marca o empresa en las redes sociales y la esfera del microblogging. Además conocer la geolocalización de los mensajes publicados.
10.3	Observatorio de marcas mineras es un proyecto que registra y analiza periódicamente las menciones en RRSS de distintas empresas mineras, generando análisis de sentimientos, geolocalización y benchmark del estado de reputación de las diversas empresas mineras del país.
10.4	A partir del análisis de datos, monitoreamos y microsegmentamos audiencias y gestionamos contenidos claves para impactar de manera efectiva a públicos cada vez más fragmentados.
10.5	Obtenga información valiosa sobre las preferencias, expectativas y preocupaciones de sus clientes, empleados, proveedores y otros actores relevantes.

En esta categoría se incluye la oferta de mercancías que prometen la elaboración de perfiles más o menos detallados y profundos de diferentes aspectos demográficos y psicológicos de las audiencias en línea, así como de sus relaciones, incluyendo elementos tales como análisis de percepciones, sentimientos y temas, interacción entre personas y geolocalización. Algunas de las formas de ofrecer estos perfilamientos sugieren procedimientos que requieren acceso a muchos más datos que los que con frecuencia están disponibles a través de las API de las

plataformas sociales, como la geolocalización (ejemplos 10.1, 10.2 y 10.3) e incluso el desarrollo de un trabajo de inteligencia en un sentido casi policial (ejemplo 10.5).

Las promesas que en este sentido hacen las AMP extranjeras son algo más modestas a la hora de ofrecer revelaciones demográficas o psicográficas, salvo por el siguiente ejemplo de la agencia Pulsar:

10.a	Obtenga una vista granular de sus segmentos de audiencia clave para comprender sus opiniones, impulsores y actitudes, luego desglose sus grupos de audiencia y stakeholders por comportamiento, demografía y psicografía.
------	---

**Categoría 11: El producto posibilita incidir directamente sobre la opinión pública o la sociedad**

11.1	El seguimiento online de las redes sociales permite frenar rápidamente el inicio de una crisis de imagen con pésima publicidad para la marca, actuando como un "detector de humo" previo al inicio del "incendio" mediático.
11.2	Monitoreamos la recepción de personas y medios en canales digitales en tiempo real de todas tus campañas e iniciativas, permitiendo potenciar mensajes con recepción positiva y evitar aquellos que no.
11.3	Desplegamos el relato de las organizaciones en el ecosistema digital y generamos conversaciones.
11.4	Desautorizamos los enlaces tóxicos que mencionan a tu marca de manera negativa y monitoreamos su evolución.

Identificamos solamente cuatro expresiones (3 % del total) que pudiéramos asociar a esta categoría. Lejos de ofrecer un producto meramente de análisis o de observación, en estos casos se ofrece una mercancía analítica que permitiría incidir de forma activa y directa en la opinión pública por diversas vías, como son “frenar” crisis de imagen, “potenciar” mensajes positivos, “generar” conversación o suprimir enlaces negativos para el cliente en redes sociales. En estos casos no se trata, como hace el marketing y la publicidad, de influir por medio de la mera persuasión, sino que mediante diversas estrategias de manipulación de las plataformas.

En el caso de los sitios no chilenos analizados, la promesa de una incidencia comunicacional directa resultó ser un poco más común (9,7 % respecto del total), como se aprecia en los siguientes ejemplos.

11.a	Con una intensa competencia por la atención, la tecnología incomparable de Alto significa que usted puede entregar un mensaje poderoso dirigido a la audiencia adecuada para obtener la máxima cuota de voz.
11.b	Identifique instantáneamente dónde están las oportunidades y desafíos y coordine respuestas poderosas basadas en evidencia para expandir su influencia con confianza y rapidez en múltiples plataformas digitales en tiempo real.
11.c	Descubra las principales comunidades en su audiencia y vea cómo cada una de ellas habla sobre el mismo tema de manera diferente, para adaptar su enfoque de contenido y de segmentación.

### **Análisis de los discursos en torno a los desafíos técnicos y metodológicos asociados al sector**

Como han sostenido autores como Driscoll y Walker (2014), para que el público pueda evaluar las interpretaciones que emergen de los datos de las plataformas sociales, necesita tener información sobre cómo se recopilan, almacenan, limpian y analizan dichos datos. Idealmente, no solo información sobre los aspectos técnicos relacionados con la minería de datos deberían estar al alcance de los potenciales clientes de las AMP, también advertencias sobre las condiciones de elaboración de esos datos dentro del sistema de las plataformas, la infraestructura involucrada, las características de las fuentes de acceso a los datos en cada momento (las diferencias entre las API gratuitas y de pago, las limitaciones de las diferentes API y la actualización constante de los cambios en sus políticas), etc. En suma, una actitud transparente en torno al negocio de la minería de datos debería ser explícita en indicar que las plataformas sociales están lejos de ser facilitadores neutrales que se limitan a dar acceso a indicios naturales de fenómenos sociales.

Para explorar cómo aborda esta problemática el sector AMP chileno, elaboramos una lista de términos clave relacionados con desafíos y dificultades discutidos por diversos investigadores en torno a la minería de datos de plataformas sociales (Tabla 5). Con estas *keywords*, llevamos a cabo un nuevo análisis de concordancia sobre nuestro corpus (descartando en este caso las tres compañías que no arrojaron resultados en la etapa previa). Luego de la limpieza del registro resultante, obtuvimos una muestra de expresiones relevantes para la discusión, que fueron distribuidas para el análisis crítico de estos discursos de acuerdo a 4 ámbitos generales.

### **Ámbito 1: Explicitación de barreras y dificultades**

Las *keyword* englobadas bajo este ámbito apuntan a identificar discusiones relativas a las dificultades técnicas relacionadas con la minería de datos de plataformas, incluyendo las barreras de acceso a los datos, las complejidades de las herramientas, los obstáculos y limitaciones que imponen las API y otros desafíos.

Sin embargo, mediante el análisis de concordancia no encontramos expresiones que abordaran de forma manifiesta estas problemáticas. En los segmentos identificados, las dificultades y barreras podían ubicarse en el lado del cliente (para ser resueltas por los productos de las AMP) o como desafíos técnicos del pasado (ya superados por la inventiva de las AMP).

### **Ámbito 2: Exposición de aspectos metodológicos**

Los conceptos de este ámbito buscan explorar si, como parte de los discursos de las AMP, se explicitan algunos de los elementos metodológicos de la minería de datos de redes sociales, incluyendo explicaciones sobre las diversas etapas del proceso de “refino” de los datos (descubrimiento, selección de fuentes, preparación, limpieza), capacidades humanas o materiales requeridas y usadas para su manejo, precauciones relacionadas con las técnicas de registro, almacenaje, filtrado, etc., controles aplicados, niveles de confianza esperables, entre otros.

A través del análisis de concordancia llevado a cabo, detectamos 5 expresiones que podrían estar haciendo referencia a algunos de estos aspectos.

A	Cómo expertos multidisciplinarios, nos enfocamos en descubrir, diseñar y ejecutar experimentos dentro de tu negocio que nos permitan escalar soluciones.
B	Analizamos Transformamos el universo de datos obtenidos en información procesada, filtrada y organizada.
C	Obtener información procesada y filtrada de una enorme cantidad de fuentes.
D	El almacenamiento de Social Data a través de sus procesos de captura, procesamiento y clasificación de ellos es el punto de partida para cualquier estrategia vinculada a la inteligencia y analítica de negocios.
E	Usamos Python, el lenguaje de programación funcional que lentamente se ha ido posicionando como una herramienta indispensable para los científicos sociales que permite analizar grandes volúmenes de información digital y automatizar procesos de alta complejidad

Como vemos, las AMP no profundizan en ningún aspecto metodológico. Las alusiones a procedimientos de investigación son, como mucho, vagas (A, B, C). Cuando los procedimientos se mencionan con algo más de precisión (D), no hay un desarrollo que provea algún tipo de inteligibilidad. En un único caso se menciona una competencia concreta usada para el análisis de los datos (E).

Aparentemente, al hablar del procesamiento de datos (su descubrimiento, organización, filtrado, etc.), las expresiones hacen referencia solamente a procedimientos sobre datos ya minados (a la selección de datos sobre lagunas ya constituidas o a su visualización), sin manifestar las limitaciones de estos datos producto de los tratamientos propios de su elaboración dentro del sistema de las plataformas y de los mecanismos de acceso a ellos por parte de las AMP.

### Ámbito 3: Relacionado con las cualidades y la calidad de los datos

En este ámbito se considera toda expresión referida a las condiciones y características de los datos empleados, en términos de su origen, fuentes, representatividad y relevancia, su integridad, orden y estructura, sus posibles defectos, así como de las dificultades relacionadas con las propiedades distintivas del *big social data* (volumen, velocidad, variedad, veracidad) y las limitaciones que podrían tener tales datos para responder a determinadas preguntas.

Las expresiones halladas en el corpus con referencias a las cualidades de los datos minados por las AMP fueron las siguientes:

F	Nuestro Proceso Desde la captura de datos en bruto hasta la entrega oportuna de información procesada.
G	Capturar la información precisa con base de datos y herramientas de búsqueda avanzada y datos de alta calidad.
H	Queremos que cada uno de nuestros trabajos se distingan por la exactitud y prolijidad. Somos agentes de la precisión, empedernidos en lograr los objetivos propuestos y perfeccionar nuestros procesos.
I	API Twitter Herramienta para descargar y analizar menciones en twitter. Nos permite visualizar, analizar y medir las tendencias, conexiones y las menciones en torno a conceptos claves.
J	Disponer de múltiples fuentes de datos digitales complementarias y relevantes para el día a día, a través de nuestra tecnología de captura digital, tales como Influencers (políticos, líderes de opinión, celebridades, etc), ONG's, Sindicatos y Movimientos Sociales Organizados.
K	Conecta tus Dashboards o bases de datos ante el desafío de consolidar data de múltiples fuentes.
L	Datos Sociales: Se denomina a cualquier contenido de acceso público que esté disponible en un ambiente digital.

Pese a ser el ámbito con mayor representación, los segmentos identificados no reflejan con claridad ni procesos ni desafíos relacionados con el tratamiento de los datos o sus particularidades. Las referencias a los procedimientos de descubrimiento y preparación de los datos son sumamente vagas (F, G, H, I) u omiten de forma intencional que los datos ya están refinados por diversas tecnologías antes del acceso a estos por parte de las AMP (F).

En cuanto a las fuentes de los datos, en las raras ocasiones en que se habla explícitamente de las fuentes (I, J, K), se hace en términos extremadamente generales, salvo por el ejemplo L, en que se apunta con un poco de claridad aunque de forma muy escueta a una característica central del origen de esos datos. De esta forma, lo más común que los datos de los que disponen las AMP parecen provenir simplemente del “mundo digital”, del “ecosistema digital” o de las redes sociales. En la única ocasión dentro del corpus global en que se hace referencia a la API de una plataforma (I), la explicación es insuficiente (¿a cuál de las API de X/Twitter nos referimos? ¿Por qué se escogió? ¿Qué limitaciones impone esta API en particular al registro de datos? Etc.).

#### **Ámbito 4: Relacionado con los valores de la investigación**

Aquí se incorporan todas las expresiones y segmentos significativos cuyo objetivo es informar sobre los componentes éticos y sociales de la investigación: normas y leyes nacionales o internacionales, estándares propios de cada compañía o del sector en general, códigos de conducta, políticas de transparencia, resguardos relativos a la privacidad de los clientes, usuarios y las personas detrás de los datos, etc.

Una faceta muy relevante dentro de este ámbito es indagar si las AMP toman algún resguardo o advierten sobre los posibles sesgos de las fuentes, del sistema de las plataformas y de los propios los investigadores e investigadoras.

El resultado de nuestro análisis de concordancia fue que al menos en los contextos de las *keywords* seleccionadas no logramos identificar expresiones que expusieran o al menos sugirieran preocupación por alguno de estos problemas de forma clara. En un caso se alude de manera vaga

a “estándares” y “buenas prácticas” (M), pero dado que se trata de una empresa con múltiples negocios más allá de sus posibles servicios vinculados a la minería de datos, es difícil sacar conclusiones de esa única expresión.

Por otro lado, si bien dos de las compañías contaban con apartados titulados “Código de Ética” en sus sitios corporativos, en uno de los casos no había ninguna posibilidad de asociar las precisiones que ahí se hacían a algún elemento relacionado con la minería de datos comercial (el texto se conectaba completamente con la faceta de *lobbying* de dicha empresa). El segundo caso (N), es más interpretable, sobre todo porque consta solo de 24 palabras, la mayoría de ellas demasiado genéricas para pretender apuntar a algo preciso.

M	Agencias socias y filiales trabajamos colaborativamente, compartiendo visión global, altos estándares de calidad, buenas prácticas y herramientas tecnológicas, para desarrollar estrategias de valor y apoyar los objetivos de nuestros clientes con un sello característico alrededor del planeta.
N	Transparencia, probidad, confianza, compromiso y colaboración son los lineamientos transversales que definen la cultura interna y nuestra relación con cada uno de nuestros stakeholders.



## DISCUSIÓN

Nuestro análisis muestra cómo, a través de sus sitios web corporativos, las agencias de monitoreo de plataformas chilenas proyectan las principales características de su sector mediante discursos que dan cuenta de lo que podríamos describir como una “amplia oferta” —al menos en términos simbólicos— de mercancías analíticas. Los discursos analizados aquí nos permiten distinguir las cualidades que las AMP consideran relevantes de sus mercancías, así como también aquello que estiman es del interés de su clientela. Dado que casi el 70 % de las empresas incluidas en este estudio (considerando solo las 15 que produjeron unidades significativas para el ACD) declara llevar más de una década en el negocio, podríamos aducir que aquello que estiman de interés para sus clientes no está basado en una mera suposición, sino que en su conocimiento del mercado.

En términos globales, podemos señalar que la oferta de las AMP se divide en dos grandes grupos. Por un lado, las expresiones que describen directamente productos o servicios de monitoreo; por otro, aquellas que ofrecen la imagen de una relación cliente-proveedor particular, caracterizada por la experiencia, multidisciplinariedad e innovación tecnológica que la empresa asegura poseer, pero también por la plasticidad de las configuraciones precisas de dicha relación, la atención constante a las necesidades de la contraparte y la connivencia.

Dentro del primer grupo, las formulaciones se caracterizaron por ofrecer cuatro énfasis principales:

- 1) Descripciones de servicios, productos o “soluciones” propiamente tales que el cliente puede adquirir, usando expresiones como “la plataforma más avanzada”, “nuestra tecnología”, “nuestras plataformas”, “el servicio está enfocado”, “contratar esta cobertura permite”, “es la plataforma ideal para”, entre otras.
- 2) Descripciones de acciones que puede llevar a cabo la empresa para el cliente o a nombre de este, con expresiones como “te mantenemos informado”, “ponemos a tu disposición”,

“queremos convertir TUS datos sociales en Información”, “te invitamos a sacar partido del potencial”, entre otras.

- 3) Descripciones de acciones que puede llevar a cabo el cliente mismo tras adquirir los productos de la compañía. En estos casos, las expresiones tomaban formas como “mantente al tanto de todo”, “monitorea”, “mapea el contexto”, “analiza”, “usa nuestra plataforma para”, “demuestra con datos”, “vas a poder visualizar”, “vas a poder apoyarte en datos”, entre otras.
- 4) Acciones generales de actividades desarrolladas por la empresa, sin referencia inmediata a su relación con el cliente, con frases como “escuchamos las redes sociales”, “implementamos”, “monitoreamos”, “analizamos”, “transformamos los datos”, “generamos informes”, etc. En este grupo también se encuentran algunas unidades que expresan valores relevantes para el negocio, como “creemos que las comunicaciones pueden y deben ser monitoreadas”, “nos gusta que las métricas nos hablen”, “tenemos una obsesión por las métricas”, entre otras.

En el segundo gran grupo de expresiones promocionales, cuyo foco es la relación cliente-agencia como mercancía, conformado principalmente por las categorías 4 (empresa exitosa y pionera) y 5 (disposición a cumplir cualquier expectativa del cliente) de nuestro análisis, se destaca el empleo de expresiones relacionadas con la confianza (“somos el mejor complemento”, “equipo multidisciplinario con expertise”, “déjalo en manos de especialistas”, etc.), la ayuda (“nuestro equipo está preparado para ayudarte”, “trabajamos junto con nuestros clientes”), la atención y el cuidado constante (“disponibles siempre”, “acompañamiento continuo de nuestro equipo”, “un equipo siempre dispuesto”, etc.), cierto grado de sumisión (“del resto nos encargamos nosotros”, “discutimos poco”) y la explicitación de la ductilidad inherente en la relación (“servicio personalizado”, “100% personalizados”, “pensados a la medida”, “un traje a la medida”, “para potenciar tu organización”).

Detectamos un tercer grupo más acotado que ofrece características particulares, representado por la categoría 3 (enfrentar un nuevo, complejo y desafiante escenario). Es un caso especial, ya que se trata de expresiones que, para sugerir la necesidad de un producto, informan de un contexto. En este caso se usan formulaciones como “en los últimos años”, “en un mundo conectado”, “en un mundo cada vez más segmentado”, “hoy, las redes sociales”, “las redes sociales son actualmente”, “en la Era de la información”, “el mundo cambia rápidamente”, “NO SE QUEDE FUERA”, entre otras.

Lo que nos muestra esta amplia variedad de formulaciones es que lo que comercia el sector AMP es un producto o serie de productos altamente plásticos, en el sentido de no presentar configuraciones fijas en cuanto a su materialidad, soporte o temporalidad, sino que cualquier composición que favorezca los propósitos del cliente. Si bien la categoría 5 de nuestro análisis (el producto como una empresa dispuesta a cumplir cualquier expectativa de sus clientes) lo expone de forma más explícita, la ambigüedad generalizada del discurso promocional de las AMP sitúa este carácter como cualidad focal de la oferta comercial del sector.

