

Facultad de Ciencia Política y Relaciones Internacionales

Licenciatura en Ciencia Política - Tesina de Grado



Big data y políticas públicas: Un estudio exploratorio sobre el uso de los datos masivos en el sector público argentino

Autora: Rocío de Frutos

Legajo: D-1151/7

Mail: roudf@hotmail.com

Octubre 2021 Rosario

Director: Mg. Lucas S. Massuco

AGRADECIMIENTOS:

A mamá, a mis amigxs (que son todo) y a la *familia*.

Por la insistencia, el aguante, el cariño y la confianza que a veces yo misma no me tengo.

Especialmente a Roque, Marito, Nacho y Fede, por entender y tirarme para adelante.

A Massu, por estar desde los inicios y siempre sumar.

A Mike, por las charlas, los gráficos y la carátula.

A Mauri, Ivito y el 2k por las horas de coworking y concentración.

A todxs los que me regalaron algo de su tiempo, interés o conocimiento para que pudiera concretar este recorrido.

INDICE

Introducción	3
1. Datos Masivos y política pública	6
La revolución de los datos masivos	6
¿Qué es (y qué no es) Big data?	7
¿Cuáles son los métodos para procesar datos masivos?	9
Problemas públicos, datos y decisiones.....	14
Ciclo de políticas públicas, el momento de tomar decisiones.....	17
Marco metodológico.....	19
2. Big data como herramienta en la política pública Argentina	20
¿Cómo se generan los datos?	23
Institucionalidad construida en el uso de datos masivos	28
Big Data en el nivel Nacional	28
Observatorio Nacional de Big Data	28
Big data en el nivel provincial.....	29
Rosario, Provincia de Santa Fe: Movilidad de los ciclistas	29
Córdoba, Provincia de Córdoba: Movilidad y transporte público	30
Neuquén, Provincia de Neuquén: Movilidad en el Área Metropolitana	31
Bahía Blanca, Provincia de Buenos Aires: ¿Qué pasa Bahía Blanca?	33
Área Metropolitana de Buenos Aires: Movilidad y exclusión social	33
Ciudad Autónoma de Buenos Aires: Movilidad de los usuarios del Ferrocarril y Subte.....	34
Intervención en el Ciclo de Políticas Públicas	36
Matriz síntesis: Momento de intervención en el Ciclo de Políticas Públicas	38
3. Desafíos y potencialidades	40
Desafíos y límites: ¿qué nos falta?	40
Calidad de los datos	42
Resguardo y privacidad de los datos	43
Capacidades estatales.....	46
Oportunidades, potencialidades y condiciones necesarias	48
4. Hallazgos e hipótesis	50
Referencias bibliográficas	52
Anexo	56
Decreto de creación de Observatorio Nacional de Big Data	56
Mapa del Estado (organigrama): Jefatura de Gabinete de Ministros de la Nación Argentina ..	61
Ley 25.326: Protección de Datos Personales	62

Introducción

En la última década se ha popularizado el término Big Data (datos masivos), expresión que refiere al fenómeno de producción y procesamiento de una cantidad ingente de datos que las personas generamos constantemente, de forma voluntaria o involuntaria, a través de dispositivos electrónicos interconectados y sensores, dejando un rastro digital que permite su monitorización.

En su libro “Big data. La revolución de los datos masivos”, Mayer-Schönberger y Cukier (2013) señalaban que en 2009 la última posible pandemia del virus H1N1 fue exitosamente frenada gracias al sistema Google Flu Trends¹ que, puesto al servicio del gobierno de Estados Unidos, procesaba los datos del motor de búsqueda prediciendo con exactitud cómo se producía la expansión geográfica del virus. La conclusión de los autores a raíz de este episodio fue: “Con ellos [los datos], cuando se produzca la próxima pandemia, el mundo dispondrá de una herramienta mejor para predecir, y por ende prevenir, su propagación”(p.8).

Observando la evolución del fenómeno con el diario del 2020, comprendemos que si bien los datos masivos son una herramienta fundamental para la toma de decisiones (que ha crecido y aumenta su potencial de uso y predicción en casi todos los aspectos de nuestras vidas), los autores pueden haber sobreestimado el peso de la información subestimando la complejidad del proceso de toma de decisiones que atraviesa el ciclo de las políticas públicas. Los datos y la evidencia son fundamentales, pero en conjunto con la persuasión y la búsqueda de consensos, elementos cruciales para alcanzar y viabilizar definiciones políticas. De acuerdo con Giandomenico Majone (1997) “la

¹ En español: Tendencias de Google sobre la gripe. Este software halló una combinación de cuarenta y cinco términos de búsqueda que, procesada por un modelo matemático, presentaba una correlación entre su predicción y las cifras oficiales de la enfermedad a lo largo del país. El modelo fue diseñado tomando las estadísticas de los Centros de Control y Prevención de Enfermedades sobre la propagación de las gripes de 2003 y 2008, comparándolas con los millones de términos de búsqueda consultados por los usuarios. De este modo, el sistema permitía identificar el avance del virus en tiempo real y por regiones específicas. Google contaba con la ventaja de ser el principal buscador y poseer una enorme cantidad de datos para analizar sobre los posibles síntomas que *googlean* los estadounidenses.

política pública está hecha de palabras. En forma escrita u oral, la argumentación es esencial en todas las etapas del proceso de formación de políticas”.

El acelerado avance en la digitalización de la información, característica fundamental de lo que va del siglo XXI, en combinación con el acceso masivo a nuevas tecnologías (teléfonos y objetos inteligentes, computadoras portátiles, internet, etc) y la necesidad apremiante que creó la pandemia de Covid-19 de tener todo disponible de modo virtual, ha extendido la creencia de que el big data no sólo es un destino ineludible de la humanidad sino, en algunos casos, el reservorio infinito donde se encuentran todas las respuestas que necesitamos.

Algo de cierto guardan estas creencias, mientras en el año 2000 sólo una cuarta parte de los datos se encontraban almacenados en formato digital, para 2007 sólo el 7% de los datos eran analógicos, incluyendo libros, fotografías, películas y música (Mayer-Schönberger & Cukier, 2013, p.15). Una característica aparejada a este fenómeno, es el crecimiento del volumen de datos en sí mismos, de acuerdo con Walter Sosa Escudero (2019), entre 2017 y 2019 “en todo el mundo hemos creado más datos que en toda la historia de la humanidad” (p. 35). Los datos digitales se duplican en pocos años, mientras la información analógica apenas presenta crecimiento. Según estimaciones de Mayer-Schönberger & Cukier, en 2013 la cantidad total de información almacenada en el mundo era de 1200 exabytes², de los cuales menos del 2% eran no digitales (p.16).

De esta situación se desprenden una serie de interrogantes: ¿Qué repercusiones tiene el optimismo sobre el potencial de los datos masivos en la gestión de las políticas públicas? ¿Cómo se insertan de hecho los datos masivos en el proceso de políticas públicas? ¿En qué fase del ciclo predominan? ¿Qué límites y potencialidades tiene el uso de big data en el ciclo de políticas públicas de Argentina? ¿Cuáles son las condiciones necesarias para que big data cumpla un rol efectivo en la política pública? ¿Cuánto valen los datos sin las capacidades estatales para abordarlos?

² Esta cantidad de datos equivale a 1.200 millones de Terabytes, siendo la capacidad de almacenamiento de una computadora de escritorio promedio, en 2021, de aproximadamente 1 Terabyte de disco rígido.

Este trabajo final integrador se propone como **objetivo general** explorar el fenómeno del big data en experiencias de la política pública argentina, analizando sus potencialidades y límites, como también los desafíos y fantasías que se imponen en torno a ella. Los **objetivos específicos** que persigue son:

- Identificar y describir las principales características del big data.
- Identificar y describir los principales usos de los datos masivos por parte de los Estados locales, provinciales y el nacional en Argentina.
- Identificar las potencialidades y los desafíos del fenómeno en el contexto de nuestro país.

1. Datos Masivos y política pública

La revolución de los datos masivos

Viktor Mayer-Schönberger y Keneth Cukier (2013), para referirse al fenómeno de los datos masivos, hablan de una revolución, es decir, un cambio brusco y radical del orden establecido de las cosas, que ya no tiene vuelta a lo anterior, el mundo ha cambiado de forma definitiva. Lo que se ha modificado profundamente son los modos de vida y el valor que tienen los datos para las empresas, como así también para los Estados y la sociedad en su conjunto.

De acuerdo con Javier Gutierrez Puebla (2018), es preciso analizar esta revolución en el contexto de las Smart Cities, ciudades sensorizadas que se conciben como “constelaciones de instrumentos conectadas a través de múltiples redes que proporcionan datos continuos sobre los movimientos de personas y mercancías y sobre el estado de estructuras y sistemas”(p.198). Oscar Oszlak (2020) denomina a este proceso de profundos cambios científicos y tecnológicos como “era exponencial”. Hay quienes se refieren también a una cuarta revolución industrial, caracterizada por la transición hacia nuevas infraestructuras digitales: sistemas ciberfísicos basados en la automatización, la robotización, la inteligencia artificial y el propio Internet de las cosas³.

En la actualidad, internet se ha convertido en un gran repositorio de datos de lo que la gente dice, hace, compra y busca, y de cómo unos conectan con otros. Estos servicios, en apariencia gratuitos (nadie paga por utilizar el buscador de Google, calcular una ruta entre dos puntos con Google Maps o utilizar las principales redes sociales), generan para las compañías sus principales activos en base a los enormes volúmenes de datos que las personas les ceden al utilizarlos, ya sean producidos por máquinas (múltiples redes de sensores y dispositivos que registran la huella digital de la actividad humana, generalmente de forma pasiva) o generados activamente por los usuarios en Internet.

³ En Inglés: Internet of Things (IoT). Este concepto se refiere a una interconexión digital de objetos cotidianos con internet que genera un ecosistema digital donde circulan los datos. El 17 de marzo de 2021 tuvo lugar por primera vez en nuestro país la Mesa Nacional de Internet de las Cosas, que tiene como objetivo impulsar el desarrollo del ecosistema en nuestro país junto al sector público y el sector privado. Recuperado el 26 de octubre de 2021, de:
<https://www.argentina.gob.ar/noticias/lanzamiento-de-la-mesa-nacional-de-internet-de-las-cosas>

Para dimensionar la magnitud de los volúmenes de datos, alcanza con decir que en 2017 cada minuto en todo el mundo se enviaban 156 millones de correos electrónicos y 452.000 tweets, se realizaban 3,5 millones de búsquedas y se descargaban 342.000 apps⁴.

Es importante, sin embargo, no reducir la “revolución de los datos masivos” sólo a los avances y desarrollos en las tecnologías de recolección, almacenamiento y procesamiento (analítica de datos), considerando que lo que se modificó fundamentalmente también fue la perspectiva sobre los posibles usos de los datos para la toma de decisiones estratégicas, creando así un nuevo paradigma donde los datos dejaron de verse estáticos y “se convirtieron en la materia prima del negocio, en un factor vital, capaz de crear una nueva forma de valor económico” (Mayer-Schönberger & Cukier, 2013, p.11).

¿Qué es (y qué no es) Big data?

El término big data define al volumen y tipo de datos producidos por la interacción de dispositivos inteligentes interconectados: computadoras, celulares, tarjetas de crédito, relojes inteligentes, cajeros automáticos y “cualquier objeto capaz de producir información y enviarla electrónicamente a otra parte” (Sosa Escudero, 2019, p.32).

Al referirse al *tipo* de datos, se busca diferenciar a los datos masivos, cuyo carácter es anárquico y espontáneo, de los datos recolectados de forma sistemática y estructurada, como es el caso de las encuestas. Esto significa que esos datos no son generados con el fin de producir información, sino como resultado de otra acción: realizar un pago, moverse de un lugar a otro, consultar un sitio web.

En 2001, Douglas Laney (analista de la consultora Gartner Chief Data Officer, considerado una de las figuras con mayor influencia en el área) resumió las principales nociones de este nuevo tipo de datos en *las tres V del Big data*: volume, velocity, variety (volumen, velocidad y variedad). Si bien con el tiempo se le han ido asignando nuevas

⁴ 2017, 22 de marzo ¿Qué sucede en internet cada minuto? De 900.000 accesos a Facebook a 3,5 millones de búsquedas de Google. El Economista. Recuperado el 26 de octubre 2021 en: <https://www.economista.es/tecnologia/noticias/8236038/03/17/Que-sucede-en-internet-cada-minuto-De-900000-accesos-a-Facebook-a-35-millones-de-busquedas-de-Google.html>

“V’s” para describirla, llegando hasta cuarenta y dos de acuerdo con Sosa Escudero (2019), a los fines de este trabajo desarrollaremos sólo las primeras tres:

- **Volumen** apunta al carácter masivo y exhaustivo de los datos, si bien la cantidad que se considera propiamente big data no se encuentra definida en ningún parámetro técnico oficial, algunas estimaciones refieren al tamaño de los archivos (se almacena en petabytes y exabytes⁵), el tipo de programas necesarios para procesarlos (exceden las capacidades de excel) y las capacidades de procesamiento necesarias en los ordenadores. Sumado a esto, los datos históricos no se borran, su valor reside en servir de apoyo para analizar procesos y realizar nuevas predicciones.
- **Velocidad** refiere a dos aspectos: por un lado, nos habla de que los datos se generan en tiempo real y se producen de forma acelerada por el mayor acceso de dispositivos electrónicos en todos los aspectos de nuestras vidas. La cantidad de datos crece cada vez más rápido. Por otro lado, la velocidad también se trata de la posibilidad de almacenar y procesar una gran cantidad de datos de forma inmediata, gracias al avance de las herramientas necesarias para el análisis de los mismos (la memoria de los ordenadores).
- **Variedad** define la heterogeneidad de los posibles tipos de datos que pueden recolectarse (minería de datos), desde estructurados (por ejemplo: una distancia recorrida captada por un GPS o una transacción bancaria) a no estructurados (un tweet

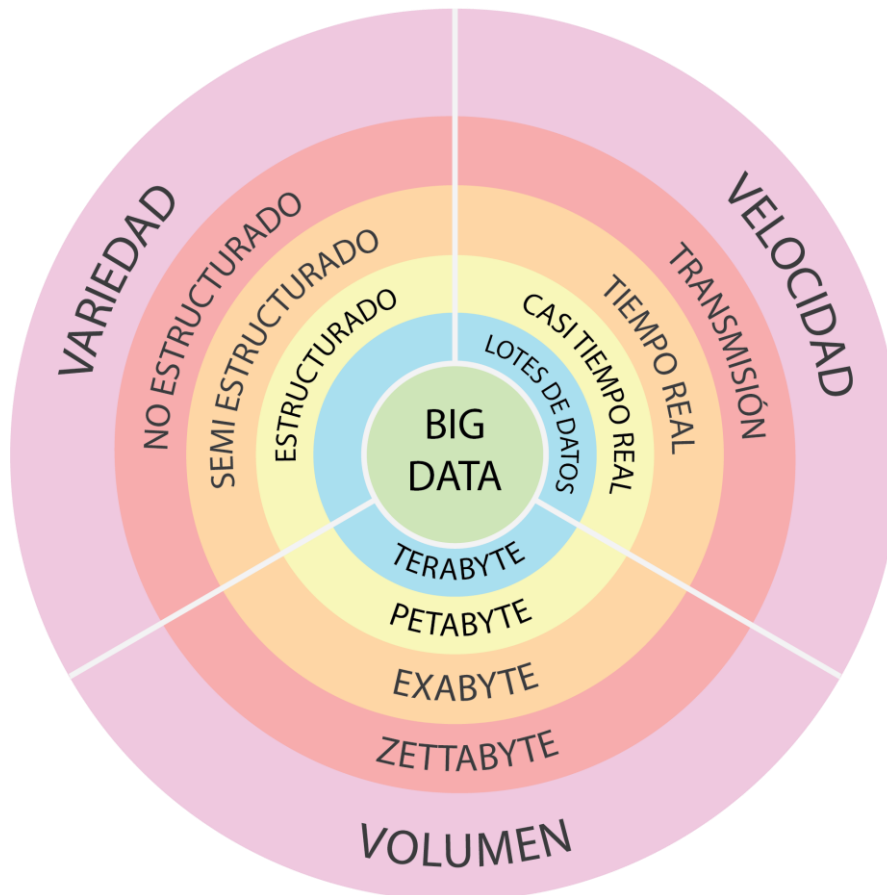
⁵ La métrica de las diferentes unidades de almacenamiento se basan en el Byte y se utilizan para indicar tanto la capacidad de almacenamiento como el tamaño de los archivos. A modo de ejemplo, las tarjetas SD o los USB suelen medirse en Gigabytes, mientras que los discos rígidos suelen hacerlo en Terabytes.

MÉTRICA	VALOR	BYTES
BYTE (B)	1	1
KILOBYTE (KB)	1.024 ¹	1.024
MEGABYTE (MB)	1.024 ²	1.048.576
GIGABYTE (GB)	1.024 ³	1.073.741.824
TERABYTE (TB)	1.024 ⁴	1.099.511.627.776
PETABYTE (PB)	1.024 ⁵	1.125.899.906.842.624
EXABYTE (EB)	1.024 ⁶	1.152.921.504.606.846.976
ZETTABYTE (ZB)	1.024 ⁷	1.180.591.620.717.411.303.424
YOTTABYTE (YB)	1.024 ⁸	1.208.925.819.614.629.174.706.176

o una foto). Esta diversidad también responde a la fuente de recolección: archivos, bases de datos, sensores, etc.

Figura 1

Las tres V's del Big data



Nota: Adaptado de Las tres V's del Big data, de Gutiérrez Puebla et al., 2019, Cómo aplicar big data en la planificación del transporte urbano: el uso de datos de telefonía móvil en el análisis de la movilidad. Figura 2. p. 18

¿Cuáles son los métodos para procesar datos masivos?

El uso de datos masivos o *Ciencia de Datos* consiste en un conjunto de técnicas, provenientes de disciplinas como las ciencias de la computación, estadística, matemática y econometría, que permiten procesar y analizar grandes cantidades de datos digitales. Dentro de la ciencia de datos, el machine learning (o aprendizaje

automatizado) consiste en que los procesadores apliquen modelos estadísticos para identificar patrones que permitan predecir tendencias, inferir probabilidades o arrojar nuevas pistas (hallazgos o insights) sobre una cuestión. El objetivo es producir evidencia que sea pertinente, de calidad y oportuna, para así fundamentar y orientar decisiones (toma de decisiones guiadas por datos, o *data-driven decision making* en inglés).

Estos sistemas funcionan de manera más acertada a medida que incorporan más datos sobre los que basar sus hallazgos. Ante la creencia de que una mayor cantidad de datos implica una lectura más profunda de los fenómenos, asistimos a lo que se conoce como un proceso de datificación. Datificar un fenómeno significa plasmarlo en un formato cuantificado para que pueda ser tabulado y analizado. Esto es muy diferente de la digitalización que es el “proceso por el que se convierte la información analógica en los unos y ceros del código binario para que los ordenadores puedan manejarla” (Mayer-Schönberger & Cukier, 2013, p. 96).

