

Elementos para una Sociología de los Datos, la Datificación y el *big data*

Elements for a Sociology of Data, Datafication, and Big Data

Gastón Becerra*

RESUMEN: El artículo analiza el fenómeno del *big data* desde una perspectiva sociológica, tomando como marco conceptual la teoría de sistemas sociales (TSS). En la primera parte, presentamos el fenómeno del big data y lo abordamos como una construcción discursiva que articula un mandato sociotécnico con promesas de objetividad y control. Ofrecemos una conceptualización del fenómeno en términos de autodescripción de la sociedad y discutimos brevemente la tesis de un impacto disruptivo en la estructura de la sociedad. En la segunda parte, se propone una primera conceptualización para una sociología de los datos y de la producción y reutilización de ellos, denominada datificación. Articulando aportes de los estudios críticos de los datos y la TSS se exploran cuatro tesis: la datificación como un proceso autopoiético, la constitución del dato en la comunicación, la exclusión/inclusión en los datos, y la emergencia de un sistema funcional.

PALABRAS CLAVE: big data, datificación, teoría de sistemas sociales, diferenciación funcional, sociología

ABSTRACT: This article analyzes the phenomenon of big data from a sociological perspective, using Social Systems Theory (SST) as its conceptual framework. In the first section, we present the phenomenon of big data and approach it as a discursive construction that articulates a sociotechnical mandate along with promises of objectivity and control. We offer a conceptualization of this phenomenon as a self-description of society and briefly discuss the thesis of its disruptive impact on the social structure. In the second section, we propose an initial conceptual framework for a sociology of data and the social practice of data production and data repurpose, referred to as datafication. Drawing on contributions from critical data studies and SST, we explore four theses: datafication as an autopoietic process, the constitution of data within communication, inclusion/exclusion dynamics in data, and the possible emergence of a functional system.

KEYWORDS: big data, datafication, social systems theory, functional differentiation, sociology

INTRODUCCIÓN

El “big data” irrumpió a principios de la década de 2010 con promesas de revolucionar todos los aspectos de nuestra realidad social, material y subjetiva (Kitchin, 2014b; Mayer-Schonberger y Cukier, 2013). Sin embargo, hacia 2020 había perdido gran parte de su crédito, en parte por escándalos relacionados con la privacidad y la manipulación política (Pentzold y Knorr, 2024). Si bien su semántica ha sufrido tal devaluación que hoy puede considerársela una *keyword* obsoleta, no es menos cierto que la producción, circulación y reutilización de datos masivos en formato digital continúa siendo una práctica extendida que ha contribuido, entre otras consecuencias, al reciente impulso de la inteligencia artificial.

En el trabajo, proponemos una caracterización del *big data* y algunos lineamientos para pensar el proceso de producción y circulación de datos en la sociedad desde la Teoría de Sistemas

* Investigador adjunto del Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Argentina. Universidad de Buenos Aires (UBA), Argentina. gaston.becerra@gmail.com, <https://orcid.org/0000-0001-9432-8848>

Sociales (TSS). En la primera parte del artículo, se analiza su emergencia como una construcción discursiva que combinó una promesa de transformación total, una retórica de disrupción tecnológica y un mandato sociotécnico de datificación. En particular, nos interesa discutir el diagnóstico que presenta al *big data* como una transformación de la estructura e integración de la sociedad contemporánea, expresado por Jasper Taekke. Alternativamente, argumentamos que esa narrativa puede entenderse como una autodescripción de la sociedad orientada a proyectar su unidad como la sociedad de la información, aunque con elementos novedosos respecto a formulaciones previas. Luego, si bien coincidimos con quienes ven cierta “inflación de expectativas” en la retórica del *big data*, sostenemos que el fenómeno no debe ser desestimado sociológicamente, ya que la autodescripción puede ser evidencia de cambios en las prácticas y en el entendimiento social de los datos.

En la segunda parte del trabajo, proponemos algunas ideas dirigidas hacia el desarrollo de una sociología sistémica de los datos y la datificación. Retomamos aportes recientes de la epistemología y los estudios críticos de los datos, especialmente de Sabina Leonelli y Rob Kitchin, e identificamos cuatro problemas centrales del *big data*: ¿qué tipo de proceso es la datificación? ¿cómo se constituye el dato? ¿cómo dar cuenta de su carácter valorativo? y ¿cómo comprender su complejidad social? Atendemos cada uno de esos temas proponiendo una clave interpretativa basada sobre la TSS que, por momentos, retoma y extiende algunos aportes de Jannis Kallinikos sobre la información y los datos. A lo largo del escrito, señalamos continuidades y rupturas respecto de las operaciones del sistema científico con el objetivo de contextualizar la novedad del fenómeno.