Aunque este no pretende ser en modo alguno un análisis cuantitativo, los énfasis discursivos que denota el peso relativo de las diferentes categorías descritas en nuestro análisis sin duda aportan indicios útiles para la evaluación crítica de la oferta de mercancías analíticas que caracteriza a las empresas de este sector en Chile.

En los siguientes apartados discutiremos los elementos que consideramos primordiales para responder nuestra primera pregunta de investigación, a saber: ¿Qué componentes ideológicos caracterizan a la oferta de mercancías o servicios que ofrece el sector de las AMP en Chile?

## Volumen, variedad y velocidad

Las dos categorías más representadas en los sitios web de AMP chilenas fueron “el producto concede una ventaja por su velocidad o poder predictivo” y “el producto concede una ventaja por la amplitud, profundidad o completitud de la información que abarca”. Se trata de dos aspectos íntimamente ligados y omnipresentes en la literatura relacionada con la minería de datos de plataformas. No por nada velocidad (categoría 1), volumen y variedad (categoría 2) son las tres primeras V con las que habitualmente se caracteriza al *big data* (Chen, 2018).

Las expresiones correspondientes a ambas categorías conforman casi un tercio del total de las unidades identificadas y analizadas en el corpus. Aunque en el corpus internacional también se destacan, en ese caso no alcanzaron a constituir un quinto del total.

¿Cómo interpretar la preeminencia en el discurso de los mensajes relacionados con la velocidad (principalmente, la rapidez para recabar los datos), el tamaño de los conjuntos de datos registrados y la variedad de estos? Este énfasis no es una tendencia reciente en el sector. Ya en 2010 lo indicaba con claridad el fundador de una de estas compañías, cuando explicaba que su ventaja respecto a las encuestas era “entregar resultados mucho más rápido” y “con muestras” de mucho mayor tamaño (BioBioChile, 2011). Se ponía de manifiesto entonces la necesidad que tiene un nuevo negocio de interpelar a un potencial cliente poniendo sus mercancías en relación con una oferta ya existente (las encuestas, los estudios de opinión), aunque se trate en este caso de productos radicalmente distintos en términos de técnicas, métodos y racionalidad subyacente. Pero, si la competencia son las encuestas tradicionales, la velocidad y el tamaño de las “muestras” son sin duda elementos centrales a destacar.

Ahora bien, después de varios lustros, la centralidad que sigue teniendo volumen, variedad y velocidad en la oferta de las AMP no puede deberse exclusivamente a su posición

relativa en el mercado respecto a las encuestadoras. Una posible respuesta alternativa, consistente con las observaciones de Kennedy (2016), es que frente a un recurso fundamentalmente defectuoso como son los datos de las plataformas sociales, caracterizados por ser desestructurados, confusos, metodológicamente defectuosos, volátiles y con frecuencia imprecisos e irrelevantes (Kennedy, 2016; Laaksonen y Pääkkönen, 2020), el sector AMP pone de relieve aquellas cualidades más incontestables, el tamaño y velocidad, aunque poco tengan que ver con otro de los aspectos que emergen dentro de esta categoría: la promesa de una cierta capacidad predictiva. Esta cualidad, sugerida algunas veces entre líneas y en otras ocasiones expuesta con total transparencia, está atada a la ideología del positivismo digital, así como a la del capitalismo de vigilancia, que según Zuboff se funda en la confianza de que la digitalización de la experiencia humana permite en última instancia la fabricación de “productos predictivos” (Zuboff, 2018).

## **Un escenario ominoso**

Una categoría emergente que se erigió como de gran relevancia durante el análisis, pese a que no se explicitaba como algo central del negocio de las AMP en la literatura consultada, fue la conformada por aquellas expresiones que exponían la existencia de un mundo nuevo, complejo y desafiante al que no era posible resistirse, pero sí posible enfrentar en mejores condiciones gracias a ciertas mercancías analíticas.

Ciertamente, las expresiones clasificadas en este grupo no son las únicas que hacen referencia a estas nuevas y extrañas condiciones. Esta es una idea que recorre todo el corpus. Aparece en la urgencia que empuja a necesitar “monitoreo en tiempo real”, “equipos que trabajan 24/7”, “predecir futuras conversaciones”, acceder a grandes volúmenes de datos variados pese a que son inmanejables, etc. La misma idea está presente en el impulso a adoptar una nueva forma de “ver” (categoría 7) y “accionar” el mundo (categoría 6) y por “obsesionarse” con las métricas (categoría 8). Algo ominoso está ocurriendo “allá afuera”, un “desafío titánico” que afecta con una rapidez y fuerza inusitada “todos los aspectos de la vida” y de los negocios.

Pero en ningún caso se trata de una conceptualización de la realidad exclusiva del contexto de esta investigación. Los tropos usados por las AMP son un reflejo de conceptos respecto a lo que es constitutivo del momento presente que son habituales en la prensa y en todo discurso sobre el capitalismo contemporáneo orientado hacia lo digital. Todas las elucubraciones actuales sobre el futuro parecen fundarse en un choque entre el tecnooptimismo que caracterizó los albores del Internet capitalista y tecnopesimismo que emergió tras la constatación de su ruptura. En este claroscuro generador de monstruos que se camuflan tras una sobrecarga de información, diferentes actores en distintos mercados buscan distinguirse con la promesa de estar capacitados para “abrirse paso a través del desorden” (Andrejevic, 2013) y con ello “proporcionar una comprensión incomparable del mundo” (Alto Analytics, s.f.).

Las expresiones que enuncian este ineludible estado de las cosas que obliga, frente al terror de lo desconocido, a conocer todo lo que se dice en todo momento, apuntan a una competencia sin tregua entre los hablantes y a un vacío o redistribución inapropiada del poder. Poder hoy fundado sobre el control de ciertos datos, que debe ser retornado a las manos correctas. Dadas esas condiciones, lo único capaz de diferenciar a los miembros de la élite (y a los profesionales que esta pueda pagar para sostener sus negocios u organizaciones) del resto de los seres humanos —en el contexto de las plataformas sociales de Internet— es la capacidad que tienen los primeros de acceder a los productos y las tecnologías que ofrecen las AMP.

## **De la recuperación de los medios de producción espiritual**

Así, por un lado tenemos que, según los discursos promocionales de las AMP, la única o la mejor forma de enfrentar las actuales condiciones comunicacionales impuestas por el despliegue global del sistema de las plataformas son los productos ofrecidos por las mismas AMP. Sin embargo, por otro lado, no es posible argüir que las virtudes de dichas mercancías analíticas estén fundadas en los elementos que las AMP decididamente enfatizan en sus discursos. La retórica en torno a la velocidad, variedad y volumen de los datos opera aquí más

bien como una máscara promocional para ensombrecer las carencias e insuficiencias en términos de “referencialidad” (Andrejevic, 2013) que tienen los datos de las redes sociales, describiendo no tanto el mundo más allá de ellos como los propios patrones de relación entre los datos. Y poniendo de relieve el “monitoreo mecanizado instantáneo y continuo de los flujos agregados” (Andrejevic, 2011) sobre cualquier otro estándar de comprensión de los fenómenos.

Considerando estos factores, ¿en qué consisten, en definitiva, los productos ofrecidos por las AMP y en qué aspectos descansaría su presunta eficacia?

La categoría 5 de nuestro análisis, el producto como una empresa dispuesta a cumplir cualquier expectativa de sus clientes, tiene un peso relativo en el corpus de aproximadamente un 10 %. Junto a la categoría 4, íntimamente relacionada, conforman aproximadamente un quinto del total de expresiones identificadas. Como hemos indicado, estas dos categorías temáticas ponen su foco en la relación cliente-agencia antes que en las cualidades técnicas o la propia eficacia de los productos y servicios del sector AMP.

En nuestro marco teórico hemos propuesto que para poder caracterizar el sector de las agencias de monitoreo de plataformas no podemos pasar por alto la comprensión del sistema de las plataformas, de las cuales las AMP son un eslabón, ni tampoco podemos ignorar el extremo opuesto de esta cadena de valor: la clientela de estas compañías.

Basándonos en los hallazgos de las investigadoras Eileen Meehan respecto al mercado de los ratings televisivos (Meehan, 1984) y Helen Kennedy con relación a un conjunto de compañías de minería de datos, junto a algunas inferencias que provienen del ejemplo paradigmático que hemos tocado en la introducción (el informe Big Data), propusimos que el papel del bloque social que adquiere las mercancías analíticas de las AMP sea fundamental para caracterizar al sector y sus productos.

El conjunto de consumidores con acceso a pagar por las mercancías elaboradas por las AMP chilena —una cierta élite local— no solo se distingue a sí mismo como clase, sino que —

siguiendo las reflexiones de Mills (Mills, 2013)— desarrolla acciones colectivas dirigidas a la conservación de sus privilegios. Es así que, frente a escenarios de incertidumbre creciente, particularmente en términos comunicacionales, donde entre otras transformaciones “surge información minuto a minuto” (ejemplo 3.1) y se altera de forma radical la proporción entre quienes hablan y quienes oyen, estos actores requieren herramientas para recuperar “el control sobre la narrativa” (ejemplo 1.3). El mercado de las AMP es, bajo esta racionalidad, uno que ofrece ante todo los medios para la recuperación de una seguridad ontológica perdida o en juego.

Pero para que una clase dominante vuelva a tener el control sobre lo que Marx y Engels llamaron “los medios para la producción espiritual”, dicha clase no puede someterse llanamente al peso de la realidad. Incluso si la realidad les entregara la razón, el riesgo sería demasiado elevado. Las élites deben, por lo tanto, imponer mediante diferentes estrategias sus teorías sobre lo que es relevante en cada momento. Determinadas élites de carácter global lo hacen a través de la banca, la vigilancia o mediante el sistema de las plataformas. Pero, ¿qué ocurre con las élites locales al margen de un sistema mediático mundializado? Estas requieren mecanismos legitimados socialmente no solo para difundir sus ideas, en general ya hegemónicas, sino y sobre todo para demostrar que sus creencias son las ideas de la razón. Si bien una parte de este trabajo lo llevan a cabo instituciones como los medios de comunicación y las universidades, así como el Estado, otra porción queda en manos organizaciones menos asentadas, como *think tanks* y una profusión de consultoras y agencias de relaciones públicas y lobby, y también de monitoreo de plataformas sociales.

En esta línea de pensamiento, lo que se puede poner en duda del papel social de las AMP, tras el análisis crítico de los discursos que ofrecen, es cualquier pretensión de objetividad, así como la esperanza de que los datos hablen por sí mismos desde la neutralidad. Lo que nos muestran los ejemplos de la categoría 5 es una contradicción clara entre un producto que se ofrece como una aproximación científica a la comprensión de un fenómeno y a la vez pone énfasis en la personalización y plasticidad de los resultados. Las percepciones internas del sector, reportadas por Kennedy (2016), que aquí hemos descrito como un autosometimiento agnotológico, se manifiestan en esta categoría.



Kennedy reportaba que los consumidores de las mercancías analíticas parecían indiferentes a la calidad de los datos y a los límites de las herramientas. A partir de nuestro propio análisis, desde la perspectiva de los discursos que caracterizan la oferta de tales mercancías, pareciera enunciarse un espacio seguro, donde esa indiferencia o sometimiento a la ignorancia deliberada está garantizado, se manejará con confidencialidad y será reconducido para “sacar el máximo provecho” (ejemplo 5.1) y realizarse “a la medida de las necesidades de su negocio” (ejemplo 5.a). En expresiones como la del ejemplo 5.2 (“convierte tus análisis cualitativos en analítica, gracias a los procesos de Machine Learning (...) 100% personalizados a las necesidades de tu negocio”), se mezcla un énfasis en el poder o el atractivo de los números (categoría 9) con la garantía de que los números resultantes no operarán por sí mismos ni mostrarán una realidad cualquiera opuesta a los intereses del cliente, sino que reproducirán con precisión aquello que el cliente busca, de acuerdo a las necesidades de su negocio. En suma, lo que cuenta como conocimiento queda en manos de quien tiene el poder para adquirirlo.

Podríamos pensar que el sector AMP saca provecho de los sesgos y prejuicios de sus clientes y les ofrecen ‘exactamente’ lo que quieren escuchar. Pero esto sería impreciso e indulgente con una élite que aparentemente sabe perfectamente lo que necesita de estas empresas. Al menos, es lo que nos muestran tanto las investigaciones de Mills y el caso estudiado por Meehan, como las observaciones de Kennedy y nuestro propio análisis en ocasión del informe Big Data en Chile. Por ello, podemos pensar que lo que ocurre es la situación opuesta: es la élite la que saca provecho de un sector dispuesto, en contra de sus propios fundamentos presuntos, a manipular lo manipulable (y como señalan los especialistas, no hay nada más manipulable que los datos de las redes sociales) en favor de ella.

## **La relevancia de lo visual**

Las falencias de las mercancías analíticas en lo concerniente a su precisión (que las interpretaciones de los datos puedan ser efectivamente tomadas como indicios de fenómenos) no

son una negación de lo factual o de la relevancia de actuar conforme a los datos por parte de las AMP (en connivencia con sus clientes), sino que, ante todo, manifiestan una disposición a considerar irrelevante o de segundo orden cualquier conocimiento referencial y a elevar la importancia de lo correlacional en pro de algo que consideran una verdad mucho más profunda: la propia razón (la razón de la élite). La pretensión comunicativa detrás del empleo de estas vías no es la mentira, que en sí misma es tan irrelevante como la verdad, sino que el resultado se adapte a sus fines, tal como señala Harry Frankfurt al referirse al charlatán:

Los valores veritativos de sus enunciados no tienen prácticamente interés para él; de lo que no hemos de darnos cuenta es de que su intención no es informar de la verdad ni tampoco ocultarla. Eso no significa que su discurso sea anárquicamente impulsivo, sino que la motivación que lo guía y lo controla prescinde de cómo son realmente las cosas de las que habla. (Frankfurt, 2006, pp. 67-68)

En este sentido, es muy interesante la importancia que tuvo dentro del corpus la categoría que describe a las mercancías analíticas como un producto cuya cualidad primordial es la de permitir visualizar los datos. Dado que la visualización de datos privilegia la correlación sobre la causación (Rettberg, 2020), su poder persuasivo puede suplir debilidades epistemológicas de los procedimientos y, como hemos indicado más arriba, potenciar la capacidad instrumental de las mercancías analíticas.

Este es un aspecto del que se ha escrito en años recientes con algún detalle (p. ej. Engebretsen y Kennedy, 2020; Coopmans, Vertesi, Lynch y Woolgar, 2014) y que pese a que hemos tocado solo marginalmente en este trabajo, sin duda juega un papel importante en la caracterización del sector AMP y otras áreas que lidian con datos digitales y *big data*. En estos contextos, la visualización de datos permite aumentar el valor comercial de estos en formas concretas, al transformar datos heterogéneos en objetos de autoridad, en información con capacidad de producir efectos y resultados empresariales (Laaksonen y Pääkkönen, 2020).

En lo que respecta a los discursos que hemos analizado, aseveraciones como que las visualizaciones de datos no son ni observaciones directas ni representaciones inocentes o neutrales de hechos (Nærland, 2020) quedan respaldadas en algunas de las expresiones que incluimos en esa categoría. Así, por ejemplo, afirmar que se mostrará con datos una gestión, discurso que expresaría una comprensión más neutral del efecto de la visualización de datos, menos sesgada u orientada hacia un fin particular, no es lo mismo que demostrar con datos “el éxito de tu gestión” (ejemplo 5 de la categoría 7), que ofrece de antemano un resultado esperado. Dado que el entendimiento generalizado es que la visualización de datos es una potente herramienta de persuasión, resultan significativos estos direccionamientos discursivos con relación a la auténtica utilidad de las mercancías ofrecidas.

## **El deseo por los símbolos legitimados de conocimiento**

En oposición a la hipótesis planteada por Kennedy (2016), aunque sin invalidar sus observaciones, la décima categoría de este análisis crítico de discurso, el producto de las AMP como algo que entrega control sobre el poder y el atractivo de los números y métricas, no pareciera, en ninguno de los corpus (nacional e internacional), ubicarse en un lugar destacado. Ello sería lo esperable si, como Kennedy afirma, las relaciones sociales de producción imperantes en el contexto de la minería de datos comercial estuvieran orientadas por un fetichismo explícito relacionado con la producción y distribución de número y métricas (independiente de su precisión).

Nuestro análisis muestra que, al menos en lo que respecta a su expresión en los discursos de las propias empresas del sector, la preponderancia de estos elementos es más bien modesta en comparación con otros factores. A la luz de los resultados del presente análisis, podríamos pensar que aquello que la autora identifica como un “deseo por los números” no es tanto una valoración de los métodos y técnicas que sostienen lo que se ha llamado la datificación de la sociedad como, en primer lugar, un deseo por poseer (y manipular) los símbolos de aquello que es socialmente tenido como “evidencia”, como señas de conocimiento. En otras palabras, lo que produciría y

vendería el sector AMP es el certificado de “tener la razón” que sus clientes requieren. Eso puede ser un número o una métrica, como observa Kennedy, pero podría ser cualquier factor legitimado que otorgara tal certificación social. Es por ello que lo esencial, en términos discursivos, de la oferta de las AMP es la explicitación de que sus productos y servicios aseguran una razón a la medida de las necesidades comunicacionales de sus clientes, con independencia de su relación con la realidad o la verdad. Y que ni siquiera los números interferirán en la producción de esos significados que operan como “sustitutos de la verdad”.

Para Kennedy es justamente la fuerza de ese deseo por los números lo que suprime la conversación entre agencias y clientes sobre las limitaciones de la minería de datos y determina, asimismo, una falta de interés en la obtención de “números precisos” (p. 225). Aquí diferimos y, como hemos señalado más arriba, dado que la mercancía analítica primordial de las AMP, según nuestro análisis, lejos de revelar un énfasis en las métricas lo que hace es mostrar garantías de una disposición a otorgar la razón sobre cualquier materia, es la propia naturaleza del vínculo agencia-cliente lo que determina relaciones sociales de producción agnotológicas que suprimen cualquier imperativo de precisión científica.