Como se mencionó más arriba, este fenómeno despierta fervientes seguidores y detractores. Para algunos es un proceso positivo, mientras otros alertan sobre no confundir la datificación del mundo con su comprensión. Autores más escépticos del fenómeno, como Ignacio Maestro Cano (2019), afirman que “la obsesión por recopilar cualquier dato, va a generar gran cantidad de ruido e información redundante” (p.2). Para Elisenda Ardévol (2016) la relevancia del big data reside, no tanto en la cantidad, como en la complejidad de los datos que pueden manejarse y relacionarse. De acuerdo con ella “el simple acceso a grandes cantidades de datos no garantiza su inteligibilidad y correcta interpretación” (p.4).

Los científicos de datos afirman que el valor de los mismos se encuentra en la potencialidad de las preguntas que se le formulan. En este sentido, Maestro Cano (2019) señala que, en el mejor de los casos, big data ofrecería la respuesta a una infinidad de preguntas indiscriminadas que (quizás por algún motivo) jamás fueron formuladas desde un ámbito científico, cuando lo que a menudo se necesitan son pocas preguntas bien escogidas. (p.3).

Una idea que aparece recurrentemente entre los autores, es que los datos sueltos no aportan un valor en sí mismos. Para que un conjunto o set de datos se transforme en

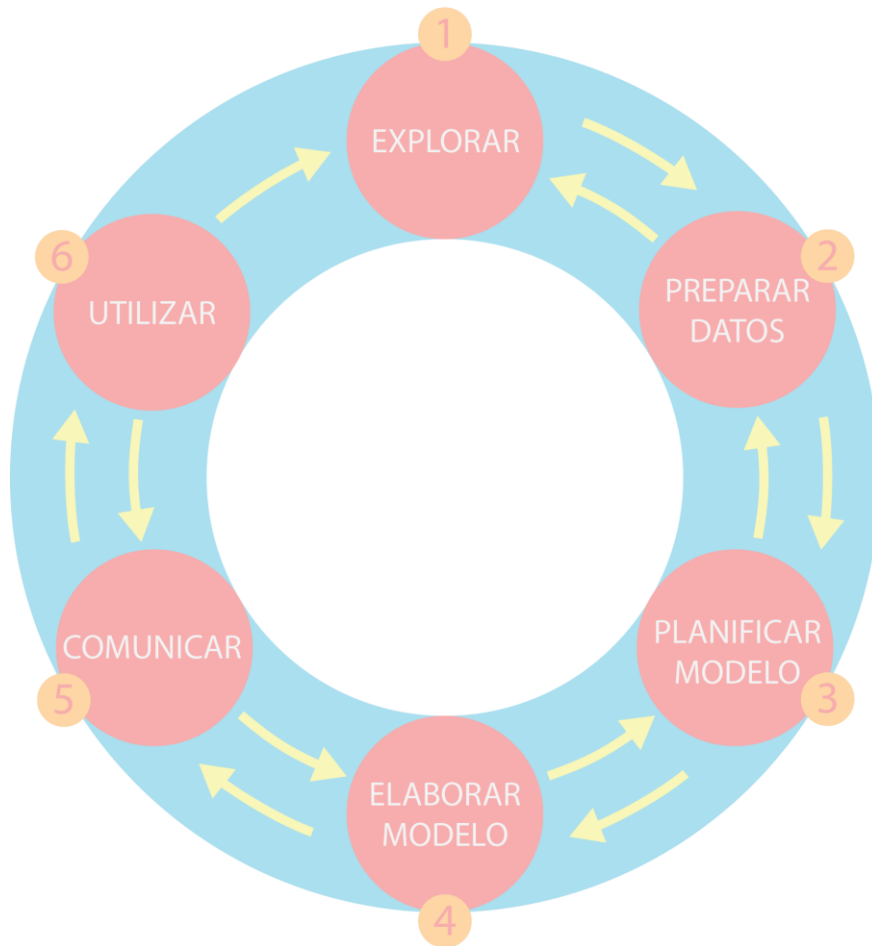
información es necesario depurarlo en base a preguntas específicas y técnicas analíticas que le den forma. Este proceso de producción de conocimiento, que algunos autores denominan como Ciclo de vida del análisis de datos (Rodríguez, P., Palomino, N., & Mondaca, J., 2017), puede desagregarse en 6 etapas:

1. **Exploración:** Explorar los datos disponibles formulando preguntas e hipótesis.
2. **Preparación:** Recopilar, limpiar y analizar la consistencia entre los datos.
3. **Modelizar - Planificar modelo:** Determinar variables explicativas y predictivas, seleccionar modelos y algoritmos a utilizar, definir métricas de desempeño.
4. **Modelizar - Elaborar modelo:** Implementar, ajustar y validar el modelo.
5. **Comunicar:** Interpretar resultados, elaborar visualizaciones y sugerir recomendaciones de mejora.
6. **Utilizar:** Tomar decisiones, definir metas y estándares, asignar recursos.

Este ciclo es un proceso iterativo, lo cual significa que es aplicado una y otra vez, pudiendo retroceder a etapas previas y reformular partes en función de la disponibilidad de nuevos datos o la reinterpretación de los mismos a la luz de nuevos hallazgos.

Figura 2

Ciclo de vida del análisis de datos masivos



Nota: Adaptado de Ciclo de vida del análisis de datos masivos, de Rodriguez, P., Palomino, N. & Mondaca, J., 2017, *El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe*. Figura 1. p.2

Los investigadores Arantza Coullat y Mario Tascón (2016) sostienen que para que los datos se conviertan en información es preciso añadirles valor mediante alguno de los siguientes procesos:

- Cálculo: los datos se procesan matemática o estadísticamente.
- Categorización: se conocen las unidades de medida que ayudan a interpretar los datos.
- Condensación: los datos se resumen en forma concisa (agregación).

- Contextualización: se aporta el contexto y propósito con el que se generan.
- Corrección: se eliminan errores o inconsistencias.

Para los autores, la información es la comunicación de conocimientos o inteligencia que sirven de marco para la incorporación de los datos, por lo tanto no se encuentra sólo en los documentos, sino también en los procesos, prácticas y normas (p.9).

Este planteo es similar al pensamiento de Oscar Oszlak en su libro “El Estado en la era exponencial”, quien destaca la distinción y relación entre datos, información y conocimiento ilustrada con la “pirámide de DICS”.

Figura 3
Pirámide de DICS



Nota: Adaptado de Pirámide de DICS, de Oszlak, O., 2020, *El Estado en la era exponencial*. p.47

La pirámide jerarquiza el proceso de reunir datos y analizar diversas fuentes de información para generar conocimiento, agregando en el escalón superior “la

sabiduría". El autor vincula la distinción entre conocimiento y sabiduría con la distinción que Platón realizaba entre doxa (conocimiento basado en opiniones sin evidencia) y episteme (conocimiento fundamentado y sistemático). Desde esta perspectiva, la distinción jerárquica de la pirámide destaca el valor de la aplicación del conocimiento y, fundamentalmente, la sabiduría a la tarea de gobernar (p.47).

Problemas públicos, datos y decisiones

Al pensar el proceso de toma de decisiones, resulta esencial establecer que no se emplean los mismos parámetros en el ámbito de las empresas que en la política pública. En el Estado los datos y la evidencia, componentes importantes y necesarios de toda definición política, no son los únicos elementos para traccionar un curso de acción.

Frente al inminente avance de los datos como el valuarte de nuestra era, es inevitable retomar a Carol Hirschon Weiss (1999) cuando señalaba que la investigación de políticas o cualquier otro tipo de investigación no determina la principal dirección de la política: "Los decisores, políticos y funcionarios poseen convicciones ideológicas y constelaciones de intereses, que en gran parte determinan el rumbo que ellos siguen" (p.378).

Aún contando con los datos concretos para responder a preguntas específicas, eso no siempre tiene efecto directo en las definiciones de un gobierno. Otros intereses pueden cobrar relevancia por sobre la información de las investigaciones y en algunos casos estas son utilizadas para respaldar lo que los actores ya han consensuado de antemano. En ocasiones, los datos sirven para ajustar y afinar sus propuestas adaptándolas a las decisiones tomadas, y no al revés: "Los mismos datos han aparecido en ambos lados de una cuestión, utilizados en apoyo de prescripciones radicalmente distintas" (Hirschon Weiss, 1999, p.380).

Giandomenico Majone (1997) sostiene que "cuando se intersectan la ciencia, la tecnología y la política entran en grave conflicto diferentes actitudes, perspectivas y reglas de la argumentación. Los criterios científicos de ver-dad chocan con las normas legales de las evidencias y con las nociones políticas de lo que constituye una base suficiente para la acción" (p.38).

El autor también traza una distinción entre la evidencia y los datos o la información. La evidencia es información seleccionada e introducida en cierto punto del argumento para persuadir a un auditorio particular. Datos o modelos poco apropiados, su introducción en un punto inadecuado del argumento o incorrectamente presentada para el auditorio al que se destina puede destruir la eficacia de la evidencia, más allá su valor cognoscitivo intrínseco (p.45).

En ciertos momentos, de acuerdo con Hirschon Weiss (1999), los datos pueden instalarse en forma de ideas, sirviendo de "ilustración" general para los elaboradores de las políticas. Es decir, la mayor parte de los aspectos metodológicos de la investigación son dejados de lado, dándole entidad a la idea por encima de lo que respaldan los datos. Al diluirse lo esencial de la investigación, no hay ninguna garantía de que las ideas sean correctas, su abreviación en consignas puede omitir elementos fundamentales, pero encontrarse instalada en el pensamiento convencional sobre una cuestión. De este modo, se moldean las suposiciones de la ciudadanía acerca de lo que es importante, de cómo abordarlo y de las soluciones que probablemente ayudarán a alcanzar los fines deseados. Muchas veces estas ideas quedan en la memoria social colectiva y resurgen a la luz de nuevos debates, como parte de los discursos políticos, más que en forma de acciones concretas.

Por último, la autora advierte sobre las investigaciones que se vuelven meros argumentos, es decir, que han derivado en una defensa o recomendación en favor de un curso de acción o una perspectiva sobre un tema. La investigación como argumentos sólo considera opciones dentro de la gama ideológica de quienes la concibieron. En este sentido, en materia de políticas, los datos tienen influencia en situaciones que cuentan con cierto consenso sobre valores y metas. También contribuyen a precisar el problema y aclarar sus parámetros, y servir de base para hacer buenas estimaciones de la eficiencia de los correctivos.

En consonancia con esta idea, Aguilar Villanueva (1993) afirma que los problemas son "construcciones", datos seleccionados y caracterizados con referencia a ciertos esquemas cognoscitivos y valorativos de los sujetos que los observan o experimentan, el problema público manifiesta toda su complejidad y variabilidad, particularmente en

el marco de sociedades abiertas, plurales y gobiernos democráticos. Esto puede comprenderse al observar el modo en que distintos gobiernos afrontan la crisis sanitaria, presentando un claro contraste entre países dentro de una misma región, como Brasil y Argentina.

Siguiendo a Majone (1997), la principal complejidad de pensar y precisar los problemas públicos, radica en que estos son “wicked problems”, es decir: problemas retorcidos, embrollados, tramposos. Que no encuentran una formulación definitiva, ni criterios que establezcan cuándo se alcanza la solución, o cuya solución no es nunca verdadera o falsa sino buena o mala y carece además de una prueba inmediata o resolutoria, problemas frecuentemente inéditos, sintomáticos de problemas de mayor trascendencia.

Aguilar Villanueva (1993) afirma que una decisión es “el resultado de la confluencia no casual, pero tampoco enteramente previsible y controlable de cuatro corrientes, cada una de las cuales tiene diversa lógica y dinámica: los problemas, las soluciones, los actores participantes y las oportunidades de elección” (p.39). El accionar gubernamental se encuentra sujeto a las posibilidades de elección dentro de una contingencia determinada, que requieren una cantidad de tiempo, recursos y capacidades estatales que permitan contrarrestar otras demandas que también resultan apremiantes.

Si tomamos como ejemplo la crisis desatada por la pandemia de Covid-19, podemos observar la complejidad de abordar un problema con la falta de experiencia, tecnologías adecuadas y frente a la magnitud de las medidas requeridas (cierre de fronteras, freno total de las actividades, etc), sumado a la cantidad de actores sociales, económicos y políticos que se ven afectados tanto por la definición de actuar, como de no hacerlo, de cada gobierno. Para quienes forman parte de la toma de decisiones, se ponen en juego numerosos factores al momento de intervenir en una cuestión: motivaciones políticas, morales, económicas, legales. Las decisiones, por lo tanto, se encuentran condicionadas por las limitaciones inescapables de la realidad (Aguilar Villanueva, 1993, p.40).

Algunas áreas dependen de forma más estricta de los datos estadísticos, ya que se ven compelidas a prestar atención a las cifras al momento de la definición política y la toma de decisiones, como es el caso de las medidas en materia de salud (observable en el seguimiento del número de casos infectados de Covid-19 para ajustar o flexibilizar las

restricciones de la cuarentena, así como las tasas de ocupación de las camas Unidad de Terapia Intensiva en los hospitales), la prevalencia del analfabetismo de adultos, o en economía, con las tasas de desempleo.

Sin embargo, como señala Majone (1997), el cálculo de soluciones óptimas o mejores dentro de las restricciones dadas es sólo la parte estática del análisis de factibilidad; la parte dinámica y más importante es el descubrimiento de instrumentos para expandir las fronteras de lo posible. Pero en este punto, las posibilidades están sujetas a lo que el sistema político considere justo o aceptable, por lo tanto, las políticas sólo pueden flexibilizarse cambiando las actitudes y los valores (p.43).

Ciclo de políticas públicas, el momento de tomar decisiones

Joan Subirats (1989) afirma que las políticas públicas son una secuencia de decisiones que implican una modificación de la realidad: decidir que existe un problema, decidir que debe intentar resolverse, decidir legislar sobre el tema, etc. Considerando las distintas decisiones como etapas de un ciclo más extenso, el autor propone un esquema de análisis con los siguientes pasos: percepción y definición del problema; intereses afectados; grado de organización; accesos a los canales representativos; consecución del estatuto propio de «tema» a incluir en el programa o agenda de actuación de los poderes públicos; formulación de una solución o de una acción de respuesta; establecimiento de objetivos y prioridades; soportes políticos, presupuestarios y administrativos de la solución propuesta; implementación o puesta en práctica de esa política; evaluación y control de los efectos producidos; mantenimiento, revisión o terminación de esa política (Subirats, 1989, p. 42).

De esta forma, el análisis de datos masivos funciona como guía y soporte en las distintas definiciones que se toman en cada una de estas etapas. Pensar las políticas como una sucesión de decisiones y posicionamientos nos brinda un marco para analizar el objetivo, el momento, la forma y condiciones bajo las cuales las administraciones públicas se sirven de este recurso. Como señala Sosa Escudero (2017), en contraste con el sector privado, el sector público es más cauto en incorporar esta herramienta, tanto por su esperable inelasticidad administrativa como por su compleja lógica de funcionamiento,

que debe internalizar cuestiones operativas, políticas y comunicacionales, así como movilizar voluntades, capacidades y recursos.

Marco metodológico

La metodología utilizada en la presente tesina es de tipo cualitativa, orientada a explorar y conocer el incipiente fenómeno de los datos masivos en las políticas públicas en Argentina. Las fuentes relevadas son secundarias: documentos, páginas web, desarrollos teóricos y notas técnicas de equipos y especialistas que han conducido o analizado experiencias vinculadas a la temática. Actualmente, big data se presenta como un concepto “de moda” del cual no existen, hasta el momento, definiciones tan exactas o tajantes sobre las dimensiones que la componen, ni uniformidad sobre el modo en que se desarrolla en el contexto de nuestro país, por lo tanto resulta primordial definir y caracterizar las propiedades del objeto en cuestión.

En el primer capítulo del trabajo desarrollamos una perspectiva teórica sobre qué son los datos masivos, sus características y el uso de estos en las definiciones que componen el ciclo de políticas públicas. En el segundo capítulo, revisaremos algunos casos ilustrativos llevados adelante por el Estado argentino en distintos niveles de gobierno, para observar en qué momento del ciclo se encuentra más presente esta herramienta y en qué nivel de institucionalidad es más comúnmente aplicada. En el tercer capítulo, analizaremos algunos desafíos y límites del uso de los datos masivos, vinculados específicamente con el contexto de nuestro país y con las particularidades y responsabilidades del uso de los mismos por parte del Estado, así como también sus potencialidades de crecimiento para nuevas propuestas y proyectos. Por último, el cuarto capítulo abordará las conclusiones e hipótesis, e intentará realizar algunos aportes en base a las impresiones recogidas en este recorrido exploratorio sobre el fenómeno analizado.

2. Big data como herramienta en la política pública Argentina

A lo largo de este capítulo revisaremos el estado del arte del big data en la política pública argentina. Para eso, realizaremos una aproximación al fenómeno, ejemplificando con distintas experiencias la forma en que el Estado en todos sus niveles se sirve de los de datos masivos en distintas fases del ciclo de políticas públicas: toma de decisiones, diseño, implementación, seguimiento o evaluación.

De acuerdo con Oscar Oszlak (2020), la transformación digital de la administración pública se ha convertido en una de las áreas de reforma estatal que más avances ha experimentado en lo que va de este siglo (p.132). Para el autor, estas herramientas son responsables de la gran mayoría de las transformaciones de los organismos públicos en América Latina en tanto son poderosas herramientas que permiten mejorar la eficacia y la eficiencia al agilizar procesos, simplificar trámites, fortalecer la transparencia y la rendición de cuentas.

Por su parte, Anthony Townsend y Patricio Zambrano-Barragán (2019) señalan que el uso del big data en las ciudades de América Latina y el Caribe aún no se encuentra generalizado pero hay un reconocimiento cada vez mayor de su potencial. Esta conclusión se funda en el análisis de distintos niveles de *madurez de big data urbana*, realizado en base a un relevamiento de 45 ciudades, y refiere a la adopción de cinco áreas funcionales:

1. los datos abiertos,
2. el cultivo de un ecosistema de datos,
3. la analítica,
4. la toma de decisiones basadas en datos,
5. los servicios urbanos digitales.

Como establecimos en el capítulo anterior, big data son los datos digitales que, preparados adecuadamente, permiten el procesamiento de la información mediante un

software. De acuerdo con Coullat y Tascón (2016), los datos son la unidad mínima, elementos primarios de información que por sí solos son irrelevantes y no suelen decir nada sobre el porqué de las cosas: un número de teléfono, un apellido, una hora, una coordenada (p.8). La utilización de los datos masivos como fuente de información presenta la ventaja de que la recolección de los mismos se realiza en forma pasiva sin que el usuario tenga que cargarlos (se registra lo que el usuario hace, no lo que dice que hace); al almacenarse de forma continua la información se mantiene siempre actualizada, permitiendo monitorizar procesos y analizar patrones de comportamiento, incluso en situaciones atípicas como feriados, movilizaciones o catástrofes.