EL BIG DATA COMO UNA AUTODESCRIPCIÓN: LA EMERGENCIA DEL BIG DATA

Las innovaciones tecnológicas suelen ser adoptadas por la sociedad de la mano de una construcción discursiva en la que, a la vez, se proyectan expectativas y angustias y se resalta su novedad y carácter disruptivo, dando lugar a nuevos símbolos y prácticas. Observar esta construcción puede ser un buen puntapié para delimitar el fenómeno y aclarar conceptualmente a qué nos referimos cuando hablamos de *big data*.

Encontramos los primeros usos de esta designación hacia 2010. En el campo de la computación, se la utilizó para designar los desafíos técnicos implicados en el manejo de grandes volúmenes de datos. Más adelante, el término fue adoptado por las consultoras de *management* que lo presentaban como una oportunidad de negocios (Diebold, 2012). Inicialmente, junto al neologismo, las definiciones solían incluir un listado de problemas computacionales que conformaron las llamadas “3V”: volumen, variedad (de formatos de datos) y velocidad (de procesamiento de datos); el *management* agregó una cuarta V al asociarlo con el valor (Kitchin y McArdle, 2016; Rijmenam, 2014). Los medios masivos se encargaron de diseminar la novedad recurriendo a una estrategia discursiva efectiva que popularizó el término en muy poco tiempo. Se reportaba una “premisa” que informaba sobre la disponibilidad de grandes volúmenes de datos, producto del

registro y digitalización de nuestros comportamientos *online*, o de sensores en el mundo material; y una “promesa”, por la cual esos datos podían explotarse para generar nuevos conocimientos, pretendidamente más objetivos y accionables (Becerra, 2021, 2022). Así desde las primeras definiciones que no incluían mucho más que unidades de medida desorbitantes y sistemas operativos que tornaban viable su procesamiento, el *big data* se fue constituyendo como un ámbito de prácticas analíticas en varias áreas de la sociedad y también como una mitología: “la creencia generalizada de que los grandes conjuntos de datos ofrecen una forma superior de inteligencia y conocimiento, capaces de generar ideas previamente imposibles, con un aura de verdad, objetividad y precisión” (boyd y Crawford, 2012: 663). El discurso es sostenido por una ideología dataísta que legitima ciertos criterios epistemológicos y nuevos actores en la distribución social de la producción del conocimiento (van Dijck, 2014).

Tanto en la prensa como en las comunicaciones comerciales, el *big data* fue asociado desde el inicio a otros desarrollos tecnológicos, como las tecnologías cripto, la inteligencia artificial, y diversos algoritmos. Su retórica se caracterizó, al igual que las otras tecnologías, por adoptar un tono disruptivo que habilitó, en simultáneo, posiciones tecnooptimistas y pesimistas que describen distopías. En el caso del *big data*, la disrupción tuvo un componente epistémico fuerte, ya que sus principales reclamos involucraban un conocimiento superior, sea porque se invocaban valores tales como objetividad, transparencia, sea porque se lo valoraba como “más accionable”. A la vez, el *big data* prometió alcance universal, pues revolucionaría todos los aspectos de nuestra vida material, subjetiva y social (Mayer-Schonberger y Cukier, 2013). Incluso, prometió cambiar nuestra relación con nosotros mismos, como evidenciaba el movimiento denominado *quantified self*, según el cual podríamos entendernos a nosotros mismos como un conjunto de datos -biométricos, psicoemocionales, comportamentales- para operar sobre nuestros rendimientos (Lupton, 2016).