Lo que aquí se señala no pretende desestimar las afirmaciones relativas a la metrificación de la existencia o los esfuerzos por entender la intensificación actual de la cuantificación en los debates en torno al *big data* y sus implicaciones sociales (Beer, 2016), ni menos descartar las relevantes observaciones respecto al uso de las cifras con fines meramente performativos —en busca de transmitir una ilusoria sensación de objetividad— que empresas como Facebook llevan a cabo (John, 2019). A lo que se apunta es a la simple constatación de que en la oferta de las AMP, el énfasis está puesto en la ductilidad de sus productos, que bien pueden tomar la forma de números como de cualquier otra cosa que funcione para ciertos fines.

No obstante, el deseo de los números está íntimamente relacionado con la capacidad de mostrar esos números como visualizaciones de datos, por lo que estos dos aspectos están íntimamente relacionados y ulteriores análisis podrían permitir interpretar una mayor preponderancia de las “métricas” en la oferta efectiva. Dado que nuestra investigación es una

primera aproximación a estas problemáticas y está limitada por las herramientas metodológicas escogidas, las conclusiones presentadas son desde luego preliminares.

## **Sobre los discursos en torno a los desafíos técnicos y metodológicos asociados al sector**

En esta sección nos haremos cargo de nuestra segunda pregunta de investigación: ¿cómo responde o se hace cargo el sector de las AMP chilenas a las diferentes problemáticas relacionadas con el sistema de las plataformas y los procedimientos de la minería de datos?

Como vimos en el apartado correspondiente del capítulo anterior, en el que describimos los resultados de nuestra aproximación a los discursos relacionados a los desafíos del sector, la discusión técnica y metodológica de las implicancias de trabajar y lidiar con los datos obtenidos a través de las API de las redes sociales es casi por completo ignorada en el corpus de sitios corporativos chilenos. Lo mismo pudimos observar en lo referente al ámbito de los valores implicados en la investigación, donde consideramos cualquier discusión sobre los preceptos socialmente consensuados (leyes, estándares, códigos o políticas) que guían la producción de conocimiento, así como los resguardos sobre al trato con datos asociados a seres humanos. Por cierto, la representación de estos aspectos tampoco resultó relevante en los sitios de empresas extranjeras escrutados.

A diferencia del tipo de reflexión que se manifiesta en la literatura técnica y académica, problemáticas cruciales tales como los factores éticos de este tipo de aproximaciones, el efecto de las diferentes API en las investigaciones, los mecanismos usados para inferir data cuando esta es de mala calidad o los desafíos en términos de modelamiento algorítmico necesarios para hacer frente a la variedad de configuraciones que toman los datos, son dejadas en lo fundamental de lado y no se explicitan al momento de desarrollar una oferta de productos y servicios.

Asimismo, no solo las debilidades metodológicas y las dificultades éticas son silenciadas, sino que los propios procedimientos son tratados en el mejor de los casos con extrema superficialidad. Al menos en sus sitios web corporativos, las AMP chilenas no consideran relevante indicar con algún detalle en qué consisten sus enfoques, más allá de los énfasis persuasivos que hemos analizado en los apartados precedentes.

Estos hallazgos resultan consistentes con los testimonios de trabajadores y trabajadoras del sector recopilados por Kennedy entre 2012 y 2013. Si bien la autora revela interesantes reflexiones personales de sus entrevistados en torno a las limitaciones metodológicas y epistemológicas de los productos de las AMP y respecto a las dimensiones éticas y de transparencia del campo, estas quedan con frecuencia eclipsadas por los imperativos comerciales de las empresas y el afán por cifras y métricas de sus clientes, lo que dificulta cualquier comunicación rigurosa sobre estos problemáticos ámbitos (Kennedy, 2016).

## CONCLUSIONES

Hemos establecido la posibilidad de que las AMP jueguen un rol en un proceso reaccionario contra la presunta “democratización” de las comunicaciones que anunciaba la irrupción global de Internet como medio de comunicación dominante. Este nuevo escenario prometía, en teoría, un plano en que las masas características del siglo XX podían devenir fácilmente en públicos. Las dinámicas que con frecuencia se han considerado constitutivas de las plataformas sociales digitales coinciden con las dimensiones que, para Wright Mills (2013), distinguen a una comunidad de públicos de una sociedad de masas: mayor equilibrio en la proporción entre hablantes y oyentes; posibilidades efectivas de impugnar a los hablantes; conexión entre la formación de opiniones y la acción social, e incremento del grado de autonomía frente a la autoridad institucional. Proponemos aquí una dimensión adicional: mayor equidad en la distribución de los símbolos socialmente legitimados de (tener) la razón.

De cara a estas transformaciones, diversas élites se encontrarían de pronto en iguales circunstancias comunicativas que todo el resto de los humanos, sometidas a que sus opiniones fueran impugnadas por quienquiera y expuestas a que en contextos fuera de su control se formaran opiniones susceptibles de convertirse en acciones sociales también fuera de su control. Este inédito marco permitía imaginar nuevos espacios desregulados, horizontales y que por su naturaleza inmaterial estarían incluso legalmente fuera del control de cualquier Estado. Pero aún más: se vaticinaba un mundo donde los símbolos socialmente legitimados de tener la razón (el conocimiento, la ciencia, los datos, etc.) se distribuirían igualitariamente. Sin importar qué tan realista fuera dicha utopía, resultaba evidentemente inadmisibile. Era inevitable que emergieran esfuerzos e instrumentos para devolverle al espacio comunicacional socialmente preponderante su eficacia como instrumento de dominio.

El sistema de las plataformas, que se estableció de facto como el guardián de las comunicaciones a través de la red y que para la enorme mayoría de los seres humanos es la única forma de acceder a Internet, funciona como un complejo filtro que en diferentes etapas y con diversos fines comerciales, como hemos visto, no solo refina los datos, sino que los modela y usa

para imponer teorías sobre lo que es y no es relevante. Sus algoritmos de entrada, que se ofrecen como sublimes mediadores desintencionados, co-crean los datos de la experiencia humana digitalizada y, mediante procesos de selección y metadatificación, producen potencialmente infinitos dobles de datos e imponen estrictas normas al comportamiento mediado por las plataformas. Determinan, en última instancia, quién dice qué, a cuántos, por cuánto tiempo y dónde; quién puede responder y cómo, y cuánta interferencia del Estado y otros poderes institucionalizados es o no es admisible. Puede parecer que en algunos momentos la información circula de forma libre y prácticamente desintermediada, pero la voluntad de decidir sus modulaciones y límites siempre está depositada en las manos de algunos.

Luego están las API, en toda su variedad direccionada a diversos fines por diferentes intereses contingentes. Estos instrumentos, nuevamente con una promesa imposible de desintermediación, imponen “vías implícitas” a aquellos y aquellas que acceden con entusiasmo a la riqueza de datos que derraman de forma supuestamente neutral y desinteresada. Esas vías implícitas quedan impresas en la realidad que se construye a partir de tales datos por diferentes procedimientos. El menos problemático de estos procesos probablemente sea el que académicas y académicos llevan a cabo cuando acceden a las API y trabajan sobre sus datos tomando, idealmente, las diversas precauciones a las que disciplinariamente se ven obligadas y obligados. Mucho menos inocua es la minería de datos comercial que hemos tratado en este estudio, donde en apariencia se dan por sentados muchos de los postulados básicos del positivismo digital que entiende a los datos y el *big social data* como portadores de una verdad sublime y a las prácticas relacionadas con ellos como maneras objetivas y privilegiadas de aproximarse a los fenómenos de la sociedad.

Entre los peligros de esta ideología se han mencionado la marginación de otras formas de conocer, la sobresimplificación de problemas complejos y la propagación deliberada de productos que no solo son referencialmente imprecisos, sino instrumentos agnotológicos. Mediante estas mercancías, las empresas dedicadas a la minería de datos comercial estarían proveyendo a una cierta élite local una seguridad ontológica perdida relacionada con la posesión de los símbolos y señas socialmente legitimadas de la razón.



De ser este el caso, dicha promesa debería estar reflejada en los discursos que caracterizan a las mercancías analíticas ofertadas por las AMP. De no serlo, los énfasis discursivos de esta oferta debería ilustrarnos respecto a otros atributos peculiares del sector y sus relaciones sociales y comerciales, como el énfasis propuesto por Kennedy (2016) en el “deseo por los números” u otros elementos vernaculares que aún no hubiésemos detectado.

En este último capítulo, sintetizaremos los principales hallazgos del análisis que hemos llevado a cabo. Para comenzar, acotaremos la definición del objeto de nuestro estudio y nos referiremos al primer objetivo específico establecido, el mapeo del sector de las AMP en Chile. Luego, condensando los elementos centrales de la discusión y de acuerdo a nuestro ACD, abordaremos de forma condensada los resultados correspondientes a nuestro segundo objetivo específico, la evaluación crítica de las mercancías analíticas de las AMP. En la tercera sección de este capítulo trazaremos finalmente las principales características del sector en Chile.

El apartado restante consta de una breve pero necesaria reflexión sobre la relevancia de este trabajo en el momento presente del sistema de las plataformas y sus rasgos distintivos al cierre del primer cuarto del siglo XXI.

Como colofón, hemos agregado una anotación sobre lo que, desde nuestro punto de vista, es la contribución metodológica que ofrece esta tesis para los estudios en el campo de la comunicación política en Chile.

## **Definición y mapeo de las AMP**

El punto de partida de nuestra investigación fue evidenciar el estado del arte de la cuestión. Bajo la apreciación esencial de que el sector de las agencias de monitoreo de plataformas sociales, especialmente en el contexto latinoamericano, ha recibido escasa atención en la literatura de ciencias sociales y las comunicaciones, así como un desinterés casi total de la

prensa, un paso ineludible de una aproximación al problema era afrontarlo mediante una perspectiva en primer lugar exploratorio-descriptiva. Esto nos permitió proveernos de las definiciones operativas básicas para hablar sobre las AMP y articular las preguntas atinentes.

La principal de estas definiciones operativas fue, desde luego, la de aquello que constituye una empresa del sector de las AMP. Ya las hemos caracterizado *in extenso* en las páginas precedentes, pero aquí proponemos el uso de la siguiente descripción en breve.

Las agencias de monitoreo de plataformas sociales son empresas que intermedian datos entre las plataformas sociales de Internet y terceros clientes de pago. El procedimiento técnico para acceder a dichos datos se realiza fundamentalmente a través de las interfaces de programación de aplicaciones (API) de las plataformas. La aplicación de estas técnicas puede quedar en manos de las propias empresas o puede ser un servicio tercerizado en otras empresas especializadas. Con estos datos, las AMP producen una variedad de mercancías, comúnmente informes y presentaciones en diversos soportes con un importante componente visual. Estas mercancías permiten a los clientes de las AMP justificar o respaldar la toma de decisiones corporativas, políticas o comunicacionales.

Esta conceptualización nos permitió, asimismo, dar cuenta del primero de nuestros objetivos específicos, a saber, desarrollar un mapa preliminar del sector en Chile. Junto a las prescripciones pertinentes a la metodología empleada para ello, el primer esbozo de ese mapeo se encuentra expuesto en la Tabla 2 (capítulo 4). En la siguiente tabla acotamos este registro solo a las empresas que, tras el análisis de concordancia, arrojaron unidades significativas de discurso. Con ello, identificamos finalmente 15 compañías (Tabla 7) que, de acuerdo a sus discursos promocionales, ofertan públicamente análisis de redes sociales a través de minería de datos con connotaciones comerciales. La tabla incluye la cantidad de años que indica tener cada empresa. Huelga decir que en el caso de consultoras de índole general o dedicadas al *lobbying* (como

Feedback, Azerta, Extend o Imaginación) o de monitoreo de medios (como LitoralPress o GlobalNews Group), que declaran llevar varias décadas en el mercado, la cantidad de años indicada no expresa necesariamente el tiempo que llevan incluyendo monitoreos de redes sociales entre sus servicios.

*Tabla 7. Mapeo de empresas del sector AMP.*

<b>Empresa</b>	<b>Sitio web</b>	<b>Años en el mercado</b>
Analitic	www.analitic.cl	11 años
Artool	www.artool.cl	13 años
Azerta	www.azerta.cl	20+ años
Brandmetric	www.brandmetric.com	10+ años
Conecta Media	conectamedia.cl	-
Edgy	www.edgy.cl	14 años
Extend	www.extend.cl	30 años
Feedback	www.feedback.cl	25 años
Global News Group	www.globalnewsgroup.com	-
Imaginación	www.imaginacion.cl	27 años
LFI	www.lfi.la	12+ años
LitoralPress	www.litoralpress.cl/sitio/mediossociales.html	39 años
Nexos	www.nexos.cl	30 años
Rompecabeza	www.rompecabeza.cl	7 años
Soul	www2.souldigital.cl	-

## **Evaluación crítica de las mercancías analíticas**

En consistencia con nuestro marco teórico, el foco de esta caracterización de las AMP está puesto en sus mercancías y en las relaciones de producción en que estas participan. De ello

emerge el segundo de nuestros objetivos específicos, orientado a explorar y evaluar desde una perspectiva crítica su oferta. En lo fundamental, esto lo hemos hecho mediante el análisis del que hemos dado cuenta en los dos capítulos previos. La elección de las fuentes, los sitios corporativos de las empresas del sector; la construcción de un corpus con dichas fuentes; la aplicación de herramientas informáticas como apoyo a un análisis de concordancia; la reducción de los datos en 11 categorías temáticas, y, finalmente, el análisis crítico de estos discursos, nos permitió determinar las siguientes siete propiedades principales.

- I. **Énfasis en la velocidad, vastedad y variedad de los datos:** las mercancías analíticas que ofertan las AMP son informes y análisis de redes sociales cuyo rasgo comercial más marcado es un énfasis palmario en la rapidez, sobre todo en el sentido de adaptarse al “dinamismo de las redes sociales” y entregar la capacidad de conocer fenómenos simultáneamente a su acaecimiento (en tiempo real, pertinentes, contingentes). La variedad y vastedad de estos datos y sus fuentes, por otra parte, es garantía de su plasticidad, su ductilidad para acoger preguntas y producir respuestas. Salvo en contadas ocasiones, la fuente inmediata y precisa de esos datos (plataformas y API particulares, si los procedimientos de obtención de datos son propios o realizados por terceras empresas, si hay revendedores de datos o *data brokers* involucrados, etc.) parece ser un detalle a resolver con posterioridad, de acuerdo a las necesidades de cada cliente.
  
- II. **Capacidades predictivas implícitas:** en asociación con las características vinculadas a las dimensiones temporales del registro y análisis, se encuentra muchas veces implícita, aunque en algunas ocasiones también de forma explícita, la oferta de una cierta capacidad predictiva. Esta promesa aparece como un aspecto regular en desarrollos contemporáneos del marketing digital, donde es habitual entender las predicciones como “adelantarse a las preferencias del consumidor” bajo las premisas de lo que se conoce como *content-based recommender systems* (Papadakis, 2023), tecnologías que nos muestran “justo lo que queremos” en sitios como Amazon y Netflix. Una aproximación relacionada es la estimación de algunos datos que no

están presentes en el registro (o han sido anonimizados), para conseguir una segmentación más compleja de los consumidores (por ejemplo, adivinar el género o la capacidad adquisitiva de los usuarios). Estos usos “predictivos” se corresponden con el diseño y los objetivos comerciales del sistema de las plataformas, pero se ubican lejos de promesas como la de “predecir futuras conversaciones y crisis comunicacionales” o la de “anteponerse a los intereses y necesidades de las comunidades” que hemos revisado en nuestro análisis de discurso de las AMP, afirmaciones en las que la separación entre los fenómenos en línea y fuera de línea queda desdibujada.

Este límite difuso es un tema frecuentes en el ámbito del *big data* y lo cierto es que se han hecho importantes avances a través de modelos gráficos probabilísticos, análisis estadísticos y de tendencias y uso de diferentes variantes de *machine learning*, en pro de utilizar grandes volúmenes de datos para la predicción de ciertos comportamientos (Dunne, 2023). No obstante, estos modelos están orientados fundamentalmente a ambientes inteligentes (que incorporan muchos sensores y procesamientos computacionales de forma ubicua) y acotados (como entornos hospitalarios y residencias de ancianos). Los modelamientos diseñados exclusivamente a partir de datos de plataformas sociales requieren, por otro lado, aproximaciones de una inusitada invasión de la privacidad de las personas (Botta, 2015; Comito, 2018 ; Fernandez Vilas, 2019). Con todo, ninguno de estos casos parecen ser el tipo de enfoques al que hacen alusión las AMP en sus ofertas, sino más probablemente refieren a estimaciones más comunes y en apariencia más sencillas, como el análisis de sentimiento (Bouazizi y Ohtsuki, 2019) o la detección de bots (Lopez-Joya, Diaz-Garcia, Ruiz y Martin-Bautista, 2023), que sin embargo también han probado ser enormemente problemáticas y poco precisas.

**III. Opacidad respecto a sus aspectos técnicos:** a través de los mensajes promocionales de las AMP en sus sitios web, es difícil, sino imposible, conocer las características técnicas de la manufactura de sus mercancías, los procedimientos y fundamentos de

sus servicios. Aquellas empresas que utilizan alguna clase de lenguaje de carácter técnico (la minoría) lo hacen de una forma vaga y a la vez poco accesible. Por ello, las características de las mercancías que aquí describimos son esencialmente de tipo cualitativo, pese a que el discurso que rodea a estas empresas reitera con insistencia que sus efectos son eminentemente cuantitativos.