Las soluciones basadas en las técnicas de procesamiento de datos encuentran aplicación en diversas áreas de las políticas públicas:

- **Transparencia electoral:** Es posible detectar irregularidades en resultados electorales por medio de técnicas de análisis de datos. En el año 2019, CIPPEC (el Centro de Implementación de Políticas Públicas para la Equidad y el Crecimiento), condujo un estudio sistemático⁶ sobre el proceso electoral en el que se definieron las autoridades a nivel provincial y municipal en las ciudades de Rosario y Santa Fe, cotejando los resultados de las PASO (elecciones Primarias Abiertas Simultáneas y Obligatorias) y las elecciones generales .
- **Planeamiento urbano:** Tanto por las herramientas de georreferenciación, sensores de polución o imágenes satelitales, es posible analizar la concentración poblacional, los flujos de movimiento, embotellamientos o áreas estratégicas dentro de una ciudad. La analítica de datos puede contribuir mediante la identificación de espacios verdes, superficies construidas, áreas protegidas, identificación de construcciones irregulares, edificaciones inapropiadas o terrenos baldíos, detectando desequilibrios y posibles ajustes.
- **Salud:** En menor escala que lo señalado al principio de este trabajo con respecto a los brotes epidemiológicos, los datos masivos también pueden contribuir a una mejora

⁶ Page, M., & Antenucci, P. (2019). *Elecciones en Santa Fe: un estudio sistémico sobre la confiabilidad del proceso electoral*. Buenos Aires: CIPPEC. Recuperado el 26 de octubre de 2021 de: <https://www.cippec.org/wp-content/uploads/2019/12/186-DT-IP-Confiabilidad-de-las-elecciones-en-Santa-Fe-Page-y-Antenucci....pdf>

en los servicios de gestión de salud pública reduciendo los tiempos de espera en los efectores públicos de emergencia, midiendo y evaluando las causas y factores que inciden en su incremento y asignando de manera inteligente los recursos (Coullaut & Tascón, 2016).

- Seguridad pública: El volumen de datos en poder de gobiernos nacionales y locales ha permitido reforzar mecanismos de seguridad para prevenir la proliferación de delitos. Las cámaras de seguridad y sistemas de identificación biométrica han generado amplias bases de datos que analizan características como: iris, rostro, ADN, etc. para contribuir con la seguridad pública y el esclarecimiento de delitos. Estos sistemas tienen por objeto unificar la información en manos del gobierno nacional. En algunos casos, los agentes cuentan con el acceso a esas bases de manera portátil, disminuyendo así el traslado innecesario de personas a las dependencias policiales (Prieto, 2020).
- Movilidad: Si bien las principales fuentes de datos para conocer las pautas de movilidad en las ciudades continúan siendo las encuestas domiciliarias de movilidad, que proporcionan información detallada y específica, estas implican costos elevados y no se realizan con tanta frecuencia por cuestiones de logística. En contrapartida, la mayor parte de los datos masivos que registran la actividad humana tiene localización espacial y temporal (tracks), por lo tanto esta fuente de datos se presenta como un excelente complemento para los estudios de movilidad urbana y la planificación del transporte.

También es posible analizar la movilidad de la población a partir de fuentes de datos específicas, como las tarjetas inteligentes de transporte, que registran los viajes realizados por sus abonados en transporte público. El uso de estos datos para hacer ajustes o tomar nuevas definiciones, contribuyen a una gestión más eficiente del transporte, no sólo mejorando el tránsito, sino también reduciendo las emisiones de carbono, abonando a la calidad de vida a los ciudadanos de manera integral.

- Estadística social: En España, el propio censo 2021 desató el debate metodológico, cuando se dió a conocer que no se elaboraría a partir de una operación censal clásica, sino a partir de registros administrativos (datos del padrón, de Hacienda, de la Seguridad Social, etc.). Estos registros deben ser cruzados para obtener una

información semejante a la facilitada en los censos tradicionales. Sin embargo, buena parte de la información habitualmente suministrada en el censo no puede ser obtenida de los registros administrativos (Gutierrez Puebla, 2018, p.214)

- Otras áreas en las que es posible utilizar datos masivos pueden ser de lo más diversas, como gestión de residuos, uso eficiente de la energía a través iluminación inteligente, detección de deserción estudiantil, estadística social, etc.

¿Cómo se generan los datos?

Cuando hablamos de datos masivos, es relevante conocer cuáles son las *fuentes de datos* para comprender las potencialidades y límites que presentan a la hora de ser explotados. Como señala Javier Gutierrez Puebla (2018), aunque los datasets⁷ del big data pueden ser exhaustivos, son a la vez una representación y una muestra, conformada por la tecnología y plataforma utilizadas, la ontología de datos empleada y el entorno regulador, y está sujeta a sesgo de muestreo (p.203). Por lo tanto, el sesgo es inherente al big data, pero varía dependiendo de la fuente de datos utilizada.

Aquí repasaremos algunas de las principales fuentes de datos usadas en políticas públicas en nuestra región. Si bien existen más opciones, a los efectos de este trabajo y en relación a las posibilidades estructurales y económicas del contexto, detallaremos las que se presentan con más frecuencia de uso.

Registros GPS (tracks):

Los primeros tracks GPS registrados aparecen a mediados de 1980s, pero se hicieron populares en 1990 cuando comenzaron a integrarse en dispositivos comerciales accesibles para compañías. Sin embargo, la verdadera revolución de esta tecnología ha tenido lugar más recientemente con la aparición de los teléfonos inteligentes en 2007 y su rápida expansión en los años sucesivos. (Gutiérrez Puebla et al., 2020, p.63)

Un gran número de los vehículos que circulan por las ciudades están equipados con dispositivos GPS: autobuses, taxis, bicicletas públicas, etc. produciendo millones de

⁷ Un set o conjunto de datos es una colección de datos, normalmente tabulada, que contiene los valores para cada una de las variables que corresponden a cada miembro del conjunto.

datos de recorridos. Los fabricantes almacenan de esta forma enormes volúmenes de datos sobre los vehículos que producen o las flotas que gestionan, como número de kilómetros recorridos, uso de combustible, estado de las rutas recorridas, tiempos de viaje, etc. Muchos datos generados por estos vehículos, se encuentran en manos de las autoridades que gestionan algunos de estos servicios, como los de bicicletas compartidas en ciudades como Buenos Aires o Rosario. Empresas como Uber cuentan con este tipo de dispositivos para analizar las rutas de sus empleados, si bien estos datos pertenecientes a las empresas son de difícil acceso, Uber ha creado un servicio web que permite visualizar tiempos de viaje y velocidades entre zonas de transporte en varias ciudades del mundo, algunas de ellas latinoamericanas, y descargar, de forma gratuita, los datos de tiempos de viaje en ficheros .csv⁸. Un sesgo posible de esta fuente de datos, tiene que ver con la menor penetración de las tecnologías GPS de los autos en barrios o regiones de condición socioeconómica más baja, por lo que los datos sobre ciertas zonas pueden ser escasos o no haber.

Tarjetas bancarias:

Cruzando estos datos se puede obtener información como gasto general según procedencia, localización de los lugares de consumo, rubros y pautas temporales del gasto (Big Transaction Data). Es posible, por ejemplo, evaluar el impacto económico de eventos turísticos con alto nivel de desagregación espacial, comparando el gasto durante el evento con un periodo anterior normal. De este modo, en Madrid con datos del BBVA, se comprobó que la semana del World Pride 2017 generó una importante inyección económica para la ciudad, con un gasto total directo de 115 millones de euros, un 15% más respecto a un periodo normal⁹.

⁸ Los archivos .csv (por las siglas en inglés *comma-separated values*) son un tipo de documento en formato abierto sencillo para representar datos en forma de tabla, en las que las columnas se separan por comas y las filas por saltos de línea. Se trata de archivos de texto cuyo formato específico les permite guardar los datos en un formato de tabla estructurada.

⁹ Peiro, M. (2017). 2,3 millones de personas y 115 millones de gasto directo durante el WorldPride 2017. El Mundo. Recuperado el 27 de octubre de 2021, de: <https://www.elmundo.es/madrid/2017/07/07/595f752eca4741976d8b45c7.html>

Telefonía móvil:

Las muestras de telefonía móvil son enormes y su nivel de sesgo es bajo (aunque dependiendo de la penetración de la operadora en la ciudad a analizar). Si bien es una tecnología usada por la mayor parte de la población, los datos de esta fuente tendrán un nivel de sesgo bajo si se dispone de datasets provenientes de compañías que tienen una alta penetración en el país y si no hay razones para pensar que esa operadora esté orientada hacia un determinado segmento del mercado. El análisis es más preciso cuando mayor sea la resolución espacial y la granularidad temporal de los datos:

- Resolución espacial: Refiere a la exactitud de la geolocalización. Los datos de telefonía móvil tienen una precisión media-alta, con errores de centenares de metros, suficiente para analizar la movilidad en las ciudades.
- Granularidad temporal: Refiere a la frecuencia con que se registran los datos.

Los datos de telefonía móvil aportan información sobre los usuarios asociada a los contratos: edad, género, domicilio, etc. Esta información normalmente no está disponible por razones de privacidad, o no suele ser muy fiable debido a que muchas terminales no tienen contrato o los clientes no suelen actualizar los datos, pero para salvar esto suele recurrirse a operaciones de enriquecimiento de los datos, ya sea de forma individual o agregada. También aportan información sobre las características sociodemográficas de los usuarios y de los viajes, basándose en las zonas de residencia y los lugares a donde se desplazan (usos de suelo), así como los consumos que realizan según el tipo de plan que tienen.

Este tipo de datos presenta sus limitaciones y difícilmente pueda llegar a alcanzar la cantidad de variables que suministra una encuesta, pero permite un amplio alcance con cierta periodicidad en la recolección.

Tarjetas inteligentes de transporte (Transport Smart Card en inglés):

Las tarjetas proporcionan una información exhaustiva sobre los viajes en transporte público en las ciudades donde su uso es obligatorio. Dependiendo del tipo de tarjeta, en muchos casos también se dispone de las características sociodemográficas del usuario (registradas al adquirir, personalizar o renovar la tarjeta). A partir de esos datos se

pueden generar matrices origen-destino y explorar las dinámicas espacio-temporales y los flujos en transporte público.

Existen dos tipos de tarjetas:

1. Tarjetas sin contacto: Registran automáticamente la entrada y salida del viajero de una estación o de un vehículo de transporte público a través de ondas de alta frecuencia.
2. Tarjetas de contacto: Exigen que el usuario pase la tarjeta por un lector, cosa que normalmente sólo hace en la entrada. En el caso de las estaciones de metro o ferrocarril, la geolocalización va asociada a cada una de ellas, así como también en las paradas de autobús cuando el registro del viaje se realiza en máquinas instaladas en las propias paradas. Al registrar solamente datos sobre la entrada del viajero, para conocer la salida se deben desarrollar algoritmos que tengan en cuenta el comportamiento del usuario a lo largo del día y de la semana, considerando los otros lugares por los que accede al transporte público.

La explotación de estos datos en estudios de movilidad entraña cierta complejidad, no sólo por el gran volumen, sino también por las lagunas de información existentes en la generación de los datos.

Internet y redes sociales:

En las redes sociales se registran opiniones sobre el funcionamiento de los sistemas de los servicios públicos, desde el transporte hasta las campañas de vacunación contra el Covid-19. A partir del método de scraping textual, es posible desentrañar las valoraciones e impresiones que circulan en la sociedad. Sin embargo, las redes sociales, al ser más utilizadas por grupos específicos, tenderán a tener un nivel de sesgo mayor.

En muchos casos los datos están geolocalizados y tienen un carácter dinámico debido a su uso constante. Estudios dinámicos basados en nuevas fuentes de datos, como Twitter, aportan indicadores de multiculturalidad y mezcla social referidos al uso del espacio a lo largo del día (segregación dinámica).

En base a los tweets que aportan información en tiempo real sobre las condiciones del tráfico y accidentes de circulación, hay un prototipo¹⁰ de la Ciudad Autónoma de Buenos Aires (CABA), donde se localizan sobre un mapa los incidentes y los tweets que los reportan. Esta información sirve de ayuda a los conductores para elegir una ruta que evite las retenciones derivadas de esos incidentes.

Planificadores de rutas (Google maps, TomTom MyDrive o Waze):

Son aplicaciones sociales de tránsito automotor en tiempo real y navegación asistida por GPS que, a diferencia de los softwares de navegación asistida por GPS tradicionales, son alimentados por los usuarios y aprenden de las rutas recorridas para proveer información de enrutamiento y actualizaciones de tráfico en tiempo real. Almacenan las coordenadas de los eventos creados por los usuarios, la hora a la que son registrados y la duración de los mismos, el sentido de circulación del vehículo, el tipo de calle y el grado de confianza en cada evento y usuario. La información es accesible a través de sus APIs¹¹ y de otros procedimientos de descarga. En el caso de Waze, la plataforma fue adquirida por Google en 2013 e integrada dentro de Google Maps.

Cámaras de tráfico y video vigilancia:

Utilizando inteligencia artificial y algoritmos, las imágenes poseen la utilidad de datos y permiten brindar información inmediata. Los contextos de aplicación son de lo más variados, desde desastres naturales, cruzando los datos del tráfico o del clima para hacer una evacuación rápida de las rutas de una ciudad, hasta control del mercado agrícola, donde los drones y tractores pueden inspeccionar las granjas y suelos, diagnosticar la vegetación y la salud de los cultivos, determinar los posibles rendimientos, rastrear al ganado, entre otras aplicaciones.

¹⁰ En el documento del BID (Banco Interamericano de Desarrollo) donde fue relevado el caso, la información dice ser de acceso público, pero el link que señalan como de libre acceso para chequearlo conduce a una página de intranet: <http://intranet.isistan.unicen.edu.ar/>)

¹¹ Sigla en inglés por Application Programming Interface. Es una interfaz de programación de aplicaciones, esto es un conjunto de rutinas que provee acceso a un determinado software. En la web las APIs permiten la descarga de datos de forma relativamente fácil.

Institucionalidad construida en el uso de datos masivos

En este apartado se describen experiencias donde los datos masivos son utilizados como herramienta al servicio de políticas públicas. Sin embargo, es menester destacar que no existe gran cantidad de información pública disponible al respecto de iniciativas de este tipo. Los casos reseñados son aquellos que se encuentran sistematizados y publicados en sitios web oficiales o en notas técnicas de libre acceso, generadas por organismos internacionales como el Banco Interamericano de Desarrollo.

Big Data en el nivel Nacional

Observatorio Nacional de Big Data

El 1 de junio de 2017 en Argentina se creó este observatorio¹², a cargo de la Secretaría de Tecnologías de la Información y las Comunicaciones (ahora Subsecretaría), dependiente de la Jefatura de Gabinete de Ministros. El observatorio cuenta con un equipo interdisciplinario que trabaja de manera coordinada con las áreas competentes de los Ministerios de Justicia y Derechos Humanos, Modernización y de Ciencia, Tecnología e Innovación Productiva. Sus integrantes realizan las tareas ad honorem, por lo cual no han sido creadas nuevas estructuras administrativas para su funcionamiento.

Este espacio es pensado desde su origen como un ámbito de estudio, investigación y formulación de regulaciones y políticas públicas que fomenten el desarrollo y la aplicación del big data. Los objetivos que figuran en la resolución de su creación apuntan a:

- Analizar la evolución del big data y sus potencialidades en el país.
- Estudiar el marco regulatorio en lo relacionado al uso de datos personales.
- Analizar riesgos y formular normas de buenas prácticas.
- Promover y crear plataformas públicas seguras y sustentables de big data para innovación y nuevas aplicaciones dentro y fuera del Estado

¹² Resolución 11-E/2017. Secretaría de Tecnologías de la Información y las Comunicaciones. Ciudad de Buenos Aires, 01/06/2017.

- Establecer intercambios de experiencias y articular investigaciones potenciando los recursos locales.

El sitio web oficial de gobierno indica que actualmente el observatorio busca “promover el intercambio de investigaciones y propuestas vinculadas con la responsabilidad de los intermediarios de internet, el tratamiento de los datos personales en entornos digitales y los impactos socioeconómicos derivados en el uso de tecnologías de procesamiento automatizado de la información”¹³.

De todos modos, hasta el momento no se encuentra disponible en la web de gobierno más información sobre lo realizado desde el observatorio, el organismo no aparece en el organigrama de la Jefatura de Ministros (Ver Anexo), y si bien el artículo 7º de su decreto de creación establecía que debía tramitarse el registro del dominio de Internet www.bigdata.gob.ar, a la fecha de entrega de este trabajo el mismo figura inexistente.

Big data en el nivel provincial

Rosario, Provincia de Santa Fe: Movilidad de los ciclistas

El objetivo de este estudio era comprender los patrones de movilidad de los ciclistas en relación con la infraestructura vial existente (sean ciclovías o no) y la relación de estos con los accidentes de tránsito, e identificar posibles mejoras de la infraestructura vial.

Para esto, se instalaron dispositivos de localización GPS en 150 bicicletas del sistema público “Mi bici, tu bici”, con un 85% de cobertura, y en 40 bicicletas de voluntarios privados, como grupo de control para complementar la información ya que su patrón de movilidad no está geográficamente limitado por la ubicación de los terminales de bicicletas públicas; permitiendo evaluar otro tipo de recorridos.

Los datos fueron recolectados en un periodo de dos semanas (bicicletas privadas) y seis semanas (bicicletas públicas). Considerando sólo los datos de días hábiles. Posteriormente, se sistematizó la cantidad de viajes, los tiempos invertidos en los

¹³Recuperado de: <https://www.argentina.gob.ar/jefatura/innovacion-publica/ssetic/grupo-de-trabajo/observatorio-nacional-big-data>

mismos, las distancias o la velocidad promedio, y los ejes viales más utilizados. Esta información se resumió en: mes, día, horario o individuo y según rango etario y género.

Para el análisis de los accidentes, se recopilaron datos de distintas agencias públicas: el Observatorio Vial de la Agencia Provincial de Seguridad Vial de Santa Fe y el Ente de la Movilidad de Rosario. También se cruzaron datos respecto a la velocidad y el recorrido de los viajes públicos y privados. Finalmente, se realizaron entrevistas focales a ciclistas privados para validar la información cuantitativa obtenida.

Los datos se visualizaron con mapas que mostraban los principales corredores utilizados por los ciclistas, la velocidad del tránsito de bicicletas, los focos de accidentes y la gravedad de dichos siniestros. El resultado fue la identificación de vías, cruces o zonas específicas que, por la frecuencia y gravedad de los accidentes, requerían mayor atención. Para complementar estos resultados, se hicieron entrevistas a ciclistas para profundizar en estos focos y en problemas generales.

Córdoba, Provincia de Córdoba: Movilidad y transporte público

Con el objetivo de incentivar el uso del transporte público y la movilidad, el gobierno de la ciudad desarrolló un sistema de seguimiento de la flota del transporte público a través de la instalación de dispositivos GPS en todos los buses del sistema público. De este modo, se recopila información respecto a tiempos de viaje, recorridos e ingresos. Por medio de la 'tarjeta única' (tarjeta de transporte), además, se recolecta información del pago, como la hora, tarifa y localización de cada transacción realizada.

La particularidad de Córdoba, es que no presenta sólo uno, sino dos sistemas de monitoreo. El servicio de buses públicos funciona con operadores privados, y en 2007 uno de estos operadores comenzó a monitorear sus buses, lanzando la aplicación *cuandollega.net* que permitía conocer la llegada de los buses a las paradas en tiempo real. En 2012, los funcionarios de la ciudad decidieron ampliar este servicio a las demás líneas en las que prestaban servicio otros operadores privados, y crear el Sistema de Monitoreo de flota del Sistema Público (SMSP), para eso instalaron dispositivos GPS en todos los buses para que transmitan en forma directa la ubicación a una nueva plataforma de datos de transporte público de la ciudad. Como resultado, hoy en día los

buses operados por operarios privados cuentan con dos dispositivos: el de la compañía y el de la ciudad.