En la sociedad, la disrupción fue reportada en múltiples sistemas. En economía, se llegó a hablar de los datos como el “nuevo petróleo” (The Economist, 2017) y se advirtió sobre la emergencia de un nuevo capitalismo de datos (Sadowski, 2019). En política, se nos habló de una nueva sociedad de control neoliberal, o un nuevo panóptico digital (Byung-Chul, 2014), que rápidamente decantó en la retórica que identificó *big data* con *Big Brother* (Portmess y Tower, 2015). Gracias al uso de datos, las ciudades evolucionarían hacia el modelo de *smart cities*, una configuración susceptible de ser monitoreada y gestionada en tiempo real (Kitchin, 2014c). Mención aparte merece el impacto en la ciencia donde, a raíz del *big data*, han surgido nuevas figuras profesionales e institucionales que estarían modificando la división de la producción del conocimiento, disputando con científicos y agencias, y donde se ha puesto en disputa el valor del método científico “tradicional”, sugiriéndose incluso que estaríamos frente a una revolución paradigmática (Kitchin, 2014a).

¿UNA NUEVA SOCIEDAD INTEGRADA POR LOS DATOS?

El impacto social más disruptivo es la promesa de una suerte de integración de lo personal, lo político, lo económico, y cualquier otra lógica o esfera de la sociedad gracias a la circulación y el entrecruzamiento de datos. Así, por ejemplo, relatando el caso de Cambridge Analytica, The Guardian (2018) describe la promesa del *big data* señalando que:

Unas pocas docenas de ‘me gusta’ pueden ofrecer una predicción bastante precisa sobre a qué partido votará un usuario, revelar su género y si su pareja es probablemente un hombre o una mujer, proporcionar indicios sólidos sobre si sus padres permanecieron juntos durante su infancia y predecir su vulnerabilidad al consumo problemático de sustancias (nuestra traducción).

Esta integración plantea un problema sociológico central, ya que, en última instancia, interroga la imagen y la estructura misma de la sociedad moderna. Para la TSS, la forma de reducción de complejidad que organiza a la sociedad es la diferenciación funcional: distintos sistemas sociales, como la economía, la política, la ciencia, entre otros, desarrollan una complejidad interna propia orientada a ofrecer soluciones especializadas a problemas sociales con su propia racionalidad y medios, códigos y programas específicos. La irrupción del *big data* podría representar, para algunos autores, un desafío a esta forma de diferenciación.

La expresión más radical de este argumento fue ofrecida por Jasper Taekke (2022, 2024), quien parte de la tesis de que los nuevos medios y tecnologías de difusión pueden ser disparadores del cambio social, para luego proponer que la adopción masiva de los medios digitales podría estar gestando una forma novedosa de diferenciación social a la que denomina “diferenciación algorítmica”:

Mi tesis es que los medios digitales marcan el fin del predominio de la diferenciación funcional porque, por un lado, se utilizan en todos los sistemas funcionales: a nivel micro en las interacciones, a nivel meso en las organizaciones y, a nivel macro como un equivalente funcional de las teorías de la reflexión en los sistemas funcionales. Por otro lado, esto implica que la sociedad en su conjunto se digitaliza y se vincula a una estructura social que se ha transformado en una arquitectura funcional digital. (...) Es a través de esta arquitectura funcional digital que los seres humanos son contruidos como personas, y que ellos y sus contribuciones son incluidos o excluidos del sistema social (Taekke, 2022: 12, nuestra traducción)

Los casos empíricos abordados explícitamente por Taekke incluyen el impacto del *big data*, por ejemplo, en el sistema de educativo¹. En publicaciones más recientes, su recorte de los medios digitales alcanza fenómenos como la inteligencia artificial (Taekke, 2025). No hay duda sobre

¹ Por caso, Taekke identifica en el ámbito educativo un desarrollo cooperativo entre varios actores de sistemas distintos, coordinados en torno a un imaginario sociotecnológico de control que “... allana el camino para un fuerte acoplamiento operativo entre los sistemas, que parece debilitar las fronteras que los separan” (Taekke, 2022: 13). El autor caracteriza el movimiento como un caso de pérdida de autonomía del sistema educativo que reconfiguraría las prácticas pedagógicas y los roles de docentes y estudiantes, subordinados a criterios externos, tales como los fines corporativos o los intentos de control político.

el gran impacto del *big data* en la sociedad, pero no parece haber evidencia suficiente para afirmar que las distintas lógicas sistémicas hayan sido efectivamente subvertidas, ni mucho menos cooptadas por una nueva capa de operatividad como sugiere Taekke. De hecho, si nos remitimos al derrotero del *big data* o de la inteligencia artificial en el sistema de la educación, no vemos un abandono del interés pedagógico². Tampoco lo vemos en la política. De hecho, el caso de Cambridge Analytica ha funcionado como un disparador para la elaboración de legislaciones específicas sobre privacidad de datos -especialmente en el marco europeo- y la reflexión sobre la soberanía de los datos que, en última instancia, reactivan discusiones tradicionales sobre el eje gobierno/oposición sólo que en el marco de una geopolítica de datos. Tampoco lo vemos en el ámbito científico, en el cual no se han producido revoluciones paradigmáticas a partir del *big data*, ni se ha abandonado la producción teórica. Más aún: los avances logrados por la analítica de grandes datos en campos como las ciencias sociales computacionales continúan evaluándose en términos de aportes a debates “clásicos” de dichas disciplinas.