- IV. Limitada incidencia directa:** la promesa en términos del potencial prescriptivo de estas mercancías, es decir, de su capacidad de incidir en el mundo, es limitada. La idea de poder generar un efecto en una audiencia, en términos de sus propiedades como construcciones infocomunicacionales, se encuentra solo marginalmente representada. En cambio, los mensajes son mucho más claros en cuanto a que estas mercancías conceden conocimientos “accionables” en materias corporativas y administrativas, para la toma de decisiones gerenciales, en la mejora o transformación de las racionalidades y sentidos al interior de las empresas o para la producción y establecimiento de políticas dentro de ellas. Es decir, el aspecto prescriptivo sobre el mundo está enmarcado a los usos que los poseedores de estos productos podrán darle, primordialmente, al interior de sus propias organizaciones.
- V. Garantías de ductilidad:** una parte importante de los esfuerzos discursivos promocionales de las AMP en torno a sus servicios y productos se concentra en describir y poner de relieve las cualidades de la vinculación que establecen con las y los clientes. Las mercancías encarnan, por esta vía, una relaciones de colaboración, ayuda y presencia permanente junto a la clientela, sus organizaciones y sus jerarquías. De esta forma, la configuración efectiva de las mercancías es “personalizada”, queda supeditada a las “necesidades” del negocio y se produce a la “medida” del cliente. Se explicita una sumisión a las agendas del mandante, así como altos niveles de confianza. La escasez de detalles, la opacidad de los servicios ofrecidos, también sugieren que se apunta a clientes que valoran un cierto grado de confidencialidad, del establecimiento relaciones cara a cara. Es notoria una orientación a dejar los detalles de la colaboración para ser conversados en reuniones

presenciales. Así, es habitual encontrar en lugares destacados de los sitios web de las AMP frases como “elige un socio tecnológico de confianza que te ayude a llegar al siguiente nivel. Agenda una reunión” (LFI) o “si necesitas mayor información sobre nuestros servicios o quieres saber cómo podemos aportar a tu organización, haz click en el botón” (Analytic), mensajes que preceden a cualquier detalle sobre los productos y servicios. En la misma línea, varias páginas presentan testimonios laudatorios de sus clientes (en un caso, hasta 12) o resaltan a quienes “han confiado” en la empresa.

- VI. Componente visual:** El factor visual de las mercancías, en informes, presentaciones y tableros virtuales o *dashboards*, es un aspecto transversal de toda la oferta. Según la imagen que proyecta el sector en sus discursos, este factor parece ser tanto una respuesta como una expresión del dinamismo que caracteriza al mundo actual, que requiere de herramientas de rápida comprensión e inmediata capacidad persuasiva, direccionada a la acción dentro de las organizaciones para la producción de efectos y resultados empresariales.
- VII. Limitado énfasis en los números y métricas:** El foco que tienen estos productos en lo que respecta a las métricas y números, el método científico y la precisión, es limitado —al menos, en cuanto a sus discursos explícitos e implícitos en sus sitios web— y siempre está subordinado a los demás factores que aquí se han mencionado como centrales, particularmente a la plasticidad que requieren las diversas “necesidades” de los negocios y la configuración de productos verdaderamente “accionables” que permitan “demostrar el éxito” de las organizaciones. Se puede decir que, al menos en la oferta explícita, los números se encuentran ausentes, salvo por una vaga promesa en relación con situarse de parte del cliente.

### 7.3. Principales características de las agencias de monitoreo de plataformas en Chile.

En el marco teórico, el importante trabajo de Eileen Meehan sobre la mercancía rating, nos permitió mostrar que las lógicas que parecen predominar en el caso de las AMP y sus productos no son necesariamente nuevas, sino que tiene antecedentes en prácticas asociadas a los medios de comunicación de masas. La concurrencia de estas observaciones junto a las de Helen Kennedy, además de los efectos públicos conocidos del informe Big Data, tratados en detalle en la introducción, arrojaron luz sobre el papel de la relación cliente-agencia en la configuración del mercado de las AMP, lo que se vio respaldado por el análisis crítico de los discursos de estas compañías en Chile.

Con estos elementos a nuestra disposición, para dar cuenta del objetivo general establecido y concluir esta investigación, podemos caracterizar a las AMP chilenas de acuerdo a los siguientes siete atributos.

1. **Las AMP en su conjunto conforman una industria comercial y culturalmente diferenciada.** Las maneras precisas en que los servicios y productos de las AMP consiguen esos objetivos son vagas, pero como oferta se configura un discurso relativamente consistente que se apoya en expresiones comunes y énfasis similares. Esto confirma preliminarmente la existencia de una industria específica y habla de una cierta estabilidad comercial y cultural para el sector. Apoya esta idea el hallazgo de que los discursos tampoco difieren en lo fundamental de aquellos que caracterizan a AMP internacionales, como Alto Data Analytics.
2. **Las AMP ofrecen diferentes enfoques corporativos y forman parte de un mercado dinámico.** No obstante, un análisis pormenorizado empresa por empresa, nos indica que dentro de esa consistencia en los discursos referentes al monitoreo de plataformas sociales, se trata de un sector variado en términos de los enfoques corporativos que presenta. Así,



forman parte del conjunto identificado compañías con un foco de ribetes más bien técnico, con discursos mucho más enfocados en el análisis del *big data*, las plataformas y las tecnologías de la información, y que fuera de toda duda ejecutan, probablemente con sus propios medios técnicos y humanos, minería de datos de tipo comercial. En el caso de otras empresas, la minería de datos de redes sociales aparece como un elemento más bien secundario dentro de una oferta de servicios de consultoría mucho más amplia (es el caso de las grandes consultoras de renombre). Finalmente, detectamos casos donde, si bien la descripción de la oferta se hace con un lenguaje apropiado para hablar de minería de datos comercial, podría tratarse de retóricas que buscan exagerar la sofisticación de aproximaciones más sencillas, como las que Kennedy cataloga de minería “dentro de la plataforma” o “gratis y simple”. De momento, los antecedentes que tenemos no nos permiten trazar líneas gruesas entre estas tres variantes, pero futuras investigaciones podrían arrojar más luz sobre estas diferencias.

3. **Las AMP ofrecen una propuesta de valor que es epistémicamente contradictoria.** En la matriz de las AMP se aprecia una importante tensión entre su vinculación, como sector, al sistema de las plataformas, como un eslabón más de la cadena de valor de los datos, por un lado; y por el otro, el énfasis que transmiten en la producción de mercancías infocomunicacionales dúctiles, que no responden a las condiciones objetivas del mundo sino a las necesidades y agendas de los clientes. Mercancías cuyos descriptores son eminentemente cualitativos y que acentúan el componente visual-persuasivo mientras oscurecen los elementos técnicos involucrados. Así, mientras en un sentido las AMP deben ser emisoras de las ideas hegemónicas de Silicon Valley y ofrecer, supuestamente una manera neutral y privilegiada (por su rapidez y envergadura) de comprender fenómenos, al mismo tiempo, deben asegurarles a sus clientes que nada de eso será un obstáculo para mostrar resultados de “éxito” o recuperar el control de las “narrativas”. Estas contradicciones tan patentes podrían configurar una cualidad inherente de las AMP, antes que de un factor a corregir, al igual que lo es la naturaleza esencialmente defectuosa de sus productos (en términos de su asociación referencial) de acuerdo a las evidencias recogidas.

4. **Las AMP ofrecen fundamentalmente herramientas de management.** La discusión técnica y metodológica sobre los desafíos de trabajar con los datos obtenidos a través de la API de las plataformas sociales, cuyos sesgos y distorsiones les son constitutivos, es casi por completo ignorada por estas empresas de cara a su clientela potencial. Si estos factores son tratados en diferentes etapas de la relación con los clientes, si se trata de asuntos que no son de interés de su clientela (como en los casos atestiguados por Kennedy) o si existe un ocultamiento por parte de las empresas, son aspectos que no podemos extrapolar ni resolver aquí. No obstante, como señala Meehan con agudeza, para aventurar refutaciones metodológicas de este tipo de mercancías o exigir una revisión entre pares que corrija sus problemas, primero deberíamos estar ante investigaciones moldeadas y desarrolladas bajo esas premisas epistémicas. En este caso, como en el discutido por Meeham, todas las observaciones que hemos recopilado indican que, por el contrario, estamos frente a mercancías sustancialmente asociadas al *management*, es decir, compuestas sobre la base de exigencias comerciales, imposiciones corporativas y necesidades del poder.
  
5. **Las AMP ofrecen productos cuyo formato y configuración particular es variable.** Las AMP proyectan, a través de sus sitios web corporativos, una amplia oferta de mercancías analíticas o de posibilidades de contribuir a los fines comerciales de sus clientes. Esta oferta, cuyas cualidades centrales fueron detalladas en la sección previa, comúnmente no es específica en cuanto a sus propiedades físicas particulares (materialidad, temporalidad, medios, etc.). Sin embargo, refleja aquello que las AMP consideran relevante de sus productos y lo que estiman es del interés de su clientela de acuerdo a su conocimiento del mercado, basado en una experiencia en el negocio que las mismas empresas proyectan como significativa.
  
6. **Las AMP se posicionan como consultores especializados en volver accesible un mundo caótico.** El negocio propiamente tal de las AMP consiste en asesorar y proveer a sus clientes de diferentes productos que les permiten lidiar con las condiciones actuales de circulación de los datos y la información. En el diagnóstico de las AMP, estas condiciones

son heterogéneas, dinámicas, caóticas, altamente competitivas y técnicamente exigentes. Se trata de un mundo que, pese a ser dibujado como ominoso y de una complejidad catastrófica, refleja en buena medida el tecnooptimismo igualitarista y sustantivamente democrático imaginado previo a la crisis de las puntocom, en que todos los hablantes gozan del mismo estatus, todas las opiniones pueden ser impugnadas y todos los humanos son sujetos de conocimiento. Las AMP se ofrecen como asesores especializados que ayudan a poner orden esta realidad y extraer de sus datos un potencial que es inaccesible por otras vías y para otros sujetos.

7. **Las AMP apuntan a un cliente que requiere insumos para acreditar conocimientos legítimos.** Los factores socioculturales que moldean al sector de las AMP son aquellos que dan forma al sistema de las plataformas, un cierto tipo de ideología que engloba exigencias comerciales, imposiciones corporativas y otras disposiciones hegemónicas. Pero a su vez, las AMP deben adaptarse a un mercado particular. En el caso que nos interesa aquí, este mercado está conformado por una pequeña élite local con el poder adquisitivo para acceder a los productos de datos que las AMP comercian, pero que ante todo tienen el interés (o la necesidad) de hacerlo. Lo que requiere de las AMP este cliente no son herramientas de persuasión dirigidas a las grandes audiencias (los mercados para aquello son otros y conocidos), sino mecanismos socialmente legitimados (datos en tiempo real, cruce de métricas heterogéneas, estimaciones de sentimientos, visualizaciones, etc.) para “tener la razón”. El uso efectivo de estos objetos de razón apunta a un nivel acotado, ya sea el organizacional (un encargado de relaciones públicas de una minera que necesita mostrarle al directorio el “éxito de su gestión” mediante métricas que atestiguan que la ruptura de un relave no tuvo impacto en la opinión pública), el corporativo (un director ejecutivo que necesita justificar una profunda reestructuración mediante cifras elocuentes sobre el pesimismo ciudadano respecto a la economía) o el político (como un ministro que busca imponer en el aparato gubernamental una teoría policial y de inteligencia que carece de evidencias materiales).

Llegados a este punto, es necesario hacer un alcance relativo a los resultados y la caracterización esbozada en este trabajo sobre el sector AMP. Si bien lo expuesto pudiera leerse

como una larga acusación contra ciertas empresas sobre unas determinadas falencias éticas, técnicas o incluso de probidad, es importante aclarar que hasta donde podemos atestiguar estas empresas cumplen con aquello que se les exige para ejecutar una función de mercado y que nuestro análisis crítico no apunta a compañías específicas, ni contra personas particulares, sino que a tratar de entender dichas funciones de mercado desde la perspectiva de la economía política de la comunicación. Aunque los procedimientos de las AMP frecuentemente aparecen como opacos, esto refleja una forma culturalmente apropiada de operar en ese mercado. Lo mismo se aplica al resto de los elementos que hemos delineado como descriptores del sector. El problema, por consiguiente, no es una cierta conducta inapropiada o poco proba del sector AMP ni de sus participantes. Como hemos indicado, el desafío estriba en entender a estas empresas con mayor precisión, para con ello acercarnos a la comprensión del rol social del sector en el que participan (sus relaciones con los diversos actores de la sociedad) y las consecuencias que de ello se derivan.

Este tipo de investigaciones se vuelven indispensables frente a la constatación de las amplias lagunas regulatorias en el ámbito de la economía de los datos, un tipo de marco que a nivel global y local recién comienza a introducirse, tímidamente, en las agendas políticas. La comprensión de los fenómenos asociados a estos mercados, considerando sus dimensiones sociales y las formas desiguales en que se distribuye el acceso a estas herramientas, son un paso previo no solo para determinar qué debe ser regulado y de qué forma, sino para resolver cuáles son los debates apropiados y las perspectivas adecuadas para abordarlos.

Así, por ejemplo, es indicativo atestiguar que entre los entrevistados de Helen Kennedy (hace más de una década) se manifestaba la creencia de que no era necesario regular el sector de la minería de datos ya que, en general, se podía confiar en la prevalencia del “sentido común”: con el tiempo, los clientes entenderían mejor los procedimientos y optarían por actores que ofrecieran mayor rectitud, dejando fuera del negocio a personas y empresas dedicadas a malas prácticas. Los hallazgos del estudio que hemos llevado a cabo indican que ese tipo de autorregulación es, cuando menos, dudosa.

Por último, nada de lo que hemos dicho hasta aquí implica que la idea de poder extraer cierto tipo de conocimiento del análisis de los datos de las plataformas sociales no tenga ninguna validez. Por el contrario, creemos que las oportunidades que ofrecen dichos datos y en general los procesos y procedimientos comprometidos en el despliegue del sistema de las plataformas son un fértil campo de estudio para diversas disciplinas. Sin embargo, la elaboración de mercancías bajo las constricciones de un mercado, a partir de los fundamentos ideológicos que caracterizan al sistema de las plataformas, no es en ningún sentido una forma de producir conocimiento.

#### **7.4. La relevancia de este estudio en la era post-API**

Para volver a encausar la reflexión dentro del contexto chileno, bajo la luz de la discusión que hemos desarrollado y las observaciones que hemos vertido, el tipo de preguntas que nos deberíamos empezar a formular es: aquellos que buscaron sacar provecho del informe Big Data, ¿eran conscientes de las limitaciones asociadas a dichas mercancías? ¿Las personas que elaboraron dicho informe dieron cuenta a sus clientes de las precauciones que cabe tener frente a estos datos? Atendiendo a las deficiencias evidentes de dicho informe, según todos los investigadores que en su momento fueron consultados, ¿qué grado de connivencia había entonces entre Alto Data Analytics y el Gobierno de Chile? Considerando el error estratégico de posicionar dicho informe como una mercancía de persuasión, expuesta al análisis crítico de expertos y a la incredulidad del público (equivocación que queda manifiesta tanto por estos efectos como por las artimañas empleadas para no transparentar por completo sus contenidos), ¿cuál era el verdadero colectivo al que iba dirigida esa mercancía, en tanto símbolo legitimado de razón? En otras palabras, ¿qué era lo accionable del informe Big Data, qué voluntades cupulares buscaba movilizar?

El propósito de este trabajo es contribuir a formular las preguntas adecuadas para la comprensión social del sector de las agencias de monitoreo de plataformas sociales. De forma secundaria, posibilita comenzar a poner en evidencia algunos procedimientos de ignorancia social deliberada y resaltar los perjuicios que pueden estar asociados a este mercado. Entender las

complejidades de las relaciones que se tejieron entre proveedores y clientes en el caso del informe Big Data a la luz de los hallazgos que hemos pormenorizado, es revelador para prepararnos para futuros escenarios. En esta sección final nos referiremos a las condiciones contingentes del sistema de las plataformas hoy, de la fase en que se encuentra el modelo de las API, de qué nos podría deparar el futuro y de cómo encajan las AMP, tal como las hemos caracterizado, en dicho esquema.

Hemos indicado con claridad que, hasta ahora, la principal vía de acceso de las AMP a los datos del sistema de las plataformas son las API que estas ofrecen, a través de las cuales es posible acceder a un gigantesco caudal de datos semiprocados de manera, tradicionalmente, gratuita. Un análisis más minucioso sobre la relación de estos elementos sociotécnicos y las AMP podría ser una importante forma de ampliar la presente investigación y entender la configuración efectiva actual del sector que hemos descrito aquí a partir de sus discursos. Esto es particularmente significativo si consideramos las transformaciones y continuidades que trae aparejada consigo lo que algunos han denominado “Era Post-API” (Freelon, 2018).

Hubo un tiempo en que el panorama de la minería de datos y del monitoreo de redes sociales pareció suficientemente consolidado (y libertarizado) para que se constituyera como un importante y dinámico mercado. Cualquier búsqueda en Google arrojaba decenas de resultados de empresas que formaban parte del negocio, con llamativos nombres como Gnip, Brandwatch, Sprinklr, Hootsuite, Topsy o Sysomos, entre otras. Este panorama afiebrado de especulación en torno a las capacidades del *big social data* llegó a su clímax hacia la primera mitad de la década de 2010, con señales claras de Silicon Valley y Wall Street respecto que el grifo de los datos no debía ser cerrado. En 2013, Apple, cuyo negocio ni siquiera está directamente vinculado con las redes sociales, compró Topsy en más de \$ 200 millones, mientras que Twitter, quizá como respuesta, se hizo con Gnip en 2014 por USD \$ 134,1 millones. Mientras las acciones de esas y otras empresas de análisis *social media* cambiaban de manos, innumerables Estados y organizaciones de la sociedad civil pusieron en estas tecnologías su confianza. Prestigiosas instituciones académicas en todo el mundo invirtieron cuantiosos recursos públicos en financiar investigaciones exclusivamente enfocadas en los datos que a través de sus API entregaban

(interesada o desinteresadamente, pero cuya elaboración no dejaba de ser opaca) empresas privadas (principalmente Twitter y Facebook) cuyo único fin era ganar dinero.