El proyecto pone en evidencia la complejidad de trabajar con empresas privadas de alto impacto, los autores denominan a este fenómeno como “sensor sprawl” (refiriéndose al despliegue excesivo de infraestructura de adquisición de datos, a veces redundante o conflictiva), y lo señalan como el resultado de la ambición de varios actores en emplear esfuerzos relacionados a generar datos con poca coordinación entre sí (Townsend & Zambrano-Barragán, 2019).

Que la ciudad posea su propio sistema de monitoreo tiene una importancia estratégica, el manejo confiable y estable de su propia infraestructura de datos le permite ejecutar proyectos de planificación de largo plazo.

Neuquén, Provincia de Neuquén: Movilidad en el Área Metropolitana

Este estudio aplicado en los municipios del departamento de Confluencia (provincia de Neuquén) y del departamento de General Roca (provincia de Río Negro), tenía por objetivos actualizar y ampliar la encuesta de movilidad domiciliaria de Neuquén-Cipolletti de 2010, estimar y caracterizar el flujo de viajes en el área metropolitana de Neuquén; determinar el dinamismo de los flujos de viaje y su evolución a lo largo del día (comparando días hábiles y fines de semana); entender las diferentes dimensiones de la movilidad, considerando movilidad interna (dentro de cada localidad) y externa (entre las diferentes localidades), así como las principales características socioeconómicas de las personas que viajan en la zona.

A partir de la herramienta Smart Steps de Telefónica, basada en tecnologías big data para telefonía móvil, se obtuvieron datos de un periodo acotado de tiempo (4 semanas de mayo 2016). La población de la muestra la constituyeron personas de más de 15 años. Los datos fueron agregados a nivel espacial para un conjunto de zonas situadas en el área de estudio y extrapolados para representar la población total de la zona.

Se estimó el número de viajes en un día típico y uno de fin de semana, desagregando por localidades y presentando los resultados en forma de matrices origen–destino. Esta información se desagregó según el motivo de los viajes (casa, trabajo, otros), franjas

horarias, género, edad y estrato económico, a partir de inferencias respecto del uso del teléfono móvil y una encuesta con mensajes SMS.

A partir de las matrices origen-destino se estimaron los principales centros productores y receptores de viajes, se calcularon el número de personas en cada zona identificando los períodos de mayor demanda y se establecieron los principales flujos de interacción entre las diferentes zonas analizadas.

El estudio ha permitido entender las diferentes dimensiones de la movilidad con el fin de mejorar la planificación urbana, del transporte y la calidad de vida de las personas. Los resultados provenientes de los datos móviles se complementaron con una encuesta realizada a diferentes usuarios a través del servicio Movistar+ sobre:

- 1) Las actividades que motivan los desplazamientos
- 2) El modo de transporte (a pie, bicicleta, moto, automóvil particular, taxi, remis, colectivo)
- 3) El nivel de educación

La identificación de los principales flujos de viaje entre las zonas estudiadas facilita la efectividad de las políticas de gestión del transporte público y privado, mejorando las conexiones entre las zonas de mayor demanda, considerando además su dinámica a lo largo del día.

Si una ciudad no dispone de medios económicos para hacer una encuesta domiciliaria de movilidad, la utilización de CDRs¹⁴ puede ser una buena alternativa. Si existen recursos económicos suficientes, y se dispone ya de una encuesta de movilidad, puede valerse de los datos de telefonía móvil para actualizar los datos.

¹⁴ Call Detail Records: El detalle de registro de llamadas (CDR por sus siglas en inglés), son los registros de las actividades (llamadas y mensajes) realizadas por los usuarios de los teléfonos móviles, también se incluyen las sesiones de datos bajo esta categoría

Bahía Blanca, Provincia de Buenos Aires: ¿Qué pasa Bahía Blanca?

Esta es una plataforma de monitoreo ambiental que amplía la red de Gobierno Abierto y brinda la apertura de datos en torno a las mediciones del Polo Petroquímico, establecimientos industriales y sistemas de tratamiento de efluentes cloacales urbanos cuyo destino final es la ría de Bahía Blanca.

La iniciativa surge a raíz de la tensión entre activistas medioambientales y la industria petroquímica de la zona. El gobierno instaló sensores que miden continuamente efluentes líquidos y la contaminación acústica de origen industrial en las vecindades de las plantas petroquímicas, y en algunas de ellas también se dispuso de cámaras que las monitorean en vivo. La información así producida y la información general de las diferentes industrias y empresas se encuentra publicada en plataforma web y aplicación móvil, permitiendo seguir en tiempo real la contaminación del aire y acústica generada.

Los datos son monitoreados por el Comité Técnico Ejecutivo del Municipio encargado del control ambiental en colaboración con la Universidad Nacional del Sur, la Universidad Tecnológica Nacional e institutos especializados como el Instituto Argentino de Oceanografía (IADO).

Área Metropolitana de Buenos Aires: Movilidad y exclusión social

En el año 2018, se llevó adelante un estudio en el asentamiento informal que rodea el barrio Costa Esperanza, en el partido de San Martín, en el Área Metropolitana de Buenos Aires. Este barrio de reciente desarrollo, sometido a un rápido proceso de crecimiento, se encuentra en condiciones de vulnerabilidad estructural: en torno a un conjunto de calles angostas no asfaltadas, carece de redes de alcantarillado, agua corriente, electricidad y gas natural, y sufre una completa ausencia de instituciones públicas: escuelas, puestos sanitarios, policiales, bomberos. Más de la mitad de su población percibe la Asignación Universal por Hijo o la asistencia del Estado para niños con madre o padre desocupados. Los niveles de desocupación y empleo informal se sitúan en torno al 30%. La población total era de 21.903 habitantes al momento de realizado el informe.

En un marco de movilidad sostenible, el sistema de transporte debe ser eficiente y fiable para garantizar que la actividad económica se desarrolle en las condiciones adecuadas

y que la población acceda a los servicios, mejorando su calidad de vida y evitando situaciones de exclusión social y desigualdad. Para ello, es necesario disponer de modelos para estimar la demanda de transporte en diferentes escenarios, maximizar la rentabilidad económica y social de las actuaciones y priorizar las inversiones.

Centrado en la población mayor de 18 años, el estudio obtuvo los datos sobre los viajes y sus características a partir del análisis de los registros de telefonía móvil del mes de agosto, completados con encuestas SMS. El nivel socioeconómico de los usuarios se infirió a partir de los datos de facturación. El grupo de edad y género se obtuvo a partir de la información que registra la operadora en el contrato utilizando un algoritmo de clasificación automática.

Los resultados fueron contrastados y completados con los de la tarjeta de transporte SUBE, para comprobar la validación de los resultados obtenidos y para añadir información complementaria, como las líneas de colectivos más utilizadas por la población del área de estudio o el grado de la utilización de la tarifa social, destinada a grupos desfavorecidos (jubilados, excombatientes, asignación universal por hijo o por embarazo, becas y pensiones no contributivas, entre otros.). El 41% de los viajes en colectivos se hicieron con tarifa social.

Este tipo de estudios con fuentes de datos combinadas da cuenta de las potencialidades existentes para el análisis de los patrones de movilidad en barrios informales y la producción de información relevante para que las autoridades públicas puedan tomar decisiones orientadas a mejorar la movilidad de la población.

Ciudad Autónoma de Buenos Aires: Movilidad de los usuarios del Ferrocarril y Subte

Este estudio tenía por objetivo conocer la demanda de pasajeros en días hábiles en diferentes sistemas de transporte. La demanda fue segmentada por perfil, diferenciando género, nivel socioeconómico y edad, además de por franjas horarias. Los resultados apuntaban a actualizar y complementar las estadísticas existentes sobre el flujo de viajeros en estas líneas.

Uno de los aspectos novedosos del estudio fue su perspectiva de género. Para ello se propuso un estudio comparativo de las pautas de movilidad entre hombres y mujeres, así como un análisis centrado en la identificación de posibles situaciones de acoso sufridas por las mujeres en el transporte público.

La base de datos utilizada correspondía a registros de eventos de red generados por teléfonos conectados a la red de Telefónica en Argentina, durante agosto de 2018. Estos datos se cruzaron con otras fuentes: los Censos, los datos geoespaciales de las zonas de interés o los datos de la tarjeta de transporte SUBE.

Los registros son anonimizados, sustituyendo el número del teléfono por un identificador único, y se agregan a nivel de zona. Una vez agregados, los datos fueron extrapolados hasta alcanzar la población total mayor de 15 años en la zona en cuestión sobre la base de los datos del Censo. Se estima su lugar de origen y destino, siendo el origen la localidad donde residen y el destino la zona donde hacen la primera parada mayor a 30 minutos.

Para ahondar en las cuestiones relacionadas con las posibles agresiones sufridas por las mujeres en el transporte público, el análisis se complementó a través de una encuesta por SMS a costo 0 para los usuarios. Las encuestas se lanzaron entre el 15 de noviembre de 2018 y el 22 de enero de 2019. Los SMS con la encuesta se enviaron a usuarios de Movistar de sexo femenino que, a través de los datos de las antenas, recibieron la clasificación de usuarios frecuentes del Ferrocarril Roca (ramal La Plata-Buenos Aires) y de la línea C de subterráneos. El desarrollo y diseño de la encuesta estuvo a cargo del Banco Interamericano de Desarrollo en conjunto con el Ministerio de Transporte.

Los datos de la tarjeta de transporte SUBE se utilizaron para validar los resultados obtenidos previamente a través de los datos de telefonía móvil. Para ello, se compararon los volúmenes de viajeros en las diferentes estaciones del Subte y del ferrocarril, flujos de viajes origen-destino, así como volumen total según día y horario. Globalmente ambas fuentes obtienen pautas de movilidad muy similares, con algunas pequeñas variaciones debidas sobre todo a diferencias metodológicas en la captación de los datos.

Intervención en el Ciclo de Políticas Públicas

A modo de síntesis, siguiendo el esquema de análisis propuesto por Subirats (1989), en el presente apartado intentaremos reconstruir a partir de las experiencias reseñadas las “fases” en que el uso de datos masivos se hace presente como herramienta en la construcción de las políticas públicas en nuestro país.

En primer lugar, el Observatorio Nacional de Big Data, como estipula en sus objetivos, apunta al estudio, investigación y formulación de regulaciones y políticas públicas que fomenten el desarrollo y la aplicación del big data. Funcionando sin estructura propia, solo con la participación de profesionales de distintos ministerios de manera ad hoc ante proyectos específicos. Las iniciativas Ad Hoc son mencionadas por Townsend y Zambrano-Barragán (2019) como esfuerzos liderados por individuos que no producen un impacto directo en la manera en que operan las organizaciones.

En el caso de movilidad de bicicletas en Rosario, los datos sirvieron para evaluar el funcionamiento del programa de bicicletas públicas y las bicisendas, identificar focos problemáticos y obtener los datos necesarios para formular mejoras o nuevas iniciativas en el área de movilidad y transporte público. El uso de los datos tuvo lugar por un período acotado, si bien el programa sigue recolectando datos por medio del sistema de GPS instalado en todas las bicicletas públicas.

Con el programa ¿Qué pasa Bahía Blanca? encontramos el uso de los datos como forma de implementar una política de control y monitoreo del cumplimiento de normas de medio ambiente por parte de las industrias. Esta iniciativa es novedosa dentro de las reseñadas, tanto por la fase del ciclo en que son utilizados los datos, como por ser la única utilizada desde medio ambiente, siendo que la mayoría de los casos se orientan a políticas de movilidad. Por último, posee la cualidad de encontrarse institucionalizada y no sólo funcionar a intervalos de tiempo o para una demanda puntual.

En el estudio de movilidad y exclusión del AMBA, los datos se usan a modo de insumo para poder conocer en profundidad las condiciones sociales y materiales de la ciudadanía y su relación con el transporte, identificando la particularidad de los problemas para tener un panorama claro a la hora de formular o pensar las mejoras

posibles. Si bien no se ha usado propiamente en la definición de un curso de acción, es uno de los destinos para los que son considerados.

En el caso del flujo de pasajeros en las líneas de Ferrocarril y Subte, sucede algo similar, lo que se intenta por medio de los datos es obtener un insumo para identificar patrones que arrojen hallazgos o evidencien problemas que puedan ser abordados para mejorar los servicios públicos. Este no funciona de forma institucionalizada e incluso no se vale exclusivamente de los datos masivos, sino que busca complementarse con otras fuentes que validen los resultados encontrados, como las encuestas vía SMS. Podríamos pensarlo como una prueba o evaluación de la recolección de datos en sí misma y sus potencialidades.

En la experiencia de Córdoba, si bien en un inicio el uso de los datos fue orientado a implementar un servicio más confortable, que incentive a los pasajeros a elegir ese medio de transporte, con la seguridad de conocer la frecuencia de circulación de los coches. La iniciativa evolucionó a una valoración más alta de los datos, en tanto insumos que el Estado también usa para monitorear el funcionamiento del servicio.

A la hora de hacer este análisis, por fuera de los objetivos formalmente establecidos por cada proyecto, no siempre se encuentra del todo claro cuál es el uso o con qué fines se recopilan y estudian los datos. En la mayoría, parece ser a modo de diagnóstico de una situación problemática o monitoreo de programas que ya se encuentran en funcionamiento, pero en ninguno de los reseñados queda tan clara su presencia en la fase de definición de cursos de acción, a excepción de los casos de Rosario y Neuquén.

En relación a las características de la modalidad de trabajo en cada experiencia, la mayoría se han dado como propuestas ad hoc, los datos han sido recabados por un plazo determinado de tiempo o el estudio de los mismos se ha realizado en forma puntual en base a alguna demanda concreta. Sólo en los casos de Bahía Blanca y Córdoba la recolección, procesamiento y uso de los datos está institucionalizada de manera permanente.

Matriz síntesis: Momento de intervención en el Ciclo de Políticas Públicas

En la siguiente tabla se presentan las experiencias reseñadas durante este capítulo, diferenciando las fases del ciclo de políticas públicas en donde el uso de datos se encuentra presente con distintos colores según su intensidad. El color verde indica las fases en los que esa propuesta incorpora el uso de datos, en amarillo figuran los casos en los que el uso interviene de modo parcial o no tan claro, y en rojo las fases en donde no se usan los datos masivos dentro de una política.

Además, se diferencian las áreas de intervención, siendo más frecuente la presencia de los datos en políticas de Movilidad. Por último, la modalidad de uso, siendo *Ad Hoc* los casos en los que los datos se utilizan durante un período de tiempo acotado en proyectos específicos; *Institucional* en los casos en que el uso de datos es una herramienta que se encuentra incorporada de forma permanente como un insumo para el Estado; *Mixta* en el caso en que la recolección de datos se da de manera permanente, pero no se realiza la debida explotación de los mismos más que en proyectos específicos.

Tabla 1:

Momento de intervención en el Ciclo de Políticas Públicas

Caso	Área	Momento de intervención en el Ciclo de Políticas Públicas					Modalidad
		Identificación del problema	Formulación	Definición del curso de acción	Implementación	Monitoreo y Evaluación	
ONBD	Monitoreo						Ad Hoc
Rosario	Movilidad						Mixta
Córdoba	Movilidad						Institucional
Neuquén	Movilidad						Ad Hoc
Bahía Blanca	Medio Ambiente						Institucional
AMBA	Movilidad						Ad Hoc
CABA	Movilidad						Ad HoC

Fuente: Elaboración propia

3. Desafíos y potencialidades

En su último libro, Oscar Oszlak (2020) trae una pregunta fundamental al respecto de la relación entre el Estado y la tecnología: ¿Qué puede esperarse hacia el futuro en la adopción de tecnologías capaces de mejorar el desempeño en la gestión pública? De acuerdo con el autor, y atendiendo al recorrido realizado durante la presente exploración, coincidimos en que esa pregunta no aparece con tanta frecuencia por fuera de Estados Unidos y Europa, donde se incentivan numerosos programas de investigación. En esta parte del globo, por el contrario, no abunda una reflexión sistemática basada en evidencia empírica por parte de la academia ni de los propios gobiernos, sobre el presente de la ciencia de datos y sus posibilidades en la gestión pública.

En un informe sobre el estado de la información pública en Argentina, publicado por CIPPEC en septiembre de 2020, Natalia Aquilino (Directora de Monitoreo y Evaluación) destacaba que “la producción de datos de calidad es todavía una deuda clave para expandir el derecho y mejorar la toma de decisiones. Los datos que produce el Estado son difíciles de leer, bajar y manipular. La recolección, producción y uso de datos estatales es mejorable”¹⁵.

En este capítulo intentaremos desarrollar un balance sobre las posibilidades y límites que evidencia el desarrollo del big data en el campo de las políticas públicas de Argentina, atendiendo a otros dos interrogantes propios: ¿Cuánto valen los datos sin las capacidades estatales para abordarlos? ¿Cuáles son las condiciones necesarias para que big data cumpla un rol efectivo en política pública?

Desafíos y límites: ¿qué nos falta?

Townsend y Zambrano-Barragán (2019) advierten que muchas organizaciones sobreestiman el poder del análisis de datos y lo utilizan para justificar decisiones que ya fueron tomadas en procesos políticos u organizacionales. Esto distorsiona los incentivos

¹⁵Aquilino, N. (2020). El estado de los datos: ¿cómo mejorar la información pública?. Recuperado 26 de octubre de 2021, de Cipepec.org website: <https://www.cipepec.org/textual/el-estado-de-los-datos-como-mejorar-la-informacion-publica-2/>

y fomenta la manipulación de la información recogida, así como también de la forma en que se codifica y se clasifica, mediante procesos que producen resultados políticos deficientes o frontalmente discriminatorios (p.9).

Como señala Sosa Escudero (2020), “el vendaval de datos de big data es una buena noticia sólo si contribuye a mejorar la señal sin aumentar el ruido, o si es capaz de proveer información útil allí donde antes había sólo silencio y oscuridad” (p.35). El autor enfatiza que las ventajas de big data no vienen de la masividad per se, sino de que los datos sean procesados a través de la tecnología y modelo adecuados, que les permitan contribuir más al orden que al caos y no sólo para responder a una tendencia imperante en el campo de la gestión.

A riesgo de emplearse como una moda, el auge del big data llega a las distintas áreas de gestión pública y privada con grandes promesas que dan lugar “tanto a analistas honestos como a aventureros, como en la época de la fiebre del oro” (Sosa Escudero, 2020, p.95). El desafío de la ciencia de datos se vincula más a la búsqueda al interior de la inmensidad y no con la inmensidad en sí misma; es decir, con la capacidad de conducir una búsqueda coherente que resuelva la tensión existente entre conjeturas y datos.

Creemos posible identificar al menos tres dimensiones en las cuales el uso de datos masivos por parte del Estado aún encuentra algunos límites o sería perfectible en función de brindar un aporte sustancial al modo en que se piensan, diseñan y ejecutan las políticas públicas. En primer lugar, pondremos el foco en la calidad de los datos, entendiendo, como dijimos hasta aquí, que el dato en sí mismo carece de valor si no es procesado y trabajado de forma correcta. En segundo lugar, la dimensión jurídica y el resguardo de la privacidad de los datos que compete al Estado en un modo más sensible que en el caso de las empresas, ya que tiene una mayor responsabilidad sobre los mismos en tanto los ciudadanos no pueden negarse a brindarlos. Por último, las capacidades estatales (Oszlak, 2020, p. 32), entendidas como la capacidad estructural del Estado a nivel material, cultural e intelectual, para adaptarse y afrontar las transformaciones, cambios científicos y tecnológicos, intentado resolver los desafíos que presenta la coyuntura.