BIG DATA COMO AUTODESCRIPCIÓN

¿Debemos entonces desestimar el fenómeno del *big data* como objeto de análisis sociológico? En términos de la TSS, la narrativa puede entenderse como una autodescripción de la sociedad en la comunicación, definida por Luhmann (2007) como “construcciones imaginarias de la unidad del sistema que posibilitan la comunicación en la sociedad — aunque no con ella sino sobre ella” (Luhmann, 2007: 687). Así las autodescripciones caracterizan la sociedad ofreciendo un relato sobre su unidad y su integración que nos permite situar en ella nuestra experiencia. En la sociedad moderna funcionalmente diferenciada, ellas pueden surgir de las reflexiones producidas en algún sistema parcial, proyectando sus particularidades como características de época de la sociedad en su conjunto. La sociología ha propuesto varias fórmulas en ese sentido, por ejemplo, advirtiendo sobre una sociedad estructurada en clases sociales. Otra fuente importante de autodescripciones son los medios de comunicación de masas, especialmente bajo la forma de comunicaciones informativas (Luhmann, 2000).

Una autodescripción muy cercana al fenómeno del *big data* es la que nos habla de una “sociedad de la información”, noción ampliamente difundida por autores como Castells (1996) y popularizada por los medios de comunicación, que asocia el desarrollo tecnológico con una nueva forma de organización social centrada en la producción, circulación y consumo de información. Desde la perspectiva de Luhmann (2007), la llamada sociedad de la información no remite a la

² Por caso, para el ámbito de la educación universitaria, UNESCO (2024) destaca el enorme potencial de la inteligencia artificial para la enseñanza, la investigación y la gestión universitaria, pero remarca que esta posibilidad está supeditada a que se la aborde con una mirada integral y reflexiva que no pierda de vista lo esencial de la educación: el desarrollo integral de los estudiantes en un marco vincular. Así, reclama: “tener presente que los aspectos relacionales y sociales, y el valor de las formas tradicionales de educación son fundamentales en las relaciones entre docentes y educandos y entre los propios educandos, y deberían tenerse en cuenta al examinar la adopción de las tecnologías de la IA en la educación” (UNESCO, 2021, p. 34)

emergencia de un sistema tecnológico, sino a una consecuencia de la diferenciación funcional de la sociedad moderna: a medida que los distintos sistemas sociales se autonomizan y se comunican bajo sus propios códigos, la información se multiplica, se especializa y se vuelve inabarcable. La abundancia informativa, facilitada por infraestructuras técnicas, no expresa una nueva función social, sino la creciente complejidad y sobrecarga comunicativa que resulta de esa diferenciación. No sería arriesgado trazar un paralelo con el caso del *big data*: si bien su narrativa -especialmente mediante las “3V”- surgió de los desafíos del manejo computacional de datos, rápidamente se expandió hacia una descripción de lo social en su conjunto, configurando lo que podría pensarse como una “sociedad de datos”.