A este periodo de abundancia de data, de un acceso casi irrestricto a los datos vinculados a la actividad de cientos de millones de personas en Internet, Deen Freelon lo ha llamado “API Age” (Freelon, 2 de abril de 2022). Para otros investigadores se trataba del “*Wild West* de la investigación de redes sociales” (Puschmann, 2019), contexto que fue aprovechado por investigadores, pero también por especuladores y criminales de la más diversa índole. Los escándalos de datos, que reflejan la opacidad y la libertariedad con que ha operado el sistema de las plataformas, han puesto en evidencia las relaciones de datos dominantes, desprestigiado a la red por completo (son parte importante de la “ruptura” de Internet) y obligando a repliegues estratégicos del mercado.

En 2018, debido a su papel en la controversia de Cambridge Analytica, Facebook restringió el acceso a su Pages API. Desde ese momento, que marca el inicio de la Era Post-API, la historia de estos instrumentos ha sido errática, pero se ha apreciado una tendencia clara: cada vez es más difícil acceder a los datos y muchas menos personas y organizaciones pueden acceder a muchos menos datos. El “APIcalypsis” (Bruns, 2019) que significó Cambridge Analytica fue un preámbulo de lo que pasaría con Twitter tras su compra por parte de Elon Musk en 2022. La conducta de tenor irracional sobre las políticas que rigen a las diferentes API de dicha plataforma y los altos costos que ahora están asociados a ellas, significan el fin de facto del acceso a esos datos para el mundo académico, pero también para la mayoría de los consultores privados. De acuerdo a la última documentación publicada por X/Twitter, acceder a 1 millón de tweets mensuales (sobre un aproximado de 500 millones de tweets diarios) tiene un costo de USD \$ 5.000 cada mes. Sobre esa cantidad de datos, los planes comienzan en los USD \$ 42.000 mensuales (Twitter, 2023a; Wired, 2023). Nunca estuvo más lejana la idea de poder “monitorear todas las conversaciones”, “escucharlo todo” u “ofrecer una visión completa” de lo que pasa en las redes sociales.

En el advenimiento de la Era Post-API no solo tiene peso el esfuerzo de los directorios de las plataformas por evitar nuevos escándalos y sus consiguientes crisis comunicacionales. También deben considerarse otros factores, como las incipientes normas regulatorias, especialmente de los países de la Unión Europea, respecto a la privacidad digital. Asimismo, actualmente aparecen como menos comercialmente justificables los costos asociados, vía subvención cruzada, para el sostenimiento del modelo de las API, cuyas lógicas de modulación e imposición de teorías sobre lo relevante no resultan evidentemente ventajosas cuando el lenguaje y los valores de las plataformas ya se han vuelto hegemónicos. En la misma línea, la codicia de los dueños y accionistas de las plataformas también juega un rol (si las API son tan valiosas para otros, ¿por qué no cobrar más?). Desde luego, la crisis global producto de la pandemia de COVID-19 y las posteriores recesiones económicas de Estados Unidos y China, que han tornado a los negocios digitales algo más conservadores; también son factores a tener en cuenta en estos cambios de enfoque.

Las limitaciones en el acceso a las API ha obligado a los integrantes más marginales del sistema de las plataformas a replantearse los supuestos con los que habían trabajado durante varios intensos lustros. Así, la academia ha discutido ampliamente, especialmente desde 2018, su posición frente al uso de los datos sociales masivos, aunque probablemente no con la profundidad y responsabilidad que ameritaba el asunto: los problemas a los que se han enfrentado académicos e instituciones debido a los cambios unilaterales en las políticas de acceso a los datos de X/Twitter (Kupferschmidt, 2023; Ledford, 2023) eran completamente predecibles.

Es de esperar que el mismo tipo de reflexión, al menos por necesidad, haya estado llevando a cabo el sector AMP durante los últimos años. Es evidente que el flujo de datos de Twitter alimentaba gran parte de la producción de mercancías de este sector. Simbólicamente, los aspectos más destacados de su oferta, son las características más destacadas de Twitter (velocidad-tiempo real, vastedad y variedad de datos, pertinencia y relación con la contingencia). Otras plataformas que permiten el análisis de datos sociales, han puesto limitaciones severas a sus API desde hace más tiempo aún (Facebook, Instagram), ofrecen menos velocidad y vastedad o



masividad (LinkedIn), o también han adoptado recientemente medidas similares de las de X/Twitter (Reddit).

Ahora bien, pese a estas transformaciones, las lógicas imperantes detrás de la cadena de producción de valor del sistema de las plataformas siguen en pie. Como ha demostrado esta investigación, las mercancías analíticas, aparentemente, continúan a la venta. ¿El lenguaje es vago y las explicaciones sucintas quizá porque el futuro no está claro para estas compañías? No podemos tener una respuesta a esa pregunta aquí. Para entender si ya ha habido modificaciones importantes en el carácter de la oferta de las AMP haría falta un estudio longitudinal que nos mostrara el estado del sector previo a 2018. Con todo, pareciera que aún hoy queda algo que venderles a los clientes. Pero si bajo las condiciones actuales los datos que antes fluían hoy escasean, entonces ¿qué constituye la “materia prima” de la producción de los productos de las AMP? Quizá los datos, propiamente dichos, no eran tan importantes como el modelo de negocios socialmente valorado.

Pese a encontrarnos en la Era Post-API, la hegemonía de las plataformas no se ha visto afectada. El terreno sembrado por el modelo de las API y por Twitter en sus años de gloria aún no es baldío (la cantidad de contendientes a convertirse en la alternativa a Twitter, pese a ser una empresa famosa por no haber producido nunca utilidades, es señal de ello). Aquello que tradicionalmente las API hacía posible no es algo que vaya a quedar sepultado con su desaparición, aunque sí tendrá que transformarse. Es por ello que la comprensión social de los fenómenos que giran en torno al sistema de las plataformas, como el sector de las AMP en Chile, es crucial para que las futuras reelaboraciones que produzca el capitalismo de orientación a lo digital no sean aún peores a lo que teníamos.

Por otro lado, nuevas plataformas de uso masivo (como TikTok), junto a las sofisticaciones que trae consigo el amplio acceso a las tecnologías basadas en los grandes modelos de lenguaje, pueden ofrecer desconocidas oportunidades a la industria de la minería de datos comercial internacional y local. El sector AMP podría morir, pero también podría verse profundamente revitalizada en los próximos años. Bastan solo ciertas modulaciones al sistema de

las plataformas para que las teorías sobre lo relevante las impongan nuevos actores, por ejemplo, desde China. De momento, TikTok ofrece un acceso muy limitado a su API (TikTok, s. f.), pero eso podría cambiar en cualquier momento. Se trata de datos más complejos de analizar que los de X/Twitter, pero similarmente vastos, veloces y contingentes. La voluntad de potencias como China por dar la batalla hegemónica dentro del sistema de las plataformas, junto a procesamiento de minería y análisis a bajo costo, potenciados por inteligencia artificial, podrían en poco tiempo volverse el estándar desde el que las AMP proyecten a futuro la posición social que les hemos adjudicado provisionalmente en esta investigación: como legitimadoras de la opinión (y, en última instancia, de la posición) de una élite.

### **7.5. Colofón: aporte metodológico de esta investigación**

Creemos que el diseño metodológico expuesto en el capítulo 4 de este trabajo, que integra las técnicas de análisis de concordancia mediante software y el uso de un amplio corpus textual, cuyo foco es una muestra importante de sitios web corporativos, puede constituir un aporte metodológico para la comunicación política en Chile. El uso de herramientas de lingüística computacional, particularmente con base a corpus, en los estudios de comunicación política no ha sido un área intensamente abordada en el contexto chileno ni latinoamericano, por lo que esperamos que la presente exploración puede ser discutida, criticada, replicada en otros contextos de análisis y mejorada por otros autores y autoras.

Proponemos que, como apoyo a los análisis críticos del discurso, el trabajo con un corpus bien seleccionado, delimitado y preparado, cuya pertinencia fue razonada de acuerdo a los encuadres teóricos particulares de las investigaciones y sus objetivos, es lo que conforma el eje de nuestra metodología. Aproximaciones bajo estos preceptos pueden constituir un campo fértil de análisis, cuyos beneficios en pro de las investigaciones críticas en comunicaciones ya hemos abordado. Esto, considerando además que en la actualidad se trata de procedimientos que, si bien tienen cierta complejidad y deben ser desarrollados con atención al método, están relativamente

al alcance de cualquier investigador o investigadora, en términos de recursos y disponibilidad de las herramientas informáticas y de la literatura requerida para su empleo apropiado.

## BIBLIOGRAFÍA

- Abutridy, J. A. (23 de diciembre de 2019). *Informe big data del gobierno- la venta de humo en la República del chamullo*. BioBioChile. Recuperado en 2023-08-25 de <https://www.biobiochile.cl/noticias/opinion/tu-voz/2019/12/23/informe-big-data-del-gobierno-la-venta-de-humo-en-la-republica-del-chamullo.shtml>
- Adey, P. (2012). Borders, identification and surveillance. In K. Ball, K. D. Haggerty, & D. Lyon (Eds.), *Routledge Handbook of Surveillance Studies* (pp. 193-200). Routledge.
- Aguirre, F. (23 de diciembre de 2019). Big Data: ¿Es posible confiar en Twitter para saber cómo pensamos los chilenos? *La Tercera*. <https://www.latercera.com/que-pasa/noticia/big-data-en-chile-es-posible-confiar-en-twitter/949961/>
- Alto Analytics (10 de noviembre de 2016). *Alto Data Analytics on RTVE*. YouTube. Recuperado en 2023-08-23 de <https://youtu.be/sjVShEOjFt0?si=fkVIZRHnQgtKWjbs>
- Alto Analytics (5 de julio de 2015). *Alejandro Romero*. Recuperado en 2023-08-23 de [https://www.alto-analytics.com/en\\_US/alejandro-romero/](https://www.alto-analytics.com/en_US/alejandro-romero/)
- Alto Analytics (s.f.). *About Us*. Recuperado en 2021-08-20 de [https://www.alto-analytics.com/en\\_US/](https://www.alto-analytics.com/en_US/)
- Alto Analytics (s.f.). *Political Debate, Public Opinion, Polarization, and Social Unrest Analytics*. Recuperado en 2023-08-24 de [https://www.alto-analytics.com/en\\_US/solutions/social-debate-public-opinion-issue-mapping/](https://www.alto-analytics.com/en_US/solutions/social-debate-public-opinion-issue-mapping/)
- América Economía (8 de abril de 2013). *Las empresas descubren un tesoro en los datos de internet*. Recuperado en 2023-10-13 de <https://www.americaeconomia.com/articulos/las-empresas-descubren-un-tesoro-en-los-datos-de-internet>
- Ames, M. G. (2018). Deconstructing the algorithmic sublime. *Big Data & Society*, 5(1), 2053951718779194. <https://journals.sagepub.com/doi/full/10.1177/2053951718779194>
- Anderson, C. (23 de junio de 2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*. <https://www.wired.com/2008/06/pb-theory/>
- Andrejevic, M. (2011). The work that affective economics does. *Cultural studies*, 25(4-5), 604-620. <https://doi.org/10.1080/09502386.2011.600551>

- Andrejevic, M. (2013). *Infoglut: How Too Much Information Is Changing the Way We Think and Know*. Routledge.
- Andrejevic, M. (2014). The big data divide. *International Journal of Communication*, 8, 17.  
<https://ijoc.org/index.php/ijoc/article/viewFile/2161/1163>
- Anthony, L. (2005). AntConc: design and development of a freeware corpus analysis toolkit for the technical writing classroom. *IPCC 2005. Proceedings. International Professional Communication Conference, 2005*, 729-737. <https://doi.org/10.1109/IPCC.2005.1494244>
- Apruebo Dignidad. (2021). *Programa de Gobierno Apruebo Dignidad*.
- Araujo, K. (2019). Desmesuras, desencantos, irritaciones y desapegos. In *Hilos Tensados. Para leer el octubre chileno*. Editorial USACH.
- Arce, G. (24 de diciembre de 2019). *Experto que recomendó no hacer el informe de “Big data”:* “Fue como la operación Huracán. La academia les dijo que no se podía hacer”.  
 Publimetro. Recuperado en 2023-08-25 de  
<https://www.publimetro.cl/cl/noticias/2019/12/24/experto-recomendo-no-informe-big-data-fue-la-operacion-huracan-la-academia-les-dijo-no-se-podia.html>
- Arefin, S. E. (2023). A Survey on Actionable Knowledge. *arXiv preprint*.  
<https://arxiv.org/pdf/2301.09317>
- Argyris, C. (2003). Actionable Knowledge. In H. Tsoukas & C. Knudsen (Eds.), *The Oxford Handbook of Organization Theory* (pp. 423- 452). Oxford University Press.
- Armano, E., Briziarelli, M., Flores, J., & Risi, E. (2022). Platforms, Algorithms and Subjectivities: Active Combination and the Extracting Value Process—An Introductory Essay. In *Digital Platforms and Algorithmic Subjectivities* (pp. 1-18). University of Westminster Press.  
<https://uwestminsterpress.co.uk/site/books/10.16997/book54/download/8894/#page=12>
- AWS (s.f.). *¿Qué es una API. Explicación de interfaz de programación de aplicaciones*.  
 Recuperado en 2023-10-30 de <https://aws.amazon.com/es/what-is/api/>
- Ayala, L. (21 de diciembre de 2019a, 20 de agosto de 2021). El big data del Gobierno: Los detalles del informe por el estallido social que entregó a la Fiscalía. *La Tercera*.  
<https://www.latercera.com/la-tercera-domingo/noticia/big-data-del-gobierno-los-detalles-del-informe-estallido-social-entrego-la-fiscalia/947967/>

- Ayala, L. (21 de diciembre de 2019b, 9 de agosto de 2023). El big data del Gobierno. Los detalles del informe por el estallido social que entregó a la Fiscalía. *La Tercera*. <https://www.latercera.com/la-tercera-domingo/noticia/big-data-del-gobierno-los-detalles-del-informe-estallido-social-entrego-la-fiscalia/947967/>
- Ayala, L., & Ojeda, J. M. (30 de diciembre de 2019). Informe 'Big Data' fue elaborado por la empresa española Alto Data Analytics. *La Tercera*. <https://www.latercera.com/nacional/noticia/informe-big-data-fue-elaborado-la-empresa-espanola-alto-data-analytics/957097/>
- Baeza-Yates, R. (28 de febrero de 2020). *Los Asombrosos Errores del Análisis de Redes Sociales Chilenas de Alto Analytics*. Medium. Recuperado en 2023-08-25 de <https://rbaeza-yates.medium.com/los-asombrosos-errores-del-2b0225c2e622>
- Baker, D. (19 de diciembre de 2017). *The internet is broken. but the web's top minds have a plan to fix it*. WIRED UK. Recuperado en 2023-09-13 de <https://www.wired.co.uk/article/is-the-internet-broken-how-to-fix-it>
- Baker, P. (2006). *Using Corpora in Discourse Analysis*. Continuum.
- Ball, K., Haggerty, K., & Lyon, D. (2012). *Routledge Handbook of Surveillance Studies*. Taylor & Francis.
- Barnes, J. A. (1954). Class and committees in a Norwegian island parish. *Human relations*, 7(1), 39-58. <https://journals.sagepub.com/doi/abs/10.1177/001872675400700102>
- Barney, N. (s. f.). *What Is Sentiment Analysis (Opinion Mining)*. TechTarget. Recuperado en 2023-11-13 de <https://www.techtarget.com/searchbusinessanalytics/definition/opinion-mining-sentiment-mining>
- Basoalto, H. (15 de enero de 2020). Fiscal nacional por informe 'Big Data': 'No hemos encontrado ningún valor y en consecuencia no ha sido necesario indagar el origen del mismo'. *La Tercera*. <https://www.latercera.com/nacional/noticia/fiscal-nacional-informe-big-data-no-encontrado-ningun-valor-consecuencia-no-ha-necesario-indagar-origen-del/974888/>
- Bayart, J. (2011). *Africa en el espejo. Colonización, Criminalidad y Estado*. Fondo de Cultura Económica, México.