Calidad de los datos

El informe de CIPPEC mencionado anteriormente, señala que en Argentina la difusión de datos falla en apertura, acceso y publicidad. Los datos se presentan desarticulados, con poca comparabilidad y trazabilidad intertemporal (Aquilino, 2020). Este aspecto de la comparabilidad es mencionado también por Sosa Escudero (2017), quien explica que la naturaleza no estructurada del big data requiere un gran esfuerzo de sistematización para garantizar la comparabilidad y compatibilización en períodos largos.

Si bien el uso y difusión de los datos abiertos se encuentra instalado en las agendas de gobierno desde hace más de una década¹⁶, el informe destaca que los datasets disponibles en Datos Argentina¹⁷ (980 al momento de la publicación del informe) presentan información administrativa escasamente vinculada con metas de gobierno, lo cual refleja más la oferta de datos disponibles que la demanda por temas específicos. Este punto resulta fundamental, en tanto pone en evidencia que se continúa trabajando en base a recursos ya existentes, intentando amoldarlos a nuevas tendencias de gestión, en lugar de encaminar esfuerzos en pos de estrategias innovadoras que faciliten la generación de datos de calidad.

De acuerdo con CIPPEC es posible pensar una agenda de reformas que garantice datos de calidad por parte del Estado, pero para esto es necesario incluir los siguientes cambios:

1. Mejorar la tecnología disponible (el software y el hardware) para procesar y publicar datos masivos;
2. desarrollar instituciones capaces de gobernar algoritmos, personas y procesos;

¹⁶ Incluso es mencionado en el primer documento publicado por el Observatorio Nacional de Big Data, en el mes de julio 2021: “siempre que sea posible, los datos deben hacerse abiertos, a menos que los riesgos de abrir los datos sobrepasen los beneficios, o existan otras bases legítimas para no hacerlo” (p.7). Recuperado de:

https://www.argentina.gob.ar/sites/default/files/2021/04/bd_benchmark_v3_-_diseno.pdf

¹⁷ El sitio oficial de datos de gobierno: <https://www.datos.gob.ar/>, ofrece datos públicos en formatos abiertos para que puedan ser utilizados, modificados y compartidos. Al 18 de octubre de 2021 se encontraban a disposición 1061 set de datos de 33 organizaciones.

3. involucrar la participación activa de la ciudadanía para garantizar datos con sentido que promuevan la transparencia y respondan en tiempo real a la rendición de cuentas.

Es necesario, además, establecer alianzas estratégicas con las organizaciones cuya misión es apoyar el uso de datos, ya que muchos de estos son producidos por empresas que ven en ellos una oportunidad de negocio y por lo tanto no suelen compartirlos de forma libre. Frente a estas limitaciones, es interesante retomar a Walter Sosa Escudero (2020) cuando lanza la pregunta acerca de si es cierto que queremos *todos los datos*, dado el desafío de la estadística es que lo mejor no atente contra lo bueno, ya que no hay peor estadística que la que nadie usa (p.44). Los datos de big data no tienen estructura, ni obedecen a ningún plan sistemático, por lo tanto, es posible que unos pocos datos bien estructurados contengan más información útil que una enorme cantidad de datos indisciplinados, cuando no sesgados. El desafío consiste en darles estructura creíble a los datos para que puedan brindar información valiosa y justa, como la provista por los datos de encuestas clásicas. Los especialistas afirman que esto no significa que en el futuro los datos masivos vayan a sustituir a las encuestas, sino que recomiendan pensar ambas fuentes complementariamente.

Resguardo y privacidad de los datos

El creciente uso de los datos digitales genera preocupación a nivel internacional, especialmente en lo que refiere al respeto a los derechos ciudadanos, la protección contra la discriminación y las vulneraciones al trato igualitario, los cuales ponen de manifiesto la necesidad de una mayor protección principalmente en la esfera de los datos personales (Prieto, 2020). Siguiendo a Sosa Escudero (2019), la espontaneidad y aparente anarquía de big data puede poner a la sociedad a confrontar entre dos valores deseables como la transparencia y la privacidad. Que una no atente contra la otra es uno de los grandes desafíos en el uso y resguardo de los datos.

Nuestra región no posee una ley única sobre protección de datos personales como sucede en la Unión Europea que tiene el Reglamento General de Protección de Datos Personales desde 2018, pero los países coinciden en la necesidad de generar un marco normativo fuerte que acompañe el desarrollo de bienes y servicios de datos innovadores

y sostenibles. En el año 2000 Argentina aprobó la Ley 25.326, que incluye un Registro Nacional de Bases De Datos, un órgano de control y una adecuación aprobada, el 30 de junio de 2003, con arreglo a la directiva 95/46/CE del Parlamento Europeo y del Consejo sobre la adecuación de la protección de los datos personales en Argentina (Observatorio Nacional de Big Data, 2020).

La ley establece que no será necesario el consentimiento cuando los datos se obtengan de fuentes de acceso público irrestricto, esto no exonera a la organización que hace uso de los datos de cumplir con las obligaciones de protección de los datos personales ni deroga el habeas data¹⁸. Los datos sensibles podrán ser tratados con finalidades estadísticas o científicas cuando no puedan ser identificados sus titulares, o cuando se trate de bases de datos con fines periodísticos, pero no pueden ser utilizados para finalidades distintas o incompatibles con aquellas que motivaron su obtención.

La ley también establece que está prohibido el registro de información en bases de datos que “no reúnan condiciones técnicas de integridad y seguridad” (Ley N° 25.326, 2000, art. 9). En este sentido, es preciso perfeccionar los controles para reducir los riesgos de violaciones de bases de datos y su explotación ilícita. Al momento de escribirse esta tesis, tuvo lugar en el plazo de una semana un hackeo a una base de datos de la Obra Social de las Fuerzas Armadas (IOSFA)¹⁹ y a otra una base de datos del Registro Nacional de las Personas (ReNaPer)²⁰. La filtración fue divulgada en la red social Twitter, por un

¹⁸Argentina incorporó el uso del Hábeas Data en la Constitución Nacional de 1994 (Art. 43), una modalidad de amparo que permite a toda persona interesada acceder al conocimiento de sus datos personales que consten en registros o bancos de datos públicos o privados destinados y a exigir su supresión, rectificación, confidencialidad o actualización, en caso de falsedad o discriminación. Hasta el momento, cualquier ciudadano puede interponer un recurso de Hábeas Data, pero no podrá entablar una demanda debido a que el Congreso Nacional aún no reglamentó la norma.

¹⁹ El documento, filtrado del organismo que funciona bajo la órbita del Ministerio de Defensa, contenía información de 1.193.316 militares argentinos y empleados de diversas agencias de defensa nacional (Prefectura Naval Argentina, del Ejército Argentino, de la Fuerza Aérea Argentina, la Armada Argentina, Gendarmería Nacional y el Ministerio de Defensa). En él se encontraban nombres completos, estado civil, sexo, dirección postal, números de teléfono, correo electrónico y rango de las personas afectadas.

²⁰ En el organismo dependiente del Ministerio del Interior, la preocupación sobre la filtración abarca no solo a los datos personales de 45 millones de personas, sino también a un dato que se volvió fundamental para ciertos trámites con el Estado: el número de trámite de DNI. Se utiliza para ingresar a Trámites a Distancia, en ANSES (incluso para cambiar clave de seguridad social). Este número tan importante está siempre a la vista y es imposible de cambiar sin tramitar un nuevo DNI y eso lo convierte en un dato sensible.

usuario identificado como @aniballeaks (en inglés: la filtración de Aníbal, en referencia al flamante Ministro de Seguridad Aníbal Fernández), ofreciendo los documentos a la venta vía un link a la Deep web. Desde ambos organismos minimizaron los hechos, en el caso de IOSFA señalando que la información correspondía a un “padrón viejo” que se encontraba desactualizado, mientras del equipo de seguridad informática del ReNaPer aseguraron que no se trató de un hackeo, sino de una divulgación de 44 perfiles públicos que fueron consultados mediante un acceso indebido al sistema, pero de un organismo autorizado²¹.

El ReNaPer provee servicio a más de 150 instituciones públicas y privadas: entre los cuales se destacan Migraciones, AFIP, ANSES, PAMI, el Ministerio de Seguridad, el Ministerio de Transporte, de Salud, el Poder Judicial, Poderes Legislativos y billeteras digitales o bancos. Su base de datos se utiliza para validar la identidad de los ciudadanos argentinos en diferentes tipos de trámites, durante la pandemia este uso se fue haciendo extensivo a más entidades. De todos modos, el suceso cuenta con antecedentes que exceden a la actual gestión de gobierno²² y puso nuevamente en agenda pública la fragilidad y complejidad del manejo de los datos personales.

Algunos autores sugieren una serie de pasos que permiten resguardar la privacidad de los individuos en el diseño de soluciones basadas en técnicas de big data (Gutierrez Puebla, et al., 2019, p.60):

- Anonimización: Los datos deben ser en todo momento anónimos, es necesario que esa anonimización sea realizada por alguien que no esté involucrado en el análisis de datos. Para ello, las compañías tienen equipos independientes que realizan este proceso y proporcionan los datos ya anonimizados.

²¹ 2021, 13 de octubre, El Renaper negó que se haya hackeado su base de datos. Ámbito. Recuperado de: <https://www.ambito.com/politica/datos/el-renaper-nego-que-se-haya-hackeado-su-base-n5297803>

²² En agosto de 2020 fue hackeado el sistema de la Dirección General de Migraciones, de donde se extrajeron bases de datos de argentinos y extranjeros, el principal afectado fue el Sistema Integral de Captura Migratoria (SiCaM) que opera en los pasos internacionales, lo que ocasionó retrasos en el ingreso y egreso al territorio nacional. En noviembre del mismo año, un ataque similar recayó en la Agencia Nacional de Seguridad Vial. En ese caso los atacantes volvieron a amenazar con publicar los datos en la Deep web.

- Información mínima: Debe utilizarse la mínima información necesaria para cada estudio. Así, por ejemplo, en estudios realizados en base a CDRs simplificados, normalmente es suficiente con: el identificador anónimo del teléfono, la fecha y hora, y las antenas celulares involucradas en el evento. En algunos casos pueden utilizarse datos como el tiempo transcurrido de una llamada y raramente se incluyen información demográfica del suscriptor, como el código postal de facturación del teléfono ni ninguna información sobre la otra parte en la comunicación.
- Resultados agregados: La presentación de los resultados se debe realizar a partir de agregaciones de la información. Siguiendo con el ejemplo del análisis de datos telefónicos, esto implica no enfocarse en teléfonos individuales, a excepción de casos en los que se usen voluntarios para revisar sus registros.

Capacidades estatales

El uso de datos masivos permite abordar viejos problemas con nuevas herramientas y analizar cuestiones que antes quedaban ocultas para investigadores y gestores, fenómenos que conocíamos, pero no podíamos medir. Sin embargo, existen limitaciones institucionales que dificultan la real incorporación de este recurso. Al hablar de capacidades estatales, nos resulta pertinente recuperar la definición de Julián Bertranou (2015, p. 39): "la capacidad estatal es la aptitud de los entes estatales para alcanzar los fines que le han sido asignados interna o externamente". En este caso, los fines obedecen a la necesidad de adaptarse y afrontar las transformaciones, cambios científicos y tecnológicos, intentado resolver los desafíos que presenta la coyuntura, como menciona Oscar Oszlak (2020, p. 32).

Dejar hablar a los datos por sí solos, puede llevar a correlaciones espurias, sin ningún valor explicativo, el análisis de correlación debe estar guiado por una teoría previa que sirva para justificar la lógica de la relación entre variables. Por lo tanto, es necesario contar con equipos interdisciplinarios que combinen conocimientos técnicos y temáticos.

No existen muchos equipos que combinen la formación necesaria para trabajar como analistas de datos masivos y planificadores. Si la voluntad de adoptar esta herramienta

es genuina por parte de las administraciones, será necesario incorporar expertos en análisis de big data, o capacitar a cada perfil (el planificador y el analista de datos) para que pueda acercarse al campo del otro, alcanzando una colaboración fructífera. El planificador necesita conocimientos básicos sobre las fuentes, métodos de analítica y soluciones basadas en big data, sus ventajas y limitaciones, para poder ir más allá del dato y analizar patrones y procesos en su contexto social, económico y político. Este proceso debe enmarcarse en una transformación de la cultura organizacional que se oriente al uso de datos, con una mirada analítica crítica y un plan claro respecto a qué preguntas son urgentes de responder.

Por otro lado, es necesario ampliar las capacidades del Estado en cuanto a las tecnologías disponibles para trabajar los datos masivos o, en su caso, realizar operaciones de filtrado en las APIs para descargar datasets que ya sean procesables con tecnologías convencionales. Esto dependerá de las posibilidades de cada administración de destinar recursos a inversiones en equipamiento, situación que en Argentina no se presenta de forma uniforme debido a las disparidades estructurales entre las provincias, incluso en lo que respecta a cuestiones básicas como el acceso a internet a lo largo del territorio nacional. Así también, es necesario desarrollar la interoperabilidad entre sistemas de las distintas agencias o áreas para generar acciones a partir de los datos.

Por último, un aspecto tal vez menos técnico pero sustancial para el crecimiento de una gobernabilidad basada en datos, tiene que ver con las capacidades de comunicabilidad. Comunicar de manera clara y sistemática qué se hace con los datos y cómo pueden éstos mejorar la eficiencia de los servicios públicos, transparentando procesos, es parte de construir credibilidad y fomentar la participación activa de la ciudadanía. El big data se basa en explotar complejas estrategias y sofisticados métodos percibidos como "caja negra" fuera del mundo científico. El desafío es desarrollar una comunidad científica y equipos de planificadores interdisciplinarios que ayuden a dar solvencia a las técnicas y los resultados que determinan ciertas decisiones y cursos de acción. Incluso al interior de los equipos decisores, una mejor comunicación de los hallazgos podría generar mayor confiabilidad en las definiciones a largo plazo.

Oportunidades, potencialidades y condiciones necesarias

Los datos masivos prometen abrir las puertas a nuevas formas de conocer, pensar y abordar los problemas y realidades que desde siempre ocupan a los investigadores y planificadores que trabajan al servicio de lo público. Se espera, entonces, que una vez vencidos algunos de los desafíos antes mencionados, la revolución tecnológica imponga modalidades de trabajo mucho más colaborativas, transversales y experimentales, que deberán sobreponerse a una cultura burocrática regida por reglas y controles (Oszlak, 2020, p.111).

Al respecto, Townsend y Zambrano-Barragán (2019) proponen algunas recomendaciones basadas en su modelo de madurez de big data urbana, que sirven de referencia para que lo que ellos denominan “una ciudad promedio de la región” logre avanzar hacia el desarrollo de big data en las cinco competencias funcionales, siendo estas:

1. Datos abiertos: datos de calidad y libre acceso.
2. Ecosistema de datos: comunidades de usuarios, mecanismos de intercambio de información y una cultura de uso.
3. Analítica: capacidades técnicas para procesar, transformar, sintetizar y visualizar los datos, obteniendo hallazgos y nuevas perspectivas.
4. Toma de decisiones basadas en datos: prácticas institucionales y una cultura de pensamiento progresivo centrado en el uso sistemático de los datos para mejorar políticas, acciones y resultados.
5. Servicios urbanos digitales: uso de datos para guiar una innovación abierta distribuida entre actores (Estado, ciudadanía, privados, etc.).

Los autores desarrollan un modelo conceptual del ciclo de vida de las innovaciones de big data urbana, demostrando cómo se organizan y se institucionalizan los catalizadores e insumos de la innovación (Townsend & Zambrano-Barragán, 2019, p.37). Este ciclo consiste en seis pasos, cada uno de los cuales es facilitado por varias categorías funcionales del modelo de madurez:

1. Evaluar las capacidades y las prioridades sobre los medios (datos abiertos, privacidad y ética del uso de datos) y los fines de la innovación de datos (sostenibilidad, equidad, etc).
2. Identificación del problema y del suministro de datos externos, expandiendo la participación tanto como sea posible para nivelar la discusión y evitar una priorización políticamente contenciosa de inconvenientes.
3. Garantizar el acceso interno, dado que las mayores barreras para el intercambio de datos a menudo existen dentro del gobierno, entre niveles y agencias.
4. Generación de análisis y perspectivas, esto significa identificar y adquirir conocimientos, infraestructuras y dominio de la ciencia de información en todas las agencias gubernamentales y complementar esto a través de asociaciones externas con universidades, empresas privadas y organizaciones civiles, traduciendo esto en perspectivas comunicables.
5. Toma de decisiones y acción que vengan acompañadas de mecanismos de monitoreo para medir rendimientos y ajustar la implementación de políticas.
6. La institucionalización entendida como la priorización y evaluación de proyectos big data que permitan la planificación a largo plazo y sostenibilidad.

En pocas palabras, el modelo de madurez de big data urbana que proponen los autores es una herramienta útil para dividir las iniciativas en partes más pequeñas, volviendo más fácil de identificar su importancia o efectividad. Esta característica resultará fundamental para el análisis que se presenta en el próximo capítulo.

4. Hallazgos e hipótesis

A modo de cierre, recapitularemos algunas cuestiones que fueron apareciendo durante este recorrido exploratorio y que considero pueden servir como una base general para iniciar futuras investigaciones sobre la temática del big data y su relación con el Estado y la gestión pública.

Los datos masivos son una herramienta con enorme potencial para el diseño, implementación, monitoreo y evaluación de políticas públicas. Sin embargo, existe cierta sobredimensión sobre cuán desarrollado está el tema en nuestro país. Las experiencias reseñadas dan cuenta de que o bien los usos del big data aún son muy incipientes, o no existe sistematización y comunicación suficiente de las experiencias verdaderamente interesantes que se llevan adelante por parte del Estado y los gobiernos. Solo dos de ellas se encuentran incorporadas institucionalmente, mientras las demás funcionan como propuestas ad hoc según demandas puntuales por períodos de tiempo acotados, en algunos casos a modo de prueba, necesitando apoyo de otras fuentes de datos para corroborar la información. Sólo en el caso de Neuquén el uso de datos masivos se da de manera clara en las instancias de definición de cursos de acción.

La mayoría de las iniciativas refieren a políticas de movilidad, esto tiene que ver con que el 80% de los datos masivos son datos espaciales, debido a que la tecnología GPS es una de las más accesibles y de uso extensivo en la mayoría de los dispositivos inteligentes.

Ante esto, es importante también recordar que los datos son construcciones humanas, cómo se recoge y categoriza la información es fundamental para su correcta interpretación. Los datos solos no revisten un valor en sí mismos, es necesario entender cómo procesarlos y trabajarlos. Incorporar big data al proceso de políticas públicas implica virar hacia una mentalidad que entienda cómo explotar datos para generar mejores resultados, esto incluye conocer las fuentes de datos y sus posibilidades en el contexto y con los recursos disponibles en nuestro país. Qué nos dicen, y qué no nos dicen, los datos masivos va a estar íntimamente ligado a las fuentes de recolección y las preguntas que se plantean para ese proceso.

Muchos datos no garantizan tener mucha información, tener todos los datos no necesariamente implica ganar en conocimiento, los datos pueden venir con ruido o incluso pueden ser manipulados para respaldar decisiones previamente definidas. Pero esto no debe ser un motivo para desestimarlos, sino para trabajar en pos de una mayor calidad en su incorporación.

Actualmente, la calidad de los datos es mejorable, su difusión falla en apertura, acceso y publicidad. Como señalan los autores, los datos se presentan desarticulados, con poca comparabilidad y trazabilidad intertemporal. El resguardo y seguridad de los datos se ve falante, sufriendo hackeos de bases de datos de manera reiterada, lo cual genera desconfianza en la ciudadanía. Mientras las capacidades estatales para la incorporación total de la herramienta se encuentran en una etapa de desarrollo embrionario, no logrando cumplir aún la mayor parte de las condiciones establecidas en el modelo de madurez de big data urbana.

El hecho de que no se haya creado aún el dominio de bigdata.gob.ar, puede estar hablándonos de que la creación de ciertas estructuras responde más a una presión por subirse a una tendencia que a una voluntad (o una posibilidad) real de encaminarse hacia una gestión de lo público basada en datos.

Los datos sirven y deben incorporarse de forma seria como herramienta de elaboración de políticas públicas, pero siempre estarán sujetos a las posibilidades del contexto y las capacidades estatales; como señalaba Aguilar Villanueva, a la confluencia entre problemas, soluciones, actores involucrados y las oportunidades de elección. Es necesario destinar presupuesto, recursos, formación de equipos y comunicar la información involucrando a la ciudadanía, para que la gestión basada en datos sea más que una expresión de deseo o un nuevo ropaje digital para las prácticas analógicas de siempre.

Referencias bibliográficas

- ❖ ¿Qué sucede en internet cada minuto? De 900.000 accesos a Facebook a 3,5 millones de búsquedas de Google. (2021). *El Economista*. Recuperado el 26 de octubre de 2021, de: <https://www.eleconomista.es/tecnologia/noticias/8236038/03/17/Que-sucede-en-internet-cada-minuto-De-900000-accesos-a-Facebook-a-35-millones-de-busquedas-de-Google.html>
- ❖ Aguilar Villanueva, L. (1993). *Problemas públicos y agenda de gobierno*. México. Miguel Ángel Porrúa Grupo Editorial.
- ❖ Aquilino, N. (2020). El estado de los datos: ¿cómo mejorar la información pública? Recuperado 26 de octubre de 2021, de: <https://www.cippecc.org/textual/el-estado-de-los-datos-como-mejorar-la-informacion-publica-2/>
- ❖ Ardévol, E. (2016). Big Data y descripción densa. *Virtualis. Revista De Cultura Digital*, 7(14), Recuperado de: <https://www.revistavirtualis.mx/index.php/virtualis/article/view/186>.
- ❖ Bertranou, J. (2015). Capacidad estatal: Revisión del concepto y algunos ejes de análisis y debate. *Revista Estado Y Políticas Públicas*, Nº 4, 37–59. Recuperado el 27 de octubre de 2021, de: https://revistaeypp.flacso.org.ar/files/revistas/1433485077_dossier-2.pdf
- ❖ Coullaut, A., & Tascón, M. (2016). *Big data y el Internet de las cosas*. Madrid: Libros de la Catarata. Madrid: Libros de la Catarata.
- ❖ Davidovsky, S. (2020). Migraciones: cómo fue el ataque del ransomware Netwalker y qué tipo de datos revela. *La Nación*. Recuperado el 26 de octubre, de: <https://www.lanacion.com.ar/tecnologia/migraciones-como-fue-ataque-del-ransomware-netwalker-nid2446451/>
- ❖ Davidovsky, S. (2021). Preocupación por la venta online de los datos de los argentinos que tiene el Registro Nacional de las Personas. *La Nación*. Recuperado el 26 de octubre de 2021, de: <https://www.lanacion.com.ar/tecnologia/preocupacion-por->

la-venta-de-los-datos-de-millones-de-argentinos-que-tiene-el-registro-nacional-de-nid13102021/

❖ Davidovsky, S. (2021). Publican información privada de casi 1,2 millones de militares y empleados de agencias de seguridad argentinas. *La Nación*. Recuperado el 26 de octubre de 2021, de: <https://www.lanacion.com.ar/tecnologia/publican-informacion-privada-de-12-millones-de-militares-y-empleados-de-agencias-de-seguridad-nid29092021/>

❖ Davidovsky, S. (2020). Un ataque ransomware amenaza con publicar información privada de la Agencia Nacional de Seguridad Vial. *La Nación*. Recuperado el 26 de octubre de 2021, de: <https://www.lanacion.com.ar/tecnologia/un-ataque-ransomware-amenaza-publicar-informacion-privada-nid2521829/>

❖ Gutiérrez Puebla, J. (2018). Big Data y nuevas geografías: la huella digital de las actividades humanas. *Documents d'Anàlisi Geogràfica 2018*, vol. 64/2 , 195–217. Recuperado de: <https://doi.org/https://doi.org/10.5565/rev/dag.526>

❖ Gutiérrez Puebla, J., Benitez, C., García Palomares, J. C., Romanillos Arroyo, G., Rubinstein, E., Leño, J. M., Rivero, K. M., Scholl, L., Moya Gómez, B., & Condeco Melhorado, A. (2020). Cómo aplicar big data en la planificación del transporte: el uso de datos de GPS en el análisis de la movilidad urbana. *Banco Interamericano de Desarrollo*. Recuperado de: <http://dx.doi.org/10.18235/0002487>

❖ Gutiérrez Puebla, J., Benitez, C., Leño, J. M., García Palomares, J. C., Condeco Melhorado, A., Mojica, C., Scholl, L., Adler, V., Vera, F., Moya Gómez, B., & Romanillos Arroyo, G. (2019). Cómo aplicar big data en la planificación del transporte urbano: el uso de datos de telefonía móvil en el análisis de la movilidad. *Banco Interamericano de Desarrollo*. Recuperado de: <https://publications.iadb.org/es/como-aplicar-big-data-en-la-planificacion-del-transporte-urbano-el-uso-de-datos-de-telefonía-movil>

❖ Hirschon Weiss, C. (1999). XIV. La investigación de políticas: ¿Datos, ideas o argumentos? En *Ciencias Sociales y Estados modernos* (pp. 377–406). México: FCE. México: FCE.

- ❖ Lanzamiento de la Mesa Nacional de Internet de las Cosas. (2021). Recuperado el 26 de octubre de 2021, de: <https://www.argentina.gob.ar/noticias/lanzamiento-de-la-mesa-nacional-de-internet-de-las-cosas>

- ❖ Maestro Cano, I. C. (2019). Sobre las limitaciones de Big Data en ciencias sociales. *Sociología Y Tecnociencia: Revista Digital De sociología Del Sistema tecnocientífico*, 9(2), Recuperado de: <https://uvadoc.uva.es/handle/10324/37533>.

- ❖ Majone, G. (1997). *Evidencia, argumentación y persuasión en la formación de políticas*. México: FCE, CNCPAP.

- ❖ Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data. La revolución de los datos masivos*. Titivilus. Recuperado de: <https://b-ok.lat/book/5832104/a81a33>

- ❖ Observatorio Nacional de Big Data. (2021). *Big Data. Análisis comparado sobre los estándares regulatorios a nivel regional*. Buenos Aires: Secretaría de Innovación Pública, Jefatura de Gabinete de Ministros Argentina. Recuperado el 26 de octubre de 2021, de: <https://www.argentina.gob.ar/noticias/big-data-presentacion-de-analisis-sobre-estandares-regulatorios>

- ❖ Oszlak, O. (2020). *El Estado en la era exponencial (1.ª ed.)*. Buenos Aires: Instituto Nacional de Administración Pública. Recuperado el 26 de octubre de 2021, de: https://www.argentina.gob.ar/sites/default/files/el_estado_en_la_era_exponencial_-_oscar_oszlak_0.pdf

- ❖ Page, M., & Antenucci, P. (2019). *Elecciones en Santa Fe: un estudio sistémico sobre la confiabilidad del proceso electoral*. Buenos Aires: CIPPEC. Recuperado el 26 de octubre de 2021, de: <https://www.cippec.org/wp-content/uploads/2019/12/186-DT-IP-Confiabilidad-de-las-elecciones-en-Santa-Fe-Page-y-Antenucci....pdf>

- ❖ Prieto, R. (2020). *Big Data. Observatorio jurídico*. Montevideo: Agencia de Gobierno Electrónico y Sociedad de la Información y el Conocimiento. Recuperado de: <https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion->

conocimiento/sites/agencia-gobierno-electronico-sociedad-informacion-
conocimiento/files/documentos/publicaciones/big-data-oj.pdf

- ❖ Rodríguez, P., Palomino, N., & Mondaca, J. (2017). El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe. Banco Interamericano de Desarrollo. Recuperado de: <https://publications.iadb.org/publications/spanish/document/El-uso-de-datos-masivos-y-sus-t%C3%A9cnicas-anal%C3%ADticas-para-el-dise%C3%B1o-e-implementaci%C3%B3n-de-pol%C3%ADticas-p%C3%BAblicas-en-Latinoam%C3%A9rica-y-el-Caribe.pdf>
- ❖ Sosa Escudero, W. (2017). ¿Será posible reemplazar al Indec con big data y algoritmos? La Nación. Recuperado el 27 de octubre de 2021, de: <https://www.lanacion.com.ar/economia/sera-posible-reemplazar-al-indecon-big-data-y-algoritmos-nid2032197/>
- ❖ Sosa Escudero, W. (2019). Big data (3.ª ed.). Buenos Aires: Siglo veintiuno. Buenos Aires: Siglo veintiuno.
- ❖ Sosa Escudero, W. (2020). Borges, big data y yo (1.ª ed.). Buenos Aires: Siglo veintiuno. Buenos Aires: Siglo veintiuno.
- ❖ Townsend, A., & Zambrano-Barragán, P. (2019). Big Data urbana: Una guía estratégica para ciudades. Banco Interamericano de Desarrollo. Recuperado de: <http://dx.doi.org/10.18235/0001918>

Anexo

○ Decreto de creación de Observatorio Nacional de Big Data

MINISTERIO DE COMUNICACIONES SECRETARÍA DE TECNOLOGÍAS DE LA INFORMACIÓN
Y LAS COMUNICACIONES

Resolución 11-E/2017

Ciudad de Buenos Aires, 01/06/2017

VISTO el Expediente N° EX-2017-05648917-APN-SSR#MCO, el Decreto N° 13/2015, el Decreto N° 268/2015 y la Resolución N° 8/2016 de esta Secretaría, y

CONSIDERANDO:

Que el Decreto N° 13 de fecha 10 de diciembre de 2015 asignó las competencias correspondientes al MINISTERIO DE COMUNICACIONES y el Decreto N° 268 de fecha 29 de diciembre de 2015 aprobó la estructura organizativa de primer nivel operativo de esta Secretaría.

Que corresponde a esta área de gobierno asistir al señor Ministro de Comunicaciones en el diseño de políticas y regulaciones que permitan un mayor desarrollo e inclusión de las comunicaciones y de los servicios postales, elaborar estudios, propuestas de regulaciones en el ámbito de su competencia; diseñar y proponer la actualización de los marcos regulatorios de telecomunicaciones, tecnologías de la información y postal; y proponer actualizaciones normativas en el ámbito de su competencia.

Que por Resolución N° 8/2016 esta Secretaría creó el Grupo de Trabajo de Servicios de Internet que tiene por objeto analizar y proponer políticas públicas y regulaciones para la promoción y el desarrollo de servicios de Internet.

Que asistimos a una verdadera revolución en el tratamiento y la producción de cantidades masivas de datos, a partir del tráfico generado en Internet y el uso de dispositivos inteligentes.

Que el desafío consiste en el desarrollo de nuevas formas de procesamiento de datos que no pueden ser analizados utilizando herramientas o procesos tradicionales.

Que, a este conjunto de datos de gran volumen, alta velocidad y/o alta variedad de información, generados a través de la red y mediante el uso de dispositivos inteligentes, que demandan nuevas formas de procesamiento y que incidirán en la toma de decisiones y en la optimización de procesos, se le denomina Big Data.

Que el Big Data, por el enorme desarrollo, potencia y capacidad de interrelación de los diversos recursos involucrados, permite anticiparse a acontecimientos futuros y, en algunos casos, predecirlos con escaso margen de error

Que inmersos en la revolución digital, los datos se han convertido en un valor añadido, convirtiéndose en un activo clave, similar a las categorías clásicas de los recursos humanos y financieros, dando lugar a innovaciones en la tecnología, el desarrollo de nuevas herramientas y nuevas habilidades.

Que el buen uso de los datos puede traer oportunidades a sectores tradicionales de la actividad económica y social, como transporte, salud, educación, agroindustria, seguridad, etc., transformando además las industrias de servicios.

Que la gestión y el análisis de datos masivos resultan en una mejor gestión en términos de eficiencia y eficacia, principios fundamentales de la administración pública, generando, además de un mayor y mejor servicio al ciudadano, un ahorro sustancial en costes, derivado de un mejor aprovechamiento de los recursos.

Que asimismo resulta propicio generar un ámbito de estudio, investigación y formulación de regulaciones y políticas públicas que fomenten el desarrollo y la aplicación del Big Data.

Que, para ello, conforme los objetivos establecidos por la Resolución N° 8/2016 de esta Secretaría y los análisis desarrollados en cumplimiento de la misma, resulta oportuna la creación de un Observatorio Nacional de Big Data en su ámbito.

Que, dada la finalidad que se persigue, resulta conveniente y necesario establecer mecanismos de cooperación con entidades y organizaciones del ámbito público y

privado vinculadas a la temática, para la puesta en práctica del Observatorio Nacional de Big Data.

Que las implicancias de la ley N° 25.326 – Protección de los Datos Personales – devienen de vital importancia en la actividad de Big Data por lo que resulta de estricta necesidad la participación de su Autoridad de Aplicación en el cumplimiento de los objetivos aquí propuestos.

Que resulta necesaria la coordinación y articulación de los objetivos establecidos en la presente, no sólo con las áreas competentes del Ministerio de Justicia y Derechos Humanos, sino también con aquellas de los Ministerios de Modernización y de Ciencia, Tecnología e Innovación Productiva en razón de sus competencias.

Que la presente medida se dicta en ejercicio de las atribuciones conferidas por el Decreto N° 268/2015.

Por ello,

EL SECRETARIO DE TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES

RESUELVE:

ARTÍCULO 1°.- Créase, en el ámbito de la Secretaría de Tecnologías de la Información y las Comunicaciones, el Observatorio Nacional de Big Data.

ARTÍCULO 2°.- El Observatorio Nacional de Big Data tendrá por objeto:

- a. Analizar la evolución de la tecnología de Big Data en materia de posibilidades de innovación, uso público, beneficios tecnológicos, políticos, económicos y sociales en un marco sustentable.
- b. Estudiar el marco regulatorio del uso de datos personales y su relación con el desarrollo y promoción de Big Data.
- c. Promover y crear plataformas públicas seguras y sustentables de Big Data para innovación y nuevas aplicaciones dentro y fuera del Estado.

- d. Establecer intercambios con Estados provinciales y municipales, Pymes, emprendedores y microemprendedores en la implementación y uso de Big Data.
- e. Analizar riesgos y formular normas de buenas prácticas orientadas a un uso sustentable del Big Data.
- f. Articular investigaciones y desarrollos dentro del Estado Nacional e instituciones académicas, potenciando los recursos locales.
- g. Promover el uso de Big Data en seguridad pública y en la lucha contra el delito.
- h. Realizar informes, auditorías, compendios de buenas prácticas, propuestas de actualización normativa, plataformas tecnológicas de uso público, entre otras propuestas.

ARTÍCULO 3°.- El Observatorio Nacional de Big Data estará a cargo de la Secretaría de Tecnologías de la Información y las Comunicaciones, la que dictará las normas aclaratorias, complementarias y modificatorias que resulten necesarias para la instrumentación de la presente, establecerá el número de miembros que lo integran, así como su designación.

ARTÍCULO 4°.- El Observatorio definirá planes anuales y plurianuales para el cumplimiento de su objeto. Para ello, podrá dictar un Reglamento de Trabajo que se aplique tanto en la celebración de reuniones, como en la participación de sus integrantes y a los documentos, proyectos, propuestas e informes que la misma produzca.

ARTÍCULO 5°.- Esta Secretaría podrá proponer al Ministro de Comunicaciones la suscripción de convenios con otras áreas y dependencias estatales en los tres (3) niveles de gobierno, organismos internacionales, cámaras empresariales, asociaciones gremiales y sindicales, universidades nacionales públicas o privadas y organizaciones no gubernamentales, organismos y entidades técnicas y académicas, con el fin de promover la colaboración, promoción y fomento del contenido de la presente resolución.

ARTÍCULO 6°.- El Observatorio Nacional de Big Data trabajará de manera coordinada con las áreas competentes de los Ministerios de Justicia y Derechos Humanos,

Modernización y de Ciencia, Tecnología e Innovación Productiva, pudiendo conformar grupos de trabajo específicos.

ARTÍCULO 7°.- Tramítese, a través de NIC Argentina, el registro del dominio de Internet www.bigdata.gob.ar.

ARTÍCULO 8°.- La tarea de los integrantes del Observatorio Nacional de Big Data será ad honorem y no demandará recursos adicionales a las partidas del Presupuesto General de la Administración, ni la creación o modificación de estructuras administrativas existentes.

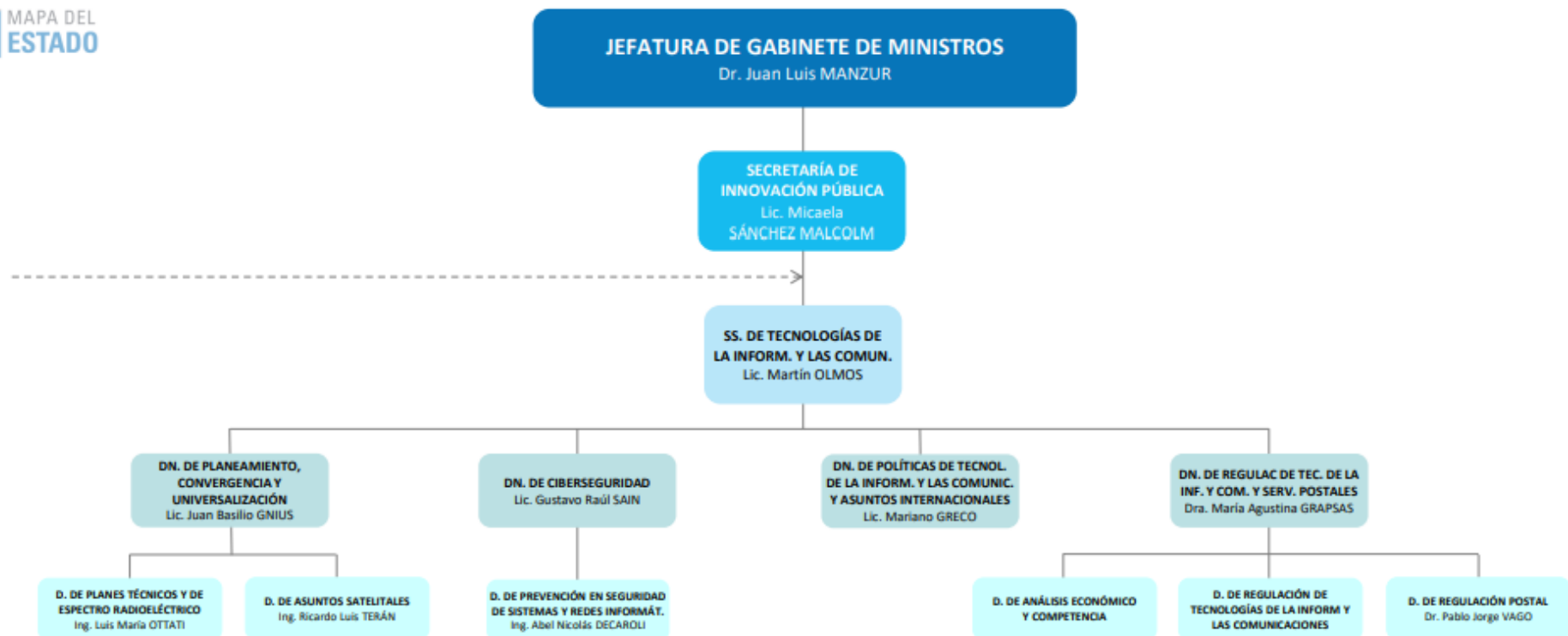
ARTÍCULO 9°.- Regístrese, comuníquese, publíquese, dése a la Dirección Nacional del Registro Oficial y archívese. — Héctor María Huici.

e. 08/06/2017 N° 38993/17 v. 08/06/2017

Fecha de publicación 08/06/2017

Fuente (consultada por última vez el 27/10/2021):
<https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-11-2017-275597/texto>

- Mapa del Estado (organigrama): Jefatura de Gabinete de Ministros de la Nación Argentina



Fuente (consultada por última vez el 27/10/2021): <https://mapadeleestado.jefatura.gob.ar/organigramas/001.pdf>

- Ley 25.326: Protección de Datos Personales

HONORABLE CONGRESO DE LA NACIÓN ARGENTINA

2000-11-02

PROTECCIÓN DE LOS DATOS PERSONALES

Ley 25.326

Disposiciones Generales. Principios generales relativos a la protección de datos. Derechos de los titulares de datos. Usuarios y responsables de archivos, registros y bancos de datos. Control. Sanciones. Acción de protección de los datos personales.