- Becerra, M., & Mastrini, G. (2014). Communication Economy Paths. A Latin American Approach. In J. Wasko, G. Murdock, & S. Helena (Eds.), *The Handbook of Political Economy of Communications* (pp. 109-125). John Wiley & Sons.
- Beer, D. (2016). *Metric Power*. Springer.
- Bekkers, V., Edwards, A., & Kool, D. D. (2013). Social media monitoring: Responsive governance in the shadow of surveillance? *Government Information Quarterly*, 30(4), 335-342. <https://doi.org/10.1016/j.giq.2013.05.024>
- Bermúdez, J. (24 de septiembre de 2020). *Dictamen E37908N20*. Contraloría General de la República. Recuperado en 2023-08-23 de <https://www.contraloria.cl/pdfbuscador/dictamenes/E37908N20/pdf>
- Betancourt, M. (2016). *The Critique of Digital Capitalism*. punctum books.
- Bhageshpur, K. (15 de noviembre de 2019). Data Is The New Oil -- And That's A Good Thing. *Forbes.com*. <https://www.forbes.com/sites/forbestechcouncil/2019/11/15/data-is-the-new-oil-and-thats-a-good-thing/?sh=6485e2c67304>
- BigCommerce (s.f.). *What is social media monitoring*. Recuperado en 2023-11-12 de <https://www.bigcommerce.com/ecommerce-answers/what-social-media-monitoring/>
- BioBioChile (19 de junio de 2011). *BrandMetric: ¿qué mide la empresa contratada por el gobierno para seguir opiniones de los chilenos?* Recuperado en 2023-11-01 de <https://www.biobiochile.cl/noticias/2011/06/19/brandmetric-que-mide-la-empresa-contratada-por-el-gobierno-para-seguir-opiniones-de-los-chilenos.shtml>
- Black, J. (2002). *A Dictionary of Economics*. Oxford University Press, USA.
- Bloom, P. (2019). *Monitored*. Pluto Press (UK).
- Blumel, G. (2023). *La vuelta larga: Crónica personal de la Crisis de Octubre*. Ediciones UC.
- Bolaño, C. R. S., Montoya, A. N., & Lopes, R. S. (2019). Apuntes sobre la economía política de la información, la comunicación y la cultura latinoamericana. *Chasqui: Revista Latinoamericana de Comunicación*, 141), 215-234. <https://dialnet.unirioja.es/servlet/articulo?codigo=7320765>
- Bonasio, A. (19 de septiembre de 2014). *What's an API and Why It Matters to You*. HuffPost UK. Recuperado en 2023-10-30 de [https://www.huffingtonpost.co.uk/alice-bonasio/whats-an-api-and-why-it-m\\_b\\_5848340.html](https://www.huffingtonpost.co.uk/alice-bonasio/whats-an-api-and-why-it-m_b_5848340.html)

- Botta, F., Moat, H. S., & Preis, T. (2015). Quantifying crowd size with mobile phone and Twitter data. *Royal Society open science*, 2(5), 150162.  
<https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.150162>
- Bouazizi, M., & Ohtsuki, T. (2019). Multi-class sentiment analysis on twitter: Classification performance and challenges. *Big Data Mining and Analytics*, 2(3), 181-194.  
<https://ieeexplore.ieee.org/abstract/document/8681053>
- boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5), 662-679.  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.441.9822&rep=rep1&type=pdf>
- Brandt, J., Buckingham, K., Buntain, C., Anderson, W., Ray, S., Pool, J.-R., & Ferrari, N. (2020). Identifying social media user demographics and topic diversity with computational social science: a case study of a major international policy forum. *Journal of Computational Social Science*, 3, 167-188. <https://doi.org/10.1007/s42001-019-00061-9>
- Bravo, D. (18 de octubre de 2022). *Las cifras del estallido social según el INDH: 3.626 víctimas, 2.987 carabineros denunciados, 13 sentencias condenatorias y más*. ADN Radio. Recuperado en 2023-08-21 de <https://www.adnradio.cl/nacional/2022/10/18/cifras-estallido-social-segun-indh-3626-victimas-2987-carabineros-13-condenas-mas.html>
- Bright, J., & Margetts, H. (2016). Big Data and Public Policy: Can It Succeed Where E-Participation Has Failed. *Policy & Internet*, 8, 218-224.  
<https://doi.org/https://doi.org/10.1002/poi3.130>
- Bruns, A. (2019). After the ‘APIcalypse’: social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544-1566.  
<https://doi.org/10.1080/09502386.2011.600551>
- Buhmann, A., & Fieseler, C. (2023). Deep learning meets deep democracy: Deliberative governance and responsible innovation in artificial intelligence. *Business Ethics Quarterly*, 33(1), 146-179. <https://doi.org/10.1017/beq.2021.42>
- Burgess, J., & Green, J. (2018). *YouTube: Online Video and Participatory Culture*. John Wiley & Sons.



- Canhoto, A. I., Vom Lehn, D., Kerrigan, F., Yalkin, C., Braun, M., & Steinmetz, N. (2015). Fall and redemption: Monitoring and engaging in social media conversations during a crisis. *Cogent Business & Management*, 2(1), 1084978.  
<https://doi.org/10.1080/23311975.2015.1084978>
- Caraway, B. (2011). Audience labor in the new media environment: A Marxian revisiting of the audience commodity. *Media, Culture & Society*, 33(5), 693-708.  
<https://journals.sagepub.com/doi/abs/10.1177/0163443711404463>
- Cárdenas, L. (11 de mayo de 2018). *De Gastón y Nova a Contralorito- El 'equipo millennial' que difunde por redes sociales los dictámenes del contralor*. La Tercera PM. Recuperado en 2023-10-21 de <https://www.latercera.com/la-tercera-pm/noticia/gaston-nova-contralorito-equipo-millennial-difunde-redes-sociales-los-dictamenes-del-contralor/161240/>
- Castells, M. (2001). *La galaxia Internet*. Arete.
- Castells, M. (2013). *Comunicación y poder*. Siglo XXI Editores.
- Catena, P. (7 de enero de 2020). Jorge Selume renuncia a la Secretaría de Comunicaciones de La Moneda. *La Tercera*. <https://www.latercera.com/politica/noticia/jorge-selume-renuncia-la-secretaria-comunicaciones-la-moneda/965198/>
- Centro de Documentación Europea Universidad de Granada (s.f.). *Ley de Servicios Digitales: nuevas normas que garantizan un entorno en línea seguro*. Recuperado en 2023-10-20 de <https://cde.ugr.es/index.php/union-europea/noticias-%20ue/1379-ley-de-servicios-digitales-nuevas-norma>
- Cesare, N., Lee, H., McCormick, T., & Spiro, E. (2018). Promises and pitfalls of using digital traces for demographic research. *Demography*, 55, 1979-1999.  
<https://doi.org/10.1007/s13524-018-0715-2>
- Chen, S. (2018). *Big Data in Computational Social Science and Humanities*. Springer.
- CNN Chile (14 de enero de 2020). *Informe Big Data: Quiñenco asegura que puso al gobierno en contacto con Alto Data Analytics, pero que no encargó estudio*. Recuperado en 2023-08-23 de [https://www.cnnchile.com/pais/big-data-quinenco-contacto-gobierno-no-encargo-estudio\\_20200114/](https://www.cnnchile.com/pais/big-data-quinenco-contacto-gobierno-no-encargo-estudio_20200114/)
- CNN Chile (19 de diciembre de 2019b). *Abbott señala que el gobierno le entregó "información de fuentes abiertas, no de inteligencia"*. Recuperado en 2023-08-21 de

[https://www.cnnchile.com/pais/abbott-gobierno-entrego-informacion-fuentes-abiertas-no-de-inteligencia\\_20191219/](https://www.cnnchile.com/pais/abbott-gobierno-entrego-informacion-fuentes-abiertas-no-de-inteligencia_20191219/)

CNN Chile (26 de diciembre de 2019c). *Piñera afirma que “muchos de los videos” sobre violaciones a los DD.HH. “son filmados fuera de Chile”*. Recuperado en 2023-08-23 de [https://www.cnnchile.com/pais/pinera-videos-violaciones-dd-hh-fuera-chile\\_20191226/](https://www.cnnchile.com/pais/pinera-videos-violaciones-dd-hh-fuera-chile_20191226/)

CNN Chile (28 de octubre de 2019a). *Ataques al Metro- Fiscal Mérida niega información sobre sospechosos cubanos y venezolanos*. Recuperado en 2023-08-22 de [https://www.cnnchile.com/pais/metro-fiscal-niega-sospechosos-cubanos-venezolanos\\_20191028/](https://www.cnnchile.com/pais/metro-fiscal-niega-sospechosos-cubanos-venezolanos_20191028/)

Cohen, N. S. (2008). The valorization of surveillance: Towards a political economy of Facebook. *Democratic Communiqué*, 22(1), 1.

Comisión Nacional de Energía. (2022). *Informe preliminar de previsión de demanda 2022-2042*. CNE, Gobierno de Chile. <https://www.cne.cl/wp-content/uploads/2022/12/Informe-Preliminar-Prevision-de-Demanda-2022-2042.pdf>

Comito, C. (2018). Human mobility prediction through Twitter. *Procedia computer science*, 134, 129-136. <https://www.sciencedirect.com/science/article/pii/S1877050918311165>

Cooperativa (31 de octubre de 2019). *Trump denunció injerencia externa en las protestas en Chile*. Recuperado en 2023-08-22 de <https://www.cooperativa.cl/noticias/pais/manifestaciones/trump-denuncio-injerencia-externa-en-las-protestas-en-chile/2019-10-31/125105.html>

Coopmans, C. (2014). Visual analytics as artful revelation. In C. Coopmans, J. Vertesi, M. E. Lynch, & S. Woolgar (Eds.), *Representation in Scientific Practice Revisited* (pp. 37-59). Cambridge, MA: MIT Press.

Coopmans, C., Vertesi, J., Lynch, M. E., & Woolgar, S. (2014). *Representation in Scientific Practice Revisited*. MIT Press.

Cossio, H. (20 de octubre de 2020). *El «Plan Zeta» de Piñera: el informe de la DINE sobre la amenaza extranjera que hizo que el Presidente hablara de guerra*. El Mostrador. Recuperado en 2023-08-22 de <https://www.elmostrador.cl/destacado/2020/10/20/el-plan-zeta-de-pinera-el-informe-de-la-dine-sobre-la-amenaza-extranjera-que-hizo-que-el-presidente-hablara-de-guerra/>

- Couldry, N., & Mejias, U. (2020). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.  
<https://doi.org/10.1515/9781503609754/html>
- Couldry, N., & Mejias, U. (2021). The decolonial turn in data and technology research: what is at stake and where is it heading. *Information, Communication & Society*, 26(4), 786-802.  
<https://doi.org/10.1080/1369118X.2021.1986102>
- Crawford, K. (2021). *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
- Dammert, L., & Sazo, D. (20 de marzo de 2021). *La teoría del complot en el Estallido chileno-un examen crítico*. Ciper. Recuperado en 2023-08-22 de  
<https://www.ciperchile.cl/2021/03/20/la-teoria-del-complot-en-el-estallido-chileno-un-examen-critico/>
- de Sousa Santos, B. (2018). *Justicia entre saberes: Epistemologías del Sur contra el epistemicidio*. Ediciones Morata.
- de Vries, A. (2023). The growing energy footprint of artificial intelligence. *Joule*.  
<https://doi.org/10.1016/j.joule.2023.09.004>
- Dencik, L., & Leistert, O. (2015). *Critical Perspectives on Social Media and Protest*.
- Dey, I. (2005). *Qualitative Data Analysis: A User Friendly Guide for Social Scientists*. Routledge.
- Diario Uchile (16 de diciembre de 2019). *Sebastián Piñera- “No habíamos encontrado el equilibrio entre desarrollo y equidad”*. Recuperado en 2023-08-23 de  
<https://radio.uchile.cl/2019/12/16/sebastian-pinera-no-habiamos-encontrado-el-equilibrio-entre-desarrollo-y-equidad/>
- Dougnac, A. (1994). *Manual de historia del derecho indiano*. Universidad Nacional Autónoma de México.
- Driscoll, K., & Walker, S. (2014). Working Within a Black Box: Transparency in the Collection and Production of Big Twitter Data. *International Journal of Communication*, 8, 20.  
<https://ijoc.org/index.php/ijoc/article/download/2171/1159>

- Dunne, R., Matthews, O., Vega, J., Harper, S., & Morris, T. (2023). Computational methods for predicting human behaviour in smart environments. *Journal of Ambient Intelligence and Smart Environments*, 15(2), 179-205. <https://doi.org/10.3233/AIS-210384>
- Engel, E. (20 de enero de 2020). Informe “Big Data”: Correlaciones vs. evidencia. *El Mercurio*. <https://espaciopublico.cl/en/informe-big-data-correlaciones-vs-evidencia/>
- El Mostrador (17 de diciembre de 2019). *Fiscal Guerra desmiente a Piñera: Ministerio Público «no tiene ningún antecedente de un grupo que haya actuado desde afuera»*. El Mostrador. Recuperado en 2023-08-23 de <https://www.elmostrador.cl/dia/2019/12/17/fiscal-guerra-desmiente-a-pinera-ministerio-publico-no-tiene-ningun-antecedente-de-un-grupo-que-haya-actuado-desde-afuera/>
- El Mostrador (27 de febrero de 2020). *Ministerio Público registra 33 muertes durante el estallido social: 4 se atribuyen a agentes del Estado*. Recuperado en 2023-08-21 de <https://www.elmostrador.cl/dia/2020/02/27/ministerio-publico-registra-33-muertes-durante-el-estallido-social-4-se-atribuyen-a-agentes-del-estado>
- Elliott, V. (7 de noviembre de 2022). *Elon Musk Has Put Twitter’s Free Speech in Danger*. WIRED. Recuperado en 2023-11-01 de <https://www.wired.com/story/twitter-free-speech-musk-takeover/>
- Engelbrechtsen, M., & Kennedy, H. (2020). *Data Visualization in Society*. Amsterdam University Press.
- Ericson, R. V., & Haggerty, K. D. (2006). *The New Politics of Surveillance and Visibility*. University of Toronto Press.
- European Commission (2023). *The EU’s Digital Services Act*. Recuperado en 2023-10-20 de [https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en)
- Fairclough, N. (2013). Critical discourse analysis. In *The Routledge handbook of discourse analysis* (pp. 9-20). Routledge. <https://doi.org/10.4324/9780203809068>
- Fernandez Vilas, A., Diaz Redondo, R. P., & Ben Khalifa, M. (2019). Analysis of crowds’ movement using Twitter. *Computational Intelligence*, 35(2), 448-472. <https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/coin.12205>

- Fisher, E. (2016). How less alienation creates more exploitation? Audience labour on social network sites. In C. Fuchs & V. Mosco (Eds.), *Marx in the Age of Digital Capitalism* (pp. 180-203). Brill. <https://www.triple-c.at/index.php/tripleC/article/view/392/357>
- Flowerdew, L. (2013). Corpus-based discourse analysis. In *The Routledge handbook of discourse analysis* (pp. 174-187). Routledge.
- France 24 (25 de octubre de 2019). *Almagro denuncia 'patrón' de desestabilización de Venezuela y Cuba en la región*. France 24. Recuperado en 2023-08-22 de <https://www.france24.com/es/20191025-almagro-denuncia-patrón-de-desestabilización-de-venezuela-y-cuba-en-la-región>
- Frankfurt, H. G. (2006). *On bullshit: sobre la manipulación de la verdad*. Ediciones Paidós Iberica.
- Freelon, D. (2 de abril de 2022). *The Post-API Age Reconsidered* [Video]. SMART Social Media Research Techniques. Recuperado en 2024-03-20 de <https://www.youtube.com/watch?v=APHms1w8RY8>
- Freelon, D. (2018). Computational Research in the Post-API Age. *Political Communication*, 35(4), 665-668. <https://doi.org/10.1080/10584609.2018.1477506>
- Fuchs, C. (2010). Labor in Informational Capitalism and on the Internet. *The Information Society*, 26(3), 179-196. <https://westminsterresearch.westminster.ac.uk/item/905x5/labor-in-informational-capitalism-and-on-the-internet>
- Fuchs, C. (2012a). Dallas Smythe Today-The Audience Commodity, the Digital Labour Debate, Marxist Political Economy and Critical Theory. *TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 10(2), 692-740. <https://www.triple-c.at/index.php/tripleC/article/view/443/414>
- Fuchs, C. (2012b). Una contribución a la crítica de la economía política del capitalismo informacional transnacional. *Nómadas*, 36), 27-41. [http://www.scielo.org.co/scielo.php?pid=S0121-75502012000100003&script=sci\\_arttext](http://www.scielo.org.co/scielo.php?pid=S0121-75502012000100003&script=sci_arttext)
- Fuchs, C. (2017). From digital positivism and administrative big data analytics towards critical digital and social media research. *European Journal of Communication*, 32(1), 37-49. <https://doi.org/10.1177/0267323116682804>
- Fuchs, C. (2021). *Social Media*. Sage Publications Limited.