Sancionada: Octubre 4 de 2000.

Promulgada Parcialmente: Octubre 30 de 2000.

Ver Antecedentes Normativos

El Senado y Cámara de Diputados de la Nación Argentina reunidos en Congreso, etc. sancionan con fuerza de Ley:

Ley de Protección de los Datos Personales

Capítulo I

Disposiciones Generales

ARTÍCULO 1° — (Objeto).

La presente ley tiene por objeto la protección integral de los datos personales asentados en archivos, registros, bancos de datos, u otros medios técnicos de tratamiento de datos, sean éstos públicos, o privados destinados a dar informes, para garantizar el derecho al honor y a la intimidad de las personas, así como también el acceso a la información que sobre las mismas se registre, de conformidad a lo establecido en el artículo 43, párrafo tercero de la Constitución Nacional.

Las disposiciones de la presente ley también serán aplicables, en cuanto resulte pertinente, a los datos relativos a personas de existencia ideal.

En ningún caso se podrán afectar la base de datos ni las fuentes de información periodísticas.

ARTÍCULO 2° — (Definiciones).

A los fines de la presente ley se entiende por:

- Datos personales: Información de cualquier tipo referida a personas físicas o de existencia ideal determinadas o determinables.
- Datos sensibles: Datos personales que revelan origen racial y étnico, opiniones políticas, convicciones religiosas, filosóficas o morales, afiliación sindical e información referente a la salud o a la vida sexual.
- Archivo, registro, base o banco de datos: Indistintamente, designan al conjunto organizado de datos personales que sean objeto de tratamiento o procesamiento, electrónico o no, cualquiera que fuere la modalidad de su formación, almacenamiento, organización o acceso.
- Tratamiento de datos: Operaciones y procedimientos sistemáticos, electrónicos o no, que permitan la recolección, conservación, ordenación, almacenamiento, modificación, relacionamiento, evaluación, bloqueo, destrucción, y en general el procesamiento de datos personales, así como también su cesión a terceros a través de comunicaciones, consultas, interconexiones o transferencias.
- Responsable de archivo, registro, base o banco de datos: Persona física o de existencia ideal pública o privada, que es titular de un archivo, registro, base o banco de datos.
- Datos informatizados: Los datos personales sometidos al tratamiento o procesamiento electrónico o automatizado.

— Titular de los datos: Toda persona física o persona de existencia ideal con domicilio legal o delegaciones o sucursales en el país, cuyos datos sean objeto del tratamiento al que se refiere la presente ley.

— Usuario de datos: Toda persona, pública o privada que realice a su arbitrio el tratamiento de datos, ya sea en archivos, registros o bancos de datos propios o a través de conexión con los mismos.

— Disociación de datos: Todo tratamiento de datos personales de manera que la información obtenida no pueda asociarse a persona determinada o determinable.

Capítulo II

Principios generales relativos a la protección de datos

ARTÍCULO 3° — (Archivos de datos – Licitud).

La formación de archivos de datos será lícita cuando se encuentren debidamente inscriptos, observando en su operación los principios que establece la presente ley y las reglamentaciones que se dicten en su consecuencia.

Los archivos de datos no pueden tener finalidades contrarias a las leyes o a la moral pública.

ARTÍCULO 4° — (Calidad de los datos).

1. Los datos personales que se recojan a los efectos de su tratamiento deben ser ciertos, adecuados, pertinentes y no excesivos en relación al ámbito y finalidad para los que se hubieren obtenido.

2. La recolección de datos no puede hacerse por medios desleales, fraudulentos o en forma contraria a las disposiciones de la presente ley.

3. Los datos objeto de tratamiento no pueden ser utilizados para finalidades distintas o incompatibles con aquellas que motivaron su obtención.

4. Los datos deben ser exactos y actualizarse en el caso de que ello fuere necesario.

5. Los datos total o parcialmente inexactos, o que sean incompletos, deben ser suprimidos y sustituidos, o en su caso completados, por el responsable del archivo o base de datos cuando se tenga conocimiento de la inexactitud o carácter incompleto de la información de que se trate, sin perjuicio de los derechos del titular establecidos en el artículo 16 de la presente ley.

6. Los datos deben ser almacenados de modo que permitan el ejercicio del derecho de acceso de su titular.

7. Los datos deben ser destruidos cuando hayan dejado de ser necesarios o pertinentes a los fines para los cuales hubiesen sido recolectados.

ARTICULO 5° — (Consentimiento).

1. El tratamiento de datos personales es ilícito cuando el titular no hubiere prestado su consentimiento libre, expreso e informado, el que deberá constar por escrito, o por otro medio que permita se le equipare, de acuerdo a las circunstancias.

El referido consentimiento prestado con otras declaraciones, deberá figurar en forma expresa y destacada, previa notificación al requerido de datos, de la información descrita en el artículo 6° de la presente ley.

2. No será necesario el consentimiento cuando:

a) Los datos se obtengan de fuentes de acceso público irrestricto;

b) Se recaben para el ejercicio de funciones propias de los poderes del Estado o en virtud de una obligación legal;

c) Se trate de listados cuyos datos se limiten a nombre, documento nacional de identidad, identificación tributaria o previsional, ocupación, fecha de nacimiento y domicilio;

d) Deriven de una relación contractual, científica o profesional del titular de los datos, y resulten necesarios para su desarrollo o cumplimiento;

e) Se trate de las operaciones que realicen las entidades financieras y de las informaciones que reciban de sus clientes conforme las disposiciones del artículo 39 de la Ley 21.526.

ARTÍCULO 6° — (Información).

Cuando se recaben datos personales se deberá informar previamente a sus titulares en forma expresa y clara:

- a) La finalidad para la que serán tratados y quiénes pueden ser sus destinatarios o clase de destinatarios;
- b) La existencia del archivo, registro, banco de datos, electrónico o de cualquier otro tipo, de que se trate y la identidad y domicilio de su responsable;
- c) El carácter obligatorio o facultativo de las respuestas al cuestionario que se le proponga, en especial en cuanto a los datos referidos en el artículo siguiente;
- d) Las consecuencias de proporcionar los datos, de la negativa a hacerlo o de la inexactitud de los mismos;
- e) La posibilidad del interesado de ejercer los derechos de acceso, rectificación y supresión de los datos.

ARTÍCULO 7° — (Categoría de datos).

1. Ninguna persona puede ser obligada a proporcionar datos sensibles.
2. Los datos sensibles sólo pueden ser recolectados y objeto de tratamiento cuando medien razones de interés general autorizadas por ley. También podrán ser tratados con finalidades estadísticas o científicas cuando no puedan ser identificados sus titulares.
3. Queda prohibida la formación de archivos, bancos o registros que almacenen información que directa o indirectamente revele datos sensibles. Sin perjuicio de ello, la Iglesia Católica, las asociaciones religiosas y las organizaciones políticas y sindicales podrán llevar un registro de sus miembros.

4. Los datos relativos a antecedentes penales o contravencionales sólo pueden ser objeto de tratamiento por parte de las autoridades públicas competentes, en el marco de las leyes y reglamentaciones respectivas.

ARTÍCULO 8° — (Datos relativos a la salud).

Los establecimientos sanitarios públicos o privados y los profesionales vinculados a las ciencias de la salud pueden recolectar y tratar los datos personales relativos a la salud física o mental de los pacientes que acudan a los mismos o que estén o hubieren estado bajo tratamiento de aquéllos, respetando los principios del secreto profesional.

ARTÍCULO 9° — (Seguridad de los datos).

1. El responsable o usuario del archivo de datos debe adoptar las medidas técnicas y organizativas que resulten necesarias para garantizar la seguridad y confidencialidad de los datos personales, de modo de evitar su adulteración, pérdida, consulta o tratamiento no autorizado, y que permitan detectar desviaciones, intencionales o no, de información, ya sea que los riesgos provengan de la acción humana o del medio técnico utilizado.

2. Queda prohibido registrar datos personales en archivos, registros o bancos que no reúnan condiciones técnicas de integridad y seguridad.

ARTÍCULO 10. — (Deber de confidencialidad).

1. El responsable y las personas que intervengan en cualquier fase del tratamiento de datos personales están obligados al secreto profesional respecto de los mismos. Tal obligación subsistirá aun después de finalizada su relación con el titular del archivo de datos.

2. El obligado podrá ser relevado del deber de secreto por resolución judicial y cuando medien razones fundadas relativas a la seguridad pública, la defensa nacional o la salud pública.

ARTÍCULO 11. — (Cesión).

1. Los datos personales objeto de tratamiento sólo pueden ser cedidos para el cumplimiento de los fines directamente relacionados con el interés legítimo del cedente y del cesionario y con el previo consentimiento del titular de los datos, al que se le debe informar sobre la finalidad de la cesión e identificar al cesionario o los elementos que permitan hacerlo.

2. El consentimiento para la cesión es revocable.

3. El consentimiento no es exigido cuando:

a) Así lo disponga una ley;

b) En los supuestos previstos en el artículo 5° inciso 2;

c) Se realice entre dependencias de los órganos del Estado en forma directa, en la medida del cumplimiento de sus respectivas competencias;

d) Se trate de datos personales relativos a la salud, y sea necesario por razones de salud pública, de emergencia o para la realización de estudios epidemiológicos, en tanto se preserve la identidad de los titulares de los datos mediante mecanismos de disociación adecuados;

e) Se hubiera aplicado un procedimiento de disociación de la información, de modo que los titulares de los datos sean inidentificables.

4. El cesionario quedará sujeto a las mismas obligaciones legales y reglamentarias del cedente y éste responderá solidaria y conjuntamente por la observancia de las mismas ante el organismo de control y el titular de los datos de que se trate.

ARTÍCULO 12. — (Transferencia internacional).

1. Es prohibida la transferencia de datos personales de cualquier tipo con países u organismos internacionales o supranacionales, que no proporcionen niveles de protección adecuados.

2. La prohibición no regirá en los siguientes supuestos:

a) Colaboración judicial internacional;

- b) Intercambio de datos de carácter médico, cuando así lo exija el tratamiento del afectado, o una investigación epidemiológica, en tanto se realice en los términos del inciso e) del artículo anterior;
- c) Transferencias bancarias o bursátiles, en lo relativo a las transacciones respectivas y conforme la legislación que les resulte aplicable;
- d) Cuando la transferencia se hubiera acordado en el marco de tratados internacionales en los cuales la República Argentina sea parte;
- e) Cuando la transferencia tenga por objeto la cooperación internacional entre organismos de inteligencia para la lucha contra el crimen organizado, el terrorismo y el narcotráfico.

Capítulo III

Derechos de los titulares de datos

ARTÍCULO 13. — (Derecho de Información).

Toda persona puede solicitar información al organismo de control relativa a la existencia de archivos, registros, bases o bancos de datos personales, sus finalidades y la identidad de sus responsables.

El registro que se lleve al efecto será de consulta pública y gratuita.

ARTÍCULO 14. — (Derecho de acceso).

1. El titular de los datos, previa acreditación de su identidad, tiene derecho a solicitar y obtener información de sus datos personales incluidos en los bancos de datos públicos, o privados destinados a proveer informes.

2. El responsable o usuario debe proporcionar la información solicitada dentro de los diez días corridos de haber sido intimado fehacientemente.

Vencido el plazo sin que se satisfaga el pedido, o si evacuado el informe, éste se estimara insuficiente, quedará expedita la acción de protección de los datos personales o de hábeas data prevista en esta ley.

3. El derecho de acceso a que se refiere este artículo sólo puede ser ejercido en forma gratuita a intervalos no inferiores a seis meses, salvo que se acredite un interés legítimo al efecto.

4. El ejercicio del derecho al cual se refiere este artículo en el caso de datos de personas fallecidas le corresponderá a sus sucesores universales.

ARTÍCULO 15. — (Contenido de la información).

1. La información debe ser suministrada en forma clara, exenta de codificaciones y en su caso acompañada de una explicación, en lenguaje accesible al conocimiento medio de la población, de los términos que se utilicen.

2. La información debe ser amplia y versar sobre la totalidad del registro perteneciente al titular, aun cuando el requerimiento sólo comprenda un aspecto de los datos personales. En ningún caso el informe podrá revelar datos pertenecientes a terceros, aun cuando se vinculen con el interesado.

3. La información, a opción del titular, podrá suministrarse por escrito, por medios electrónicos, telefónicos, de imagen, u otro idóneo a tal fin.

ARTÍCULO 16. — (Derecho de rectificación, actualización o supresión).

1. Toda persona tiene derecho a que sean rectificadas, actualizados y, cuando corresponda, suprimidos o sometidos a confidencialidad los datos personales de los que sea titular, que estén incluidos en un banco de datos.

2. El responsable o usuario del banco de datos, debe proceder a la rectificación, supresión o actualización de los datos personales del afectado, realizando las operaciones necesarias a tal fin en el plazo máximo de cinco días hábiles de recibido el reclamo del titular de los datos o advertido el error o falsedad.

3. El incumplimiento de esta obligación dentro del término acordado en el inciso precedente, habilitará al interesado a promover sin más la acción de protección de los datos personales o de hábeas data prevista en la presente ley.

4. En el supuesto de cesión, o transferencia de datos, el responsable o usuario del banco de datos debe notificar la rectificación o supresión al cesionario dentro del quinto día hábil de efectuado el tratamiento del dato.

5. La supresión no procede cuando pudiese causar perjuicios a derechos o intereses legítimos de terceros, o cuando existiera una obligación legal de conservar los datos.

6. Durante el proceso de verificación y rectificación del error o falsedad de la información que se trate, el responsable o usuario del banco de datos deberá o bien bloquear el archivo, o consignar al proveer información relativa al mismo la circunstancia de que se encuentra sometida a revisión.

7. Los datos personales deben ser conservados durante los plazos previstos en las disposiciones aplicables o en su caso, en las contractuales entre el responsable o usuario del banco de datos y el titular de los datos.

ARTÍCULO 17. — (Excepciones).

1. Los responsables o usuarios de bancos de datos públicos pueden, mediante decisión fundada, denegar el acceso, rectificación o la supresión en función de la protección de la defensa de la Nación, del orden y la seguridad públicos, o de la protección de los derechos e intereses de terceros.

2. La información sobre datos personales también puede ser denegada por los responsables o usuarios de bancos de datos públicos, cuando de tal modo se pudieran obstaculizar actuaciones judiciales o administrativas en curso vinculadas a la investigación sobre el cumplimiento de obligaciones tributarias o previsionales, el desarrollo de funciones de control de la salud y del medio ambiente, la investigación de delitos penales y la verificación de infracciones administrativas. La resolución que así lo disponga debe ser fundada y notificada al afectado.

3. Sin perjuicio de lo establecido en los incisos anteriores, se deberá brindar acceso a los registros en cuestión en la oportunidad en que el afectado tenga que ejercer su derecho de defensa.

ARTÍCULO 18. — (Comisiones legislativas).

Las Comisiones de Defensa Nacional y la Comisión Bicameral de Fiscalización de los Organos y Actividades de Seguridad Interior e Inteligencia del Congreso de la Nación y la Comisión de Seguridad Interior de la Cámara de Diputados de la Nación, o las que las sustituyan, tendrán acceso a los archivos o bancos de datos referidos en el artículo 23 inciso 2 por razones fundadas y en aquellos aspectos que constituyan materia de competencia de tales Comisiones.

ARTÍCULO 19. — (Gratuidad).

La rectificación, actualización o supresión de datos personales inexactos o incompletos que obren en registros públicos o privados se efectuará sin cargo alguno para el interesado.

ARTÍCULO 20. — (Impugnación de valoraciones personales).

1. Las decisiones judiciales o los actos administrativos que impliquen apreciación o valoración de conductas humanas, no podrán tener como único fundamento el resultado del tratamiento informatizado de datos personales que suministren una definición del perfil o personalidad del interesado.

2. Los actos que resulten contrarios a la disposición precedente serán insanablemente nulos.

Capítulo IV

Usuarios y responsables de archivos, registros y bancos de datos

ARTÍCULO 21. — (Registro de archivos de datos. Inscripción).

1. Todo archivo, registro, base o banco de datos público, y privado destinado a proporcionar informes debe inscribirse en el Registro que al efecto habilite el organismo de control.

2. El registro de archivos de datos debe comprender como mínimo la siguiente información:

a) Nombre y domicilio del responsable;

- b) Características y finalidad del archivo;
 - c) Naturaleza de los datos personales contenidos en cada archivo;
 - d) Forma de recolección y actualización de datos;
 - e) Destino de los datos y personas físicas o de existencia ideal a las que pueden ser transmitidos;
 - f) Modo de interrelacionar la información registrada;
 - g) Medios utilizados para garantizar la seguridad de los datos, debiendo detallar la categoría de personas con acceso al tratamiento de la información;
 - h) Tiempo de conservación de los datos;
 - i) Forma y condiciones en que las personas pueden acceder a los datos referidos a ellas y los procedimientos a realizar para la rectificación o actualización de los datos.
- 3) Ningún usuario de datos podrá poseer datos personales de naturaleza distinta a los declarados en el registro.

El incumplimiento de estos requisitos dará lugar a las sanciones administrativas previstas en el capítulo VI de la presente ley.

ARTÍCULO 22. — (Archivos, registros o bancos de datos públicos).

1. Las normas sobre creación, modificación o supresión de archivos, registros o bancos de datos pertenecientes a organismos públicos deben hacerse por medio de disposición general publicada en el Boletín Oficial de la Nación o diario oficial.
2. Las disposiciones respectivas, deben indicar:
 - a) Características y finalidad del archivo;
 - b) Personas respecto de las cuales se pretenda obtener datos y el carácter facultativo u obligatorio de su suministro por parte de aquéllas;
 - c) Procedimiento de obtención y actualización de los datos;

d) Estructura básica del archivo, informatizado o no, y la descripción de la naturaleza de los datos personales que contendrán;

e) Las cesiones, transferencias o interconexiones previstas;

f) Órganos responsables del archivo, precisando dependencia jerárquica en su caso;

g) Las oficinas ante las que se pudiesen efectuar las reclamaciones en ejercicio de los derechos de acceso, rectificación o supresión.

3. En las disposiciones que se dicten para la supresión de los registros informatizados se establecerá el destino de los mismos o las medidas que se adopten para su destrucción.

ARTÍCULO 23. — (Supuestos especiales).

1. Quedarán sujetos al régimen de la presente ley, los datos personales que por haberse almacenado para fines administrativos, deban ser objeto de registro permanente en los bancos de datos de las fuerzas armadas, fuerzas de seguridad, organismos policiales o de inteligencia; y aquellos sobre antecedentes personales que proporcionen dichos bancos de datos a las autoridades administrativas o judiciales que los requieran en virtud de disposiciones legales.

2. El tratamiento de datos personales con fines de defensa nacional o seguridad pública por parte de las fuerzas armadas, fuerzas de seguridad, organismos policiales o inteligencia, sin consentimiento de los afectados, queda limitado a aquellos supuestos y categoría de datos que resulten necesarios para el estricto cumplimiento de las misiones legalmente asignadas a aquéllos para la defensa nacional, la seguridad pública o para la represión de los delitos. Los archivos, en tales casos, deberán ser específicos y establecidos al efecto, debiendo clasificarse por categorías, en función de su grado de fiabilidad.

3. Los datos personales registrados con fines policiales se cancelarán cuando no sean necesarios para las averiguaciones que motivaron su almacenamiento.

ARTÍCULO 24. — (Archivos, registros o bancos de datos privados).

Los particulares que formen archivos, registros o bancos de datos que no sean para un uso exclusivamente personal deberán registrarse conforme lo previsto en el artículo 21.

ARTÍCULO 25. — (Prestación de servicios informatizados de datos personales).

1. Cuando por cuenta de terceros se presten servicios de tratamiento de datos personales, éstos no podrán aplicarse o utilizarse con un fin distinto al que figure en el contrato de servicios, ni cederlos a otras personas, ni aun para su conservación.

2. Una vez cumplida la prestación contractual los datos personales tratados deberán ser destruidos, salvo que medie autorización expresa de aquel por cuenta de quien se prestan tales servicios cuando razonablemente se presume la posibilidad de ulteriores encargos, en cuyo caso se podrá almacenar con las debidas condiciones de seguridad por un período de hasta dos años.

ARTÍCULO 26. — (Prestación de servicios de información crediticia).

1. En la prestación de servicios de información crediticia sólo pueden tratarse datos personales de carácter patrimonial relativos a la solvencia económica y al crédito, obtenidos de fuentes accesibles al público o procedentes de informaciones facilitadas por el interesado o con su consentimiento.

2. Pueden tratarse igualmente datos personales relativos al cumplimiento o incumplimiento de obligaciones de contenido patrimonial, facilitados por el acreedor o por quien actúe por su cuenta o interés.

3. A solicitud del titular de los datos, el responsable o usuario del banco de datos, le comunicará las informaciones, evaluaciones y apreciaciones que sobre el mismo hayan sido comunicadas durante los últimos seis meses y el nombre y domicilio del cesionario en el supuesto de tratarse de datos obtenidos por cesión.

4. Sólo se podrán archivar, registrar o ceder los datos personales que sean significativos para evaluar la solvencia económico-financiera de los afectados durante los últimos cinco años. Dicho plazo se reducirá a dos años cuando el deudor cancele o de otro modo extinga la obligación, debiéndose hacer constar dicho hecho.

5. La prestación de servicios de información crediticia no requerirá el previo consentimiento del titular de los datos a los efectos de su cesión, ni la ulterior comunicación de ésta, cuando estén relacionados con el giro de las actividades comerciales o crediticias de los cesionarios.

ARTÍCULO 27. — (Archivos, registros o bancos de datos con fines de publicidad).

1. En la recopilación de domicilios, reparto de documentos, publicidad o venta directa y otras actividades análogas, se podrán tratar datos que sean aptos para establecer perfiles determinados con fines promocionales, comerciales o publicitarios; o permitan establecer hábitos de consumo, cuando éstos figuren en documentos accesibles al público o hayan sido facilitados por los propios titulares u obtenidos con su consentimiento.

2. En los supuestos contemplados en el presente artículo, el titular de los datos podrá ejercer el derecho de acceso sin cargo alguno.

3. El titular podrá en cualquier momento solicitar el retiro o bloqueo de su nombre de los bancos de datos a los que se refiere el presente artículo.

ARTÍCULO 28. — (Archivos, registros o bancos de datos relativos a encuestas).

1. Las normas de la presente ley no se aplicarán a las encuestas de opinión, mediciones y estadísticas relevadas conforme a Ley 17.622, trabajos de prospección de mercados, investigaciones científicas o médicas y actividades análogas, en la medida que los datos recogidos no puedan atribuirse a una persona determinada o determinable.

2. Si en el proceso de recolección de datos no resultara posible mantener el anonimato, se deberá utilizar una técnica de disociación, de modo que no permita identificar a persona alguna.

Capítulo V

Control

ARTÍCULO 29. — (Órgano de Control).

1. El órgano de control deberá realizar todas las acciones necesarias para el cumplimiento de los objetivos y demás disposiciones de la presente ley. A tales efectos tendrá las siguientes funciones y atribuciones:

a) Asistir y asesorar a las personas que lo requieran acerca de los alcances de la presente y de los medios legales de que disponen para la defensa de los derechos que ésta garantiza;

b) Dictar las normas y reglamentaciones que se deben observar en el desarrollo de las actividades comprendidas por esta ley;

c) Realizar un censo de archivos, registros o bancos de datos alcanzados por la ley y mantener el registro permanente de los mismos;

d) Controlar la observancia de las normas sobre integridad y seguridad de datos por parte de los archivos, registros o bancos de datos. A tal efecto podrá solicitar autorización judicial para acceder a locales, equipos, o programas de tratamiento de datos a fin de verificar infracciones al cumplimiento de la presente ley;

e) Solicitar información a las entidades públicas y privadas, las que deberán proporcionar los antecedentes, documentos, programas u otros elementos relativos al tratamiento de los datos personales que se le requieran. En estos casos, la autoridad deberá garantizar la seguridad y confidencialidad de la información y elementos suministrados;

f) Imponer las sanciones administrativas que en su caso correspondan por violación a las normas de la presente ley y de las reglamentaciones que se dicten en su consecuencia;

g) Constituirse en querellante en las acciones penales que se promovieran por violaciones a la presente ley;

h) Controlar el cumplimiento de los requisitos y garantías que deben reunir los archivos o bancos de datos privados destinados a suministrar informes, para obtener la correspondiente inscripción en el Registro creado por esta ley.

2. (Punto vetado por art. 1° del Decreto N° 995/2000 B.O. 2/11/2000)

3. (Punto vetado por art. 1° del Decreto N° 995/2000 B.O. 2/11/2000)

El Director tendrá dedicación exclusiva en su función, encontrándose alcanzado por las incompatibilidades fijadas por ley para los funcionarios públicos y podrá ser removido por el Poder Ejecutivo por mal desempeño de sus funciones.

ARTÍCULO 30. — (Códigos de conducta).

1. Las asociaciones o entidades representativas de responsables o usuarios de bancos de datos de titularidad privada podrán elaborar códigos de conducta de práctica profesional, que establezcan normas para el tratamiento de datos personales que tiendan a asegurar y mejorar las condiciones de operación de los sistemas de información en función de los principios establecidos en la presente ley.

2. Dichos códigos deberán ser inscriptos en el registro que al efecto lleve el organismo de control, quien podrá denegar la inscripción cuando considere que no se ajustan a las disposiciones legales y reglamentarias sobre la materia.

Capítulo VI

Sanciones

ARTÍCULO 31. — (Sanciones administrativas).

1. Sin perjuicio de las responsabilidades administrativas que correspondan en los casos de responsables o usuarios de bancos de datos públicos; de la responsabilidad por daños y perjuicios derivados de la inobservancia de la presente ley, y de las sanciones penales que correspondan, el organismo de control podrá aplicar las sanciones de apercibimiento, suspensión, multa de mil pesos (\$ 1.000.-) a cien mil pesos (\$ 100.000.-), clausura o cancelación del archivo, registro o banco de datos.

2. La reglamentación determinará las condiciones y procedimientos para la aplicación de las sanciones previstas, las que deberán graduarse en relación a la gravedad y extensión de la violación y de los perjuicios derivados de la infracción, garantizando el principio del debido proceso.

ARTÍCULO 32. — (Sanciones penales).

1. Incorporase como artículo 117 bis del Código Penal, el siguiente:

1°. Será reprimido con la pena de prisión de un mes a dos años el que insertara o hiciera insertar a sabiendas datos falsos en un archivo de datos personales.

2°. La pena será de seis meses a tres años, al que proporcionara a un tercero a sabiendas información falsa contenida en un archivo de datos personales.

3°. La escala penal se aumentará en la mitad del mínimo y del máximo, cuando del hecho se derive perjuicio a alguna persona.

4°. Cuando el autor o responsable del ilícito sea funcionario público en ejercicio de sus funciones, se le aplicará la accesoria de inhabilitación para el desempeño de cargos públicos por el doble del tiempo que el de la condena'.

2. Incorpórase como artículo 157 bis del Código Penal el siguiente:

Será reprimido con la pena de prisión de un mes a dos años el que:

1°. A sabiendas e ilegítimamente, o violando sistemas de confidencialidad y seguridad de datos, accediere, de cualquier forma, a un banco de datos personales;

2°. Revelare a otro información registrada en un banco de datos personales cuyo secreto estuviere obligado a preservar por disposición de una ley.

Cuando el autor sea funcionario público sufrirá, además, pena de inhabilitación especial de uno a cuatro años'.

Capítulo VII

Acción de protección de los datos personales

ARTÍCULO 33. — (Procedencia).

1. La acción de protección de los datos personales o de hábeas data procederá:

a) para tomar conocimiento de los datos personales almacenados en archivos, registros o bancos de datos públicos o privados destinados a proporcionar informes, y de la finalidad de aquéllos;

b) en los casos en que se presume la falsedad, inexactitud, desactualización de la información de que se trata, o el tratamiento de datos cuyo registro se encuentra prohibido en la presente ley, para exigir su rectificación, supresión, confidencialidad o actualización.

ARTÍCULO 34. — (Legitimación activa).

La acción de protección de los datos personales o de hábeas data podrá ser ejercida por el afectado, sus tutores o curadores y los sucesores de las personas físicas, sean en línea directa o colateral hasta el segundo grado, por sí o por intermedio de apoderado.

Cuando la acción sea ejercida por personas de existencia ideal, deberá ser interpuesta por sus representantes legales, o apoderados que éstas designen al efecto.

En el proceso podrá intervenir en forma coadyuvante el Defensor del Pueblo.

ARTÍCULO 35. — (Legitimación pasiva).

La acción procederá respecto de los responsables y usuarios de bancos de datos públicos, y de los privados destinados a proveer informes.

ARTÍCULO 36. — (Competencia).

Será competente para entender en esta acción el juez del domicilio del actor; el del domicilio del demandado; el del lugar en el que el hecho o acto se exteriorice o pudiera tener efecto, a elección del actor.

Procederá la competencia federal:

- a) cuando se interponga en contra de archivos de datos públicos de organismos nacionales, y
- b) cuando los archivos de datos se encuentren interconectados en redes interjurisdicciones, nacionales o internacionales.

ARTÍCULO 37. — (Procedimiento aplicable).

La acción de hábeas data tramitará según las disposiciones de la presente ley y por el procedimiento que corresponde a la acción de amparo común y supletoriamente por las

normas del Código Procesal Civil y Comercial de la Nación, en lo atinente al juicio sumarísimo.

ARTÍCULO 38. — (Requisitos de la demanda).

1. La demanda deberá interponerse por escrito, individualizando con la mayor precisión posible el nombre y domicilio del archivo, registro o banco de datos y, en su caso, el nombre del responsable o usuario del mismo.

En el caso de los archivos, registros o bancos públicos, se procurará establecer el organismo estatal del cual dependen.

2. El accionante deberá alegar las razones por las cuales entiende que en el archivo, registro o banco de datos individualizado obra información referida a su persona; los motivos por los cuales considera que la información que le atañe resulta discriminatoria, falsa o inexacta y justificar que se han cumplido los recaudos que hacen al ejercicio de los derechos que le reconoce la presente ley.

3. El afectado podrá solicitar que mientras dure el procedimiento, el registro o banco de datos asiente que la información cuestionada está sometida a un proceso judicial.

4. El Juez podrá disponer el bloqueo provisional del archivo en lo referente al dato personal motivo del juicio cuando sea manifiesto el carácter discriminatorio, falso o inexacto de la información de que se trate.

5. A los efectos de requerir información al archivo, registro o banco de datos involucrado, el criterio judicial de apreciación de las circunstancias requeridas en los puntos 1 y 2 debe ser amplio.

ARTÍCULO 39. — (Trámite).

1. Admitida la acción el juez requerirá al archivo, registro o banco de datos la remisión de la información concerniente al accionante. Podrá asimismo solicitar informes sobre el soporte técnico de datos, documentación de base relativa a la recolección y cualquier otro aspecto que resulte conducente a la resolución de la causa que estime procedente.

2. El plazo para contestar el informe no podrá ser mayor de cinco días hábiles, el que podrá ser ampliado prudencialmente por el juez.

ARTÍCULO 40. — (Confidencialidad de la información).

1. Los registros, archivos o bancos de datos privados no podrán alegar la confidencialidad de la información que se les requiere salvo el caso en que se afecten las fuentes de información periodística.

2. Cuando un archivo, registro o banco de datos público se oponga a la remisión del informe solicitado con invocación de las excepciones al derecho de acceso, rectificación o supresión, autorizadas por la presente ley o por una ley específica; deberá acreditar los extremos que hacen aplicable la excepción legal. En tales casos, el juez podrá tomar conocimiento personal y directo de los datos solicitados asegurando el mantenimiento de su confidencialidad.

ARTÍCULO 41. — (Contestación del informe).

Al contestar el informe, el archivo, registro o banco de datos deberá expresar las razones por las cuales incluyó la información cuestionada y aquellas por las que no evacuó el pedido efectuado por el interesado, de conformidad a lo establecido en los artículos 13 a 15 de la ley.

ARTÍCULO 42. — (Ampliación de la demanda).

Contestado el informe, el actor podrá, en el término de tres días, ampliar el objeto de la demanda solicitando la supresión, rectificación, confidencialidad o actualización de sus datos personales, en los casos que resulte procedente a tenor de la presente ley, ofreciendo en el mismo acto la prueba pertinente. De esta presentación se dará traslado al demandado por el término de tres días.

ARTÍCULO 43. — (Sentencia).

1. Vencido el plazo para la contestación del informe o contestado el mismo, y en el supuesto del artículo 42, luego de contestada la ampliación, y habiendo sido producida en su caso la prueba, el juez dictará sentencia.

2. En el caso de estimarse procedente la acción, se especificará si la información debe ser suprimida, rectificada, actualizada o declarada confidencial, estableciendo un plazo para su cumplimiento.

3. El rechazo de la acción no constituye presunción respecto de la responsabilidad en que hubiera podido incurrir el demandante.

4. En cualquier caso, la sentencia deberá ser comunicada al organismo de control, que deberá llevar un registro al efecto.

ARTÍCULO 44. — (Ambito de aplicación).

Las normas de la presente ley contenidas en los Capítulos I, II, III y IV, y artículo 32 son de orden público y de aplicación en lo pertinente en todo el territorio nacional.

Se invita a las provincias a adherir a las normas de esta ley que fueren de aplicación exclusiva en jurisdicción nacional.

La jurisdicción federal regirá respecto de los registros, archivos, bases o bancos de datos interconectados en redes de alcance interjurisdiccional, nacional o internacional.

ARTICULO 45. — El Poder Ejecutivo Nacional deberá reglamentar la presente ley y establecer el organismo de control dentro de los ciento ochenta días de su promulgación.

ARTICULO 46. — (Disposiciones transitorias).

Los archivos, registros, bases o bancos de datos destinados a proporcionar informes, existentes al momento de la sanción de la presente ley, deberán inscribirse en el registro que se habilite conforme a lo dispuesto en el artículo 21 y adecuarse a lo que dispone el presente régimen dentro del plazo que al efecto establezca la reglamentación.

(Nota Infoleg: por art. 2° del Decreto N° 1558/2001 B.O. 3/12/2001 se establece en CIENTO OCHENTA (180) días el plazo previsto en el presente artículo)

ARTICULO 47. — Los bancos de datos destinados a prestar servicios de información crediticia deberán eliminar y omitir el asiento en el futuro de todo dato referido a obligaciones y calificaciones asociadas de las personas físicas y jurídicas cuyas

obligaciones comerciales se hubieran constituido en mora, o cuyas obligaciones financieras hubieran sido clasificadas con categoría 2, 3, 4 ó 5, según normativas del Banco Central de la República Argentina, en ambos casos durante el período comprendido entre el 1º de enero del año 2000 y el 10 de diciembre de 2003, siempre y cuando esas deudas hubieran sido canceladas o regularizadas al momento de entrada en vigencia de la presente ley o lo sean dentro de los 180 días posteriores a la misma. La suscripción de un plan de pagos por parte del deudor, o la homologación del acuerdo preventivo o del acuerdo preventivo extrajudicial importará la regularización de la deuda, a los fines de esta ley.

El Banco Central de la República Argentina establecerá los mecanismos que deben cumplir las Entidades Financieras para informar a dicho organismo los datos necesarios para la determinación de los casos encuadrados. Una vez obtenida dicha información, el Banco Central de la República Argentina implementará las medidas necesarias para asegurar que todos aquellos que consultan los datos de su Central de Deudores sean informados de la procedencia e implicancias de lo aquí dispuesto.

Toda persona que considerase que sus obligaciones canceladas o regularizadas están incluidas en lo prescripto en el presente artículo puede hacer uso de los derechos de acceso, rectificación y actualización en relación con lo establecido.

Sin perjuicio de lo expuesto en los párrafos precedentes, el acreedor debe comunicar a todo archivo, registro o banco de datos al que hubiera cedido datos referentes al incumplimiento de la obligación original, su cancelación o regularización.

(Artículo incorporado por art. 1º de la Ley N° 26.343 B.O. 9/1/2008)

ARTICULO 48. — Comuníquese al Poder Ejecutivo.

DADA EN LA SALA DE SESIONES DEL CONGRESO ARGENTINO, EN BUENOS AIRES, A LOS CUATRO DIAS DEL MES DE OCTUBRE DEL AÑO DOS MIL.

— REGISTRADO BAJO EL N° 25.326 —

RAFAEL PASCUAL. — JOSE GENOUD. — Guillermo Aramburu. — Mario L. Pontaquarto.

Antecedentes Normativos

- Artículo 47 vetado por art. 2° del Decreto N° 995/2000 B.O. 2/11/2000.

Fuente (consultada por última vez el 27/10/2021):

<https://www.argentina.gob.ar/normativa/nacional/ley-25326-64790/actualizacion>