- Galli, J. F. (3 de abril de 2020). *Ordinario N° 9.573 en respuesta a la solicitud de información N° AB001T0001252*. Ministerio del Interior y Seguridad Pública. Recuperado en 2023-08-23 de <https://www.interior.gob.cl/transparenciaactiva/doc/ActosSecretos/3967.pdf>
- Gandy, O. H. (2011). The Political Economy of Personal Information. In J. Waski, G. Murdock, & H. Sousa (Eds.), *The Handbook of Political Economy of Communications* (pp. 436-457). John Wiley & Sons.  
<https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781444395402#page=446>
- Gatica, C. (19 de junio de 2011). *Gobierno iniciará seguimiento web de comentarios sobre administración de Sebastián Piñera*. BioBioChile. Recuperado en 2023-08-23 de <https://www.biobiochile.cl/noticias/2011/06/19/gobierno-iniciara-monitoreo-web-de-comentarios-sobre-administracion-de-sebastian-pinera.shtml>
- Gibney, E. (29 de marzo de 2018). *The scant science behind Cambridge Analytica's controversial marketing techniques*. Nature. Recuperado en 2021-08-24 de <https://www.nature.com/articles/d41586-018-03880-4>
- Gitelman, L. (2013). *Raw Data Is an Oxymoron*. MIT Press.
- Goggin, B. (8 de octubre de 2023). *Verified accounts spread fake news release about a Biden \$8 billion aid package to Israel*. NBC News. Recuperado en 2023-11-01 de <https://www.nbcnews.com/tech/internet/verified-accounts-spread-fake-news-release-biden-8-billion-aid-package-rcna119372>
- Golding, P., & Murdock, G. (1997). *The Political Economy of the Media* (1). Edward Elgar Pub.
- Google (1 de abril de 2012). *Our history in depth*. Internet Archive. Recuperado en 2023-10-20 de <https://web.archive.org/web/20120401005940/http://www.google.com/about/company/history/>
- Gregory, A., & Half, G. (2020). The damage done by big data-driven public relations. *Public Relations Review*, 46(2), 101902.  
<https://www.sciencedirect.com/science/article/pii/S0363811120300278>
- Gutiérrez, C. (28 de diciembre de 2019). *¿Inteligencia artificial o estupidez natural- Breve introducción a 'Big Data'*. Ciper. Recuperado en 2023-08-23 de

<https://www.ciperchile.cl/2019/12/28/inteligencia-artificial-o-estupidez-natural-breve-introduccion-a-big-data/>

- Hacking, I. (1996). *Representar e intervenir* (S. Martínez, Trans.). Paidós.
- Hacking, I. (2007). Kinds of people: Moving targets. *Proceedings-British Academy*, 151, 285.
- Hacking, I. (2015). Biopower and the avalanche of printed numbers. In V. W. C. & N. Morar (Eds.), *Biopower: Foucault and beyond*. University of Chicago Press Chicago, IL.
- Hallahan, K., Holtzhausen, D., Van Ruler, B., Verčič, D., & Sriramesh, K. (2007). Defining strategic communication. *International journal of strategic communication*, 1(1), 3-35.  
<https://doi.org/10.1080/15531180701285244>
- Hardy, J. (2014). *Critical Political Economy of the Media*. Routledge.
- Heath, A. (31 de octubre de 2023). *Elon Musk's 'everything app' plan for X, in his own words*. The Verge. Recuperado en 2023-11-01 de <https://www.theverge.com/23940924/elon-musk-x-twitter-all-hands-linda-yaccarino-super-app>
- Herrero, R. D. (2016). A twofold commodification of “place” in hotel websites and its consequences for the discursive creation of a tourist identity. *Ibérica*, 31), 63-82.  
<http://www.revistaiberica.org/index.php/iberica/article/download/189/148>
- Herrero, V. (14 de enero de 2020). *Exclusivo- polémico informe de Big Data habría sido entregado al gobierno por Luksic*. Interferencia. Recuperado en 2023-08-23 de <https://interferencia.cl/articulos/exclusivo-polemico-informe-de-big-data-habria-sido-entregado-al-gobierno-por-luksic>
- Herrero, V. (16 de enero de 2020). *No sólo Hinzpeter- otro alto ejecutivo de Grupo Luksic participó en reunión de inteligencia con la ANI*. Interferencia. Recuperado en 2023-08-23 de <https://interferencia.cl/articulos/no-solo-hinzpeter-otro-alto-ejecutivo-de-grupo-luksic-participo-en-reunion-de-inteligencia>
- French, H. (23 de noviembre de 2022). *What We Lose if We Lose Twitter*. Foreign Policy. Recuperado en 2023-11-01 de <https://foreignpolicy.com/2022/11/23/elon-musk-twitter-free-speech-democracy/>
- Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. *Big Data & Society*, 3(2), 205395171667423. <https://doi.org/10.1177/2053951716674238>

- Infobae (28 de octubre de 2019). *Medios chilenos aseguran que cubanos y venezolanos participaron en las violentas protestas contra el gobierno de Sebastián Piñera*. Recuperado en 2023-08-22 de <https://www.infobae.com/america/america-latina/2019/10/28/medios-chilenos-aseguran-que-cubanos-y-venezolanos-participaron-en-las-violentas-protestas-contra-el-gobierno-de-sebastian-pinera/>
- Jakes, L. (21 de enero de 2020). Con las protestas en Sudamérica también aparecieron troles rusos en Twitter. *The New York Times*.  
<https://www.nytimes.com/es/2020/01/21/espanol/america-latina/troles-rusos-sudamerica.html>
- Jensen, K. B. (2020). *A Handbook of Media and Communication Research: Qualitative and Quantitative Methodologies*. Routledge.
- Jeřábek, H. (2017). *Paul Lazarsfeld and the Origins of Communications Research*. Routledge.
- John, N. A. (2019). Social media bullshit: What we don't know about facebook. com/peace and why we should care. *Social Media+ Society*, 5(1), 2056305119829863.  
<https://journals.sagepub.com/doi/pdf/10.1177/2056305119829863>
- Jurgenson, N. (24 de febrero de 2011). *Digital dualism versus augmented reality*. The Society Pages - Cyborgology. Recuperado en 2023-10-24 de <https://thesocietypages.org/cyborgology/2011/02/24/digital-dualism-versus-augmented-reality/>
- Kalning, K. (11 de marzo de 2007). If Second Life isn't a game, what is it? NBC News. Recuperado en 2023-10-10 de <https://www.nbcnews.com/id/wbna17538999>
- Kang, H., & McAllister, M. P. (2011). Selling you and your clicks: examining the audience commodification of Google. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*. <https://triple-c.at/index.php/tripleC/article/view/255/234>
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons*, 53(1), 59-68.  
<https://www.sciencedirect.com/science/article/pii/S0007681309001232>
- Kast, J. A. (2021). *Plan para el futuro de Chile. Programa de Gobierno de José Antonio Kast 2022-2025*.



- Keen, A. (2015). *The Internet is Not the Answer*. Atlantic Books Ltd.
- Kelsey, T. (2017). *Introduction to Social Media Marketing*. Apress.
- Kennedy, H. (2016). *Post, Mine, Repeat: Social Media Data Mining Becomes Ordinary*. Palgrave Macmillan.
- Kennedy, H., & Engebretsen, M. (2020). Introduction : The relationships between graphs, charts, maps and meanings, feelings, engagements. In *Data Visualization in Society*. Amsterdam University Press. [https://doi.org/10.5117/9789463722902\\_ch01](https://doi.org/10.5117/9789463722902_ch01)
- Kennedy, H., & Hill, R. L. (2017). The feeling of numbers: Emotions in everyday engagements with data and their visualisation. *Sociology*, 52(4), 830-848. <https://journals.sagepub.com/doi/abs/10.1177/0038038516674675>
- Kent, M. (2012). When community becomes a commodity. In *Digital Dialogues and Community 2.0* (pp. 243-252). Elsevier. <https://www.sciencedirect.com/science/article/pii/B9781843346951500170>
- Knoche, M. (2021). Capitalisation of the Media Industry From a Political Economy Perspective. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 19(2), 325-342. <https://doi.org/10.31269/triplec.v19i2.1283>
- Kollock, P., & Smith, M. (2002). *Communities in Cyberspace*. Routledge.
- Kumar Gharai, D. (2 de agosto de 2023). *Social Media Measurement- A Guide to Success*. Sprinklr.com. Recuperado en 2023-11-13 de <https://www.sprinklr.com/blog/social-media-measurement/>
- Kupferschmidt, K. (2023). Twitter's threat to curtail free data access angers scientists. *Science*, 379(6633), 624-625. <https://www.science.org/doi/pdf/10.1126/science.adh1696>
- La Tercera (14 de enero de 2020). Quién es el ejecutivo español que viajó a Chile en medio de la crisis. *La Tercera*. <https://www.latercera.com/la-tercera-pm/noticia/informe-bigdata-quien-ejecutivo-espanol-viajo-chile-medio-la-crisis/973202/>
- La Tercera (28 de octubre de 2019). Gobierno rastrea rol de venezolanos en Twitter. *La Tercera*. <https://www.latercera.com/nacional/noticia/gobierno-rastrea-rol-venezolanos-twitter/880408/>

- Laaksonen, S.-M., & Pääkkönen, J. (2020). Between automation and interpretation: Using data visualization in social media analytics companies. In M. Engebretsen & H. Kennedy (Eds.), *Data Visualization in Society* (pp. 95-109). Amsterdam University Press Amsterdam. <https://library.oapen.org/bitstream/handle/20.500.12657/22273/9789048543137.pdf?sequence=1#page=96>
- Ledford, H. (2023). Researchers scramble as Twitter plans to end free data access. *Nature*. <https://doi.org/10.1038/d41586-023-00460-z>
- Lévi-Strauss, C. (1964). *El pensamiento salvaje*. Fondo De Cultura Economica.
- Limón, R. (1 de octubre de 2018). *La fiebre del 'Big Data' llega al empleo*. El País. Recuperado en 2023-10-31 de [https://elpais.com/tecnologia/2018/09/25/actualidad/1537892876\\_048111.html](https://elpais.com/tecnologia/2018/09/25/actualidad/1537892876_048111.html)
- Lopez-Joya, S., Diaz-Garcia, J. A., Ruiz, M. D., & Martin-Bautista, M. J. (2023). *Bot Detection in Twitter: An Overview*. In *International Conference on Flexible Query Answering Systems* (pp. 131-144). Springer. [https://link.springer.com/chapter/10.1007/978-3-031-42935-4\\_11](https://link.springer.com/chapter/10.1007/978-3-031-42935-4_11)
- Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. (2023). A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature human behaviour*, 7(1), 74-101. <https://doi.org/10.1038/s41562-022-01460-1>
- Louridas, P. (2020). *Algorithms*. MIT Press.
- Ludwig, D., Hilborn, R., & Walters, C. (1993). Uncertainty, resource exploitation, and conservation: lessons from history. *Science*, 260(5104), 17-36. <https://www.science.org/doi/pdf/10.1126/science.260.5104.17>
- Lyon, D. (2003). *Surveillance as Social Sorting*. Routledge.
- Manheim, K., & Kaplan, L. (2019). Artificial intelligence: Risks to privacy and democracy. *Yale JL & Tech*. [https://yjolt.org/sites/default/files/21\\_yale\\_j.l.\\_tech.\\_106\\_0.pdf](https://yjolt.org/sites/default/files/21_yale_j.l._tech._106_0.pdf)
- Marx, G. T. (2016). *Windows Into the Soul*. University of Chicago Press.
- Marx, K. (2008). *El Capital. Libro primero, Vol. 1*. Siglo XXI editores.
- Matei, S. A., Jullien, N., & Goggins, S. P. (2017). *Big Data Factories*. Springer.
- McChesney, R. (2013). *Desconexión digital. Cómo el capitalismo está poniendo a Internet en contra de la democracia*. El Viejo Topo.

- Meehan, E. R. (1984). Ratings and the institutional approach: A third answer to the commodity question. *Critical Studies in Mass Communication*, 1:2, 216-225.  
<https://doi.org/10.1080/15295038409360032>
- Mejias, U., & Couldry, N. (2019). Colonialismo de datos: repensando la relación de los datos masivos con el sujeto contemporáneo. *Virtualis*, 10(18), 78-97.  
<https://www.revistavirtualis.mx/index.php/virtualis/article/download/289/290>
- Méndez, C. (30 de marzo de 2023). *Data center de Google podría utilizar un máximo de 7.600.000 litros de agua potable por día*. La Diaria. Recuperado en 2023-10-31 de <https://ladiaria.com.uy/ambiente/articulo/2023/3/data-center-de-google-podria-utilizar-un-maximo-de-7600000-litros-de-agua-potable-por-dia/>
- Meta (s. f.). *Cómo funcionan las subastas de anuncios*. Facebook.com. Recuperado en 2023-10-18 de <https://web.facebook.com/business/ads/ad-auction>
- Metricool (3 de agosto de 2021). *Analítica en redes sociales - herramientas recomendadas*. Metricool. Recuperado en 2023-11-12 de <https://metricool.com/es/analitica-redes-sociales/>
- Milan, S., & Treré, E. (2019). Big data from the South (s): Beyond data universalism. *Television & New Media*, 20(4), 319-335.  
<https://journals.sagepub.com/doi/pdf/10.1177/1527476419837739>
- Miller, D., & Slater, D. (2020). *The Internet: An Ethnographic Approach*. Routledge.
- Mills, C. W. (2013). *La Elite del poder*. Fondo de Cultura Económica.
- Monaco, D., & Arnaudo, D. (2020). *Análisis de datos para monitoreo de redes sociales*. NDI.
- Montes, R. (2019). Más de un millón de personas protestan en Chile en una histórica marcha. *El País*. [https://elpais.com/internacional/2019/10/25/america/1572033004\\_292219.html](https://elpais.com/internacional/2019/10/25/america/1572033004_292219.html)
- Montes, R., & González, E. (9 de noviembre de 2019). “No supimos entender el clamor por una sociedad más justa”. *El País*.  
[https://elpais.com/internacional/2019/11/09/actualidad/1573313845\\_751195.html](https://elpais.com/internacional/2019/11/09/actualidad/1573313845_751195.html)
- Morozov, E. (4 de febrero de 2019). *Capitalism new clothes*. The Baffler. Recuperado en 2023-10-25 de <https://thebaffler.com/latest/capitalisms-new-clothes-morozov>
- Mosco, V. (2005). *The Digital Sublime: Myth, Power, and Cyberspace*. MIT Press.
- Mosco, V. (2009). *The Political Economy of Communication*. SAGE.
- Mosco, V. (2015). *To the Cloud*. Routledge.

- Mosco, V. (2016). Marx in the Cloud. In *Marx in the age of digital capitalism* (pp. 516-535). Brill. <http://digamo.free.fr/fuchsmosco.pdf#page=525>
- Mosco, V. (2017). *Becoming Digital*. Emerald Group Publishing.
- Mozur, P. (15 de octubre de 2018). A Genocide Incited on Facebook, With Posts From Myanmar's Military. *The New York Times*.  
<https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html>
- Mullet, D. R. (2018). A general critical discourse analysis framework for educational research. *Journal of Advanced Academics*, 29(2), 116-142.  
<https://journals.sagepub.com/doi/abs/10.1177/1932202X18758260>
- Muñoz, A., & Catena, P. (23 de diciembre de 2019). Las repercusiones del big data del gobierno. El hijo que nadie quiere reconocer en La Moneda. *La Tercera*.  
<https://www.latercera.com/la-tercera-pm/noticia/las-repercusiones-del-big-data-del-gobierno-hijo-nadie-quiere-reconocer-la-moneda/949656/>
- Nærland, T. U. (2020). The political significance of data visualization: Four key perspectives. In *Data Visualization in Society*. Amsterdam University Press.  
[https://doi.org/10.5117/9789463722902\\_ch04](https://doi.org/10.5117/9789463722902_ch04)
- Nau, C., Quan-Haase, A., & McCay-Peet, L. (2022). Defining Social Media and Asking Social Media Research Questions: How Well Does the Swiss Army Knife Metaphor Apply? In *The SAGE Handbook of Social Media Research Methods* (pp. 13-26). SAGE Publications London, UK.
- Newberry, C., & Macready, H. (1 de diciembre de 2022). *What is Social Listening, Why it Matters + 14 Tools To Help*. Hootsuite Blog. Recuperado en 2023-11-13 de <https://blog.hootsuite.com/social-listening-business/>
- Nixon, B. (2014). Toward a political economy of 'audience labour' in the digital era. *tripleC: Communication*. <https://www.triple-c.at/index.php/triplec/article/download/535/621>
- O'Neil, C. (2018). *Armas de destrucción matemática: Cómo el Big Data aumenta la desigualdad y amenaza la democracia*. Capitán Swing Libros.
- O'Reilly, T. (2005). *What Is Web 2.0 Design Patterns and Business Models for the Next Generation of Software*. Oreilly.com. Recuperado en 2023-10-10 de <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>

- Olshannikova, E., Olsson, T. T., Huhtamäki, J., & Kärkkäinen, H. (2017). Conceptualizing big social data. *Journal of Big Data*. <https://doi.org/https://doi.org/10.1186/s40537-017-0063-x>
- Papadakis, H., Papagrigoriou, A., Kosmas, E., Panagiotakis, C., Markaki, S., & Fragopoulou, P. (2023). Content-based recommender systems taxonomy. *Foundations of Computing and Decision Sciences*, 48(2), 211-241. <https://doi.org/10.2478/fcds-2023-0009>
- Pardo, N. G. (2007). *Cómo hacer análisis crítico del discurso: una perspectiva latinoamericana*. Universidad Nacional de Colombia.
- Parkin, S. (10 de junio de 2023). Who needs the Metaverse? Meet the people still living on Second Life. *The Guardian*. <https://www.theguardian.com/technology/2023/jun/10/who-needs-the-metaverse-meet-the-people-still-living-on-second-life>
- Peters, J. D. (2017). *Hablar al aire: una historia de la idea de comunicación*. Fondo de Cultura Económica.
- Pfeffer, J., Mooseder, A., Lasser, J., Hammer, L., Stritzel, O., & Garcia, D. (2023). *This Sample seems to be good enough! Assessing Coverage and Temporal Reliability of Twitter's Academic API*. In *Proceedings of the International AAAI Conference on Web and Social Media 17* (pp. 720-729). <https://ojs.aaai.org/index.php/ICWSM/article/download/22182/21961>
- Pfeiffer, S. (2022). *Digital Capitalism and Distributive Forces*. transcript Verlag. <https://doi.org/10.14361/9783839458938>
- Piñera, S. (20 de octubre de 2019). *Presidente Piñera: "Estoy seguro de que, con la unidad de todos los chilenos, vamos a derrotar a los violentistas y vamos a recuperar el país en paz y el país con libertad que todos queremos"*. Prensa Presidencia de Chile. Recuperado en 2023-08-21 de <https://prensa.presidencia.cl/comunicado.aspx?id=103689>
- Pinkowska, M. (19 de febrero de 2023). *A complete guide to social media analysis*. Brand24. Recuperado en 2023-11-12 de <https://brand24.com/blog/guide-to-social-media-analysis/>
- Pollach, I. (2005). Corporate self-presentation on the WWW: Strategies for enhancing usability, credibility and utility. *Corporate Communications: An International Journal*, 10(4), 285-301. <https://doi.org/10.1108/13563280510630098/full/html>

- Postigo, H. (2016). The socio-technical architecture of digital labor: Converting play into YouTube money. *New Media & Society*, 18(2), 332-349.  
<https://doi.org/10.1177/1461444814541527>
- Puebla, D. (23 de diciembre de 2019). *Decano de Universidad Católica e informe 'Big Data': 'Está lejos de ser de inteligencia'*. Radio Futuro. Recuperado en 2023-08-25 de  
<https://www.futuro.cl/2019/12/decano-de-universidad-catolica-e-informe-big-data-esta-lejos-de-ser-de-inteligencia>
- Puschmann, C. (2019). An end to the wild west of social media research: A response to Axel Bruns. *Information, Communication & Society*, 22(11), 1582-1589.  
<https://doi.org/10.1080/13645579.2019.1682840>
- Puschmann, C., & Burgess, J. (2014). Metaphors of big data. *International Journal of Communication*, 8, 1690-1709. <https://eprints.qut.edu.au/73094>
- Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023). *Robust speech recognition via large-scale weak supervision*. In *International Conference on Machine Learning* (pp. 28492-28518). PMLR.  
<https://proceedings.mlr.press/v202/radford23a.html>
- Raley, R. (2013). Dataveillance and Countervailance. In L. Gitelman (Ed.), *"Raw Data" Is an Oxymoron* (pp. 121-145). MIT Press.
- Rettberg, J. W. (2020). Ways of knowing with data visualizations. In *Data visualization in society* (pp. 35-48). Amsterdam University Press Amsterdam.  
<https://library.oapen.org/bitstream/handle/20.500.12657/22273/9789048543137.pdf?sequence=1#page=36>
- Ricaurte, P. (2019). Data epistemologies, the coloniality of power, and resistance. *Television & New Media*, 20(4), 350-365.  
<https://journals.sagepub.com/doi/abs/10.1177/1527476419831640>
- Richtel, M. (22 de octubre de 2011). A Silicon Valley School That Doesn't Compute. *The New York Times*. <https://www.nytimes.com/2011/10/23/technology/at-waldorf-school-in-silicon-valley-technology-can-wait.html>

- Robinson, A. (5 de diciembre de 2019). ¿Quién incendió América Latina? *La Vanguardia*.  
<https://www.lavanguardia.com/internacional/20191206/472070690924/america-latina-protestas-incidentes-conspiraciones-chile-venezuela-colombia-ecuador.html>
- Rogers, R. (2013). *Digital Methods*. MIT Press.
- Rutherford, D. (1995). *Routledge Dictionary of Economics*. Routledge.
- Sadin, É. (2018). *La siliconización del mundo*. Caja Negra.
- Sadowski, J. (2020). *Too Smart: How Digital Capitalism is Extracting Data, Controlling Our Lives, and Taking Over the World*. MIT Press.
- Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). *Earthquake shakes twitter users: real-time event detection by social sensors*. In *Proceedings of the 19th International Conference on World Wide Web* (pp. 851-860). <https://dl.acm.org/doi/abs/10.1145/1772690.1772777>
- Saleh, F. (27 de marzo de 2018). *Jorge Selume el señor big data a cargo de la Secom*. El Mostrador. Recuperado en 2023-08-23 de  
<https://www.elmostrador.cl/noticias/pais/2018/03/27/jorge-selume-el-senor-big-data-a-cargo-de-la-secom>
- Sánchez Bella, I., Hera, A. D. L., & Díaz Rementería, C. (1992). *Historia del derecho indiano*.
- Satariano, A., & Mozur, P. (7 de febrero de 2023). The People Onscreen Are Fake. The Disinformation Is Real. *The New York Times*.  
<https://www.nytimes.com/2023/02/07/technology/artificial-intelligence-training-deepfake.html>
- Savage, M., & Burrows, R. (2007). The coming crisis of empirical sociology. *Sociology*, 41(5), 885-899. <https://journals.sagepub.com/doi/abs/10.1177/0038038507080443>
- Schaffar, W. (2016). New social media and politics in Thailand: The emergence of fascist vigilante groups on Facebook. *ASEAS-Austrian Journal of South-East Asian Studies*, 9(2), 215-234. [https://www.ssoar.info/ssoar/bitstream/handle/document/62850/ssoar-aseas-2016-2-schaffar-New\\_Social\\_Media\\_and\\_Politics.pdf?sequence=1&isAllowed=y](https://www.ssoar.info/ssoar/bitstream/handle/document/62850/ssoar-aseas-2016-2-schaffar-New_Social_Media_and_Politics.pdf?sequence=1&isAllowed=y)
- Schiffer, Z., & Newton, C. (14 de febrero de 2023). *Yes, Elon Musk created a special system for showing you all his tweets first*. Platformer. Recuperado en 2023-11-01 de  
<https://www.platformer.news/p/yes-elon-musk-created-a-special-system>
- Schiller, D. (1999). *Digital Capitalism: Networking the Global Market System*. MIT Press (MA).

- Schiller, H. I. (2019). *Revival: Communication and Cultural Domination (1976)*. Routledge.
- Scholz, T. (2013). *Digital Labor*. Routledge.
- Schroeder, R. (2018). *Social Theory after the Internet*. UCL Press.
- Shane, S., & Isaac, M. (21 de septiembre de 2017). Facebook to Turn Over Russian-Linked Ads to Congress. *The New York Times*.  
<https://www.nytimes.com/2017/09/21/technology/facebook-russian-ads.html>
- Senado de Chile (4 de octubre de 2020). *Comisión de DD.H.H. revisa cifras a un año del estallido social*. Senado.cl. Recuperado en 2023-08-21 de  
<https://www.senado.cl/noticias/carabineros/comision-de-dd-h-h-revisa-cifras-a-un-ano-del-estallido-social>
- Sevignani, S. (2013). The commodification of privacy on the Internet. *Science and Public Policy*, 40(6), 733-739. <https://academic.oup.com/spp/article-abstract/40/6/733/1619824>
- Sheldon, P., Rauschnabel, P., & Honeycutt, J. M. (2019). *The Dark Side of Social Media*. Academic Press.
- Shepherd, T. (2013). Gendering the commodity audience in social media. In *The Routledge companion to media & gender* (pp. 157-167). Routledge.  
<https://prism.ucalgary.ca/bitstreams/3ecdc43d-b7e1-4ccb-a963-67dc2cae1079/download>
- Shklovski, I., Vertesi, J., & Troshynski..., E. (2009). The commodification of location: dynamics of power in location-based systems. *Proceedings of the 11th international conference on Ubiquitous computing*. <https://doi.org/10.1145/1620545.1620548>
- Sichel, S. (2021). *Programa de Gobierno 2022-2026. Las personas al centro*.
- Siebert, F. (23 de diciembre de 2019). 'En base a nuestros análisis, no creo que el estallido social en Chile sea producto de la manipulación de redes sociales, ni interna ni externamente'. Facultad de Ciencias Sociales. Recuperado en 2023-08-25 de  
<https://facso.uchile.cl/noticias/160227/no-creo-que-el-estallido-social-sea-producto-de-la-manipulacion>
- Siebert, B. (2015). *Cultural Techniques: Grids, Filters, Doors, and Other Articulations of the Real*. Fordham Univ Press.
- Sloan, L., Morgan, J., Housley, W., Williams, M., Edwards, A., Burnap, P., & Rana, O. (2013). Knowing the tweeters: Deriving sociologically relevant demographics from Twitter.



*Sociological research online*, 18(3), 74-84.

<https://journals.sagepub.com/doi/abs/10.5153/sro.3001>

Smythe, D. W. (2006). On the audience commodity and its work. In M. G. Durham & D. M.

Kellner (Eds.), *Media and cultural studies: Keywords* (pp. 230-256). Blackwell New York.

Snee, H., Hine, C., Morey, Y., Roberts, S., & Watson, H. (2015). *Digital Methods for Social Science*. Palgrave Macmillan.

Spiryn, K. (27 de octubre de 2023). *What is Social Media Intelligence (SOCMINT)?* Brand24.

Recuperado en 2023-11-12 de <https://brand24.com/blog/what-is-social-media-intelligence/>

Srnicek, N. (2018). *Capitalismo de Plataformas*. Caja Negra.

Stieglitz, S., & Dang-Xuan, L. (2013). Social media and political communication: a social media analytics framework. *Social network analysis and mining*, 3(4), 1277-1291.

<https://link.springer.com/article/10.1007/s13278-012-0079-3>.

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics—Challenges in topic discovery, data collection, and data preparation. *International journal of information management*, 39, 156-168.

<https://www.sciencedirect.com/science/article/pii/S0268401217308526>

Terranova, T. (2000). Free labor: Producing culture for the digital economy. *Social text*.

<https://muse.jhu.edu/article/31873/summary>

Thatcher, J. E., & Dalton, C. M. (2021). *Data Power: Radical Geographies of Control and Resistance*. Pluto Press (UK).

Thrift, N. (2006). Re-inventing invention: new tendencies in capitalist commodification.

*Economy and society*. <https://doi.org/10.1080/03085140600635755>

TikTok (s. f.). *Research API*. TikTok for Developers. Recuperado en 2023-10-31 de

<https://developers.tiktok.com/products/research-api/>

Timcke, S. (2017). *Capital, State, Empire*. University of Westminster Press.

Timcke, S. (2021). *Algorithms and the End of Politics*. Bristol University Press.

Tran, T. (3 de marzo de 2020). *What is Social Listening, Why it Matters, and 10 Tools to Make it Easier*. Hootsuite Blog. Recuperado en 2021-08-24 de <https://blog.hootsuite.com/social-listening-business/>

- Tringali, G. (9 de noviembre de 2016). *Alto Data Analytics, 'preguntamos a los datos'*. Cuatro. Recuperado en 2023-08-23 de [https://www.cuatro.com/blogs/hablemosdeempresa/alto-data-analytics-preguntamos-datos\\_6\\_2272815010.html](https://www.cuatro.com/blogs/hablemosdeempresa/alto-data-analytics-preguntamos-datos_6_2272815010.html)
- Trottier, D. (2015). Coming to terms with social media monitoring: Uptake and early assessment. *Crime, Media, Culture: An International Journal*, 11(3), 317-333. <https://doi.org/10.1177/1741659015593390>
- Tufekci, Z. (2017). *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. Yale University Press.
- Twitter (16 de julio de 2020). “Presentamos una nueva y mejorada API de Twitter. Recuperado en 2023-10-31 de [https://blog.twitter.com/es\\_la/topics/product/2020/Presentamos-una-nueva-y-mejorada-API-de-Twitter](https://blog.twitter.com/es_la/topics/product/2020/Presentamos-una-nueva-y-mejorada-API-de-Twitter)
- Twitter (2023a). *Getting Started with the Twitter API*. Recuperado en 2024-03-20 de <https://developer.twitter.com/en/docs/twitter-ads-api/getting-started>
- Twitter (2023b). *Preguntas frecuentes sobre la ubicación que se incluye en los Tweets*. Ayuda de Twitter. Recuperado en 2023-08-22 de <https://help.twitter.com/es/safety-and-security/tweet-location-settings>
- Twitter (s. f.). *Search Tweets: Standard v1.1*. Twitter. Recuperado en 2023-10-31 de <https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/overview>
- Uddin, M. F., & Gupta, N. (2014). *Seven V's of Big Data understanding Big Data to extract value*. In *Proceedings of the 2014 zone 1 conference of the American Society for Engineering Education* (pp. 1-5). IEEE. <https://ieeexplore.ieee.org/abstract/document/6820689>
- Unesco. (2023). *Directrices para la gobernanza de las plataformas digitales: Salvaguardar la libertad de expresión y el acceso a la información con un enfoque de múltiples partes interesadas*. UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000387360>
- Unión Europea (2022). *Reglamento de Servicios Digitales*. EUR-Lex. Recuperado en 2023-10-20 de <https://eur-lex.europa.eu/ES/legal-content/summary/digital-services-act.html>
- ur Rehman, I. (2019). Facebook-Cambridge Analytica data harvesting: What you need to know. *Library Philosophy and Practice (e-journal)*, 2497, 1-11.

- Vadoya, S. (28 de octubre de 2019). Policía identifica a uno de los autores de incendios en estaciones de Metro. *La Tercera*. <https://www.latercera.com/nacional/noticia/policia-identifica-uno-los-autores-incendios-estaciones-metro/880381/>
- van der Vlist, F. N., Helmond, A., Burkhardt, M., & Seitz, T. (2022). API Governance: The Case of Facebook's Evolution. *Social Media + Society*, 8(2), 205630512210862. <https://doi.org/10.1177/20563051221086228>
- van Dijck, J. (2013). *The Culture of Connectivity*. Oxford University Press.
- van Dijck, J., Poell, T., & de Waal, M. (2018). *The Platform Society*. Oxford University Press.
- van Dijk, T. A. (2001). Critical discourse analysis. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The Handbook of Discourse Analysis* (pp. 352-371). Blackwell.
- van Dijk, T. A. (2003). La multidisciplinaridad del análisis crítico del discurso: un alegato en favor de la diversidad. In R. Wodak & M. Meyer (Eds.), *Métodos de análisis crítico del discurso* (pp. 143-177). Gedisa.
- Varoufakis, Y. (2024). *Tecnofeudalismo: El sigiloso sucesor del capitalismo*. Deusto.
- Violani, N., & Riesco, J. (24 de diciembre de 2019). Big Data: las gerstiones de Selume y Díaz para conseguir un informe. *La Segunda*.
- Wang, Z., & Ye, X. (2018). Social media analytics for natural disaster management. *International Journal of Geographical Information Science*, 32(1), 49-72. <https://doi.org/10.1080/13658816.2017.1367003>
- Waring, H. Z. (2017). *Discourse Analysis: The Questions Discourse Analysts Ask and How They Answer Them*. Taylor & Francis.
- Wasko, J., Murdock, G., & Sousa, H. (2014). *The Handbook of Political Economy of Communications*. John Wiley & Sons.
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- Wiencierz, C., & Röttger, U. (2019). Big data in public relations: A conceptual framework. *Public relations journal*, 12(3).
- Wiesenberg, M., Zerfass, A., & Moreno, A. (2017). Big data and automation in strategic communication. *International journal of strategic communication*, 11(2), 95-114. <https://doi.org/10.1080/1553118X.2017.1285770>

- Wiggers, K. (6 de agosto de 2020). *Researchers discover evidence of gender bias in major computer vision APIs*. VentureBeat. Recuperado en 2023-10-31 de <https://venturebeat.com/business/researchers-discover-evidence-of-gender-bias-in-major-computer-vision-apis/>
- Williams, A., & Gilbert, J. (2022). *Hegemony Now: How Big Tech and Wall Street Won the World (And How We Win it Back)*. Verso Books.
- Williams, M. L., Burnap, P., & Sloan, L. (2017). Towards an ethical framework for publishing Twitter data in social research: Taking into account users' views, online context and algorithmic estimation. *Sociology*, 51(6), 1149-1168. <https://journals.sagepub.com/doi/full/10.1177/0038038517708140>
- Williams, S. M., & Pei, C.-A. H. W. (1999). Corporate social disclosures by listed companies on their web sites: An international comparison. *The International Journal of Accounting*, 34(3), 389-419. <https://www.sciencedirect.com/science/article/pii/S0020706399000163>
- WIRED (30 de marzo de 2023). *Twitter actualiza los precios de su API*. Recuperado en 2024-03-22 de <https://es.wired.com/articulos/twitter-actualiza-los-precios-de-su-api>
- Wodak, R. (2014). Critical discourse analysis. In *The Routledge companion to English studies* (pp. 302-316). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315852515-22/critical-discourse-analysis-ruth-wodak>
- X (s. f.). *About your For you timeline on X*. X. Recuperado en 2023-10-18 de <https://help.twitter.com/en/using-x/x-timeline>
- Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112(4), 1036-1040. <https://www.pnas.org/doi/full/10.1073/pnas.1418680112>
- Yu, S., & Guo, S. (2016). *Big Data Concepts, Theories, and Applications*. Springer.
- Zhu, Y., & Chen, H. (2015). Social media and human need satisfaction: Implications for social media marketing. *Business horizons*, 58(3), 335-345. <https://www.sciencedirect.com/science/article/pii/S0007681315000075>
- Zuboff, S. (2018). *The Age of Surveillance Capitalism*. PublicAffairs.

## ANEXOS

### **1. Documentos públicos disponibles relacionados con el informe Big Data**

- A. Ordinario 6282 de la subsecretaría del Interior, con fecha 25 de febrero de 2020, en el que el entonces subsecretario, Juan Francisco Galli, responde a un requerimiento (AB001C0004308) de la Contraloría General de la República (CGP). Allí, se acoge a la Ley N° 19.974 para no transparentar el informe Big Data.
  
- B. Solicitud AB001C0004308 de la CGP dirigida al ministro del Interior Gonzalo Blumel. En referencia a una solicitud no fue respondida realizada a ese ministerio por parte de un particular acogiéndose a la Ley de Transparencia, la CGP requiere al ministro la entrega de una copia del informe Big Data, además de información sobre los organismos que lo elaboraron y sus costos.
  
- C. Solicitud a la CGP por parte de un particular requiriendo una copia del informe Big Data, acusando no respuesta por parte del ministerio del Interior.
  
- D. Respuesta del director de la Agencia Nacional de Inteligencia (ANI) a la solicitud AB003W0000080 que requería, a través de la Ley de Transparencia, una copia del informe Big Data. Gustavo Jordán se acoge a la ley 19.974 (que crea la ANI) para negar la entrega de esta información.
  
- E. Dictamen de Contraloría E37908N20 en respuesta a requerimiento del senador Pedro Araya por el informe Big Data. Aludiendo a la calidad de muestra promocional gratuita del informe de Alto Analytics, lo que implica que no se incurrió en contrataciones del Estado, y señalando nuevamente la ley 19.974, que transforma en secretos los antecedentes en poder de la ANI y garantiza el derecho de sus funcionarios a mantener en secreto la identidad de sus fuentes, la CGP da por justificadas la negativa del ministerio del Interior a ofrecer mayores antecedentes sobre el informe Big Data.

Estos documentos, en formato PDF, pueden descargarse del siguiente link:

<https://is.gd/GdA7du>

## **2. Registro completo de las unidades significativas consideradas en el análisis**

Se encuentra disponible una planilla con el registro completo de las unidades significativas, distribuidas de acuerdo a las categorías discutidas, que fueron consideradas en el análisis crítico del discurso de los sitios web corporativos de las compañías estudiadas.

El documento, en formato xlsx (para Microsoft Excel), puede descargarse del siguiente link: <https://is.gd/aIndOp>