Esto lleva a interrogarnos qué diferencias introduce el manejo del sentido cuando se trabaja con (grandes) datos en lugar de información. Un primer elemento diferenciador puede ser el alcance o la pretensión de aprehender exhaustivamente al objeto descrito. A diferencia de la información en los albores de la masificación de las tecnologías de la información y la comunicación de los años ‘70 y ‘80, el *big data* se construye sobre un supuesto de ubicuidad, según el cual existen datos para todo, y todo es potencialmente convertible en dato. Un avance en las condiciones que permiten este mandato ha sido la irrupción de dispositivos inteligentes que se integran a la vida cotidiana hasta volverse prácticamente invisibles; y, en las últimas décadas, el desarrollo de tecnologías en la nube que abaratan los costos de almacenamiento y procesamiento (Lyytinen y Yoo, 2002). La formulación más extrema de esta pretensión se observa en su disputa con el método científico tradicional, cuando se afirma que las técnicas de muestreo habrían perdido sentido, porque el *big data* trabajaría con la totalidad ($N = 100\%$) y eliminaría tanto el error muestral como la incertidumbre inferencial. Si bien ese reclamo es fácilmente rebatible, revela una aspiración latente: los datos podrían duplicar la realidad de manera completa. Más interesante es el cambio en la dimensión temporal. De acuerdo con Luhmann (2007: 864), “la información es un producto que se desintegra. En cuanto se actualiza, se desvanece”. Así la informatividad u originalidad resultan ser medidas relativas al estado de un sistema. Sin embargo, la temporalidad de los datos parece diferir de la temporalidad de la información. El valor del dato reside en su disponibilidad para el procesamiento futuro. Por ello, se promueve su recolección incluso sin un plan analítico definido, pues se mantiene la expectativa de una explotación *a posteriori* de ellos. De hecho, una de las principales oportunidades del *big data* sería la reutilización y readecuación [*re-purposing*] de los datos en relación con nuevos intereses. De esta manera, los datos quedan abiertos a las diferencias de un observador aún no especificado, esperando que puedan ser transformados y revalorizados en otros ámbitos bajo lógicas y prácticas distintas.

HACIA UNA SOCIOLOGÍA DE LOS DATOS Y LA DATIFICACIÓN

En suma, si bien el *big data* no parece haber generado el impacto estructuralmente disruptivo que proclamaba su *hype* o “inflación de expectativas”, hay elementos en su narrativa que introducen algunos elementos interesantes acerca de las maneras en que la sociedad produce, circula y comprende los datos. Por ello, lejos de desecharlo como objeto sociológico, proponemos entender al

big data como un indicio de la necesidad de desarrollar una sociología de los datos y de la datificación en la sociedad. Para contribuir con dicho objetivo, en lo que sigue esbozamos programáticamente una propuesta de conceptualización desde la teoría de sistemas sociales de Luhmann (1998), atendiendo diversos problemas de la filosofía y los estudios críticos de los datos. A lo largo del desarrollo, también señalaremos algunas diferencias y continuidades con las operaciones del sistema científico a fin de contextualizar mejor su novedad.

DATIFICACIÓN COMO PROCESO SOCIAL

El creciente interés por indagar la producción y circulación de datos en la sociedad se refleja en la emergencia de conceptualizaciones en torno a la “datificación” como proceso social. Revisiones sistemáticas (Becerra et al., 2025; Flensburg y Lomborg, 2023) ubican el desarrollo del concepto en la discusión crítica de la definición originaria propuesta por Mayer-Schonberger y Cukier en “Big Data. A Revolution That Will Transform How We Live, Work, and Think” (2013), donde afirman que “datificar un fenómeno significa ponerlo en un formato cuantificado para que pueda ser tabulado y analizado” (Mayer-Schonberger y Cukier, 2013: 48). Desde entonces, la literatura ha complejizado ese proceso en dos direcciones: por un lado, diseccionando las operaciones epistémicas necesarias para construir el dato y, por otro lado, ubicándolo en un contexto socio-político de producción y circulación. En este apartado, colocaremos el foco sobre el primer interés, retomando las contextualizaciones en los próximos apartados.

Diversos autores coinciden en describir la datificación como un proceso selectivo, mediante el cual la realidad es transformada materialmente para volverla maleable, recombinable y operativa (Baack, 2015; D’Ignazio y Klein, 2020; Lycett, 2013). La metáfora de la transformación de la realidad -a veces presentada como un cambio de densidad o una “liquificación”- apunta, en general, a subrayar que los datos poseen una materialidad propia que condiciona su comunicación y manipulación. En la actualidad, la materialidad de los datos los ubica en un sustrato digital, es decir, en la codificación de bajo nivel en términos de unos y ceros, que es posible operar y transformar a través de computaciones. Luego, los datos son comunicaciones estructuradas: desde simples colecciones de registros, sin mayores restricciones sobre los elementos que contienen, hasta *datasets* organizados como matrices o estructuras más flexibles, como los *data objects*, que distinguen entidades, propiedades y relaciones, y que en ocasiones se vinculan con esquemas y ontologías que añaden especificaciones semánticas a los valores contenidos. Todos permiten distinguir entre un “adentro”, conformado por elementos y valores discretos de observaciones, y un “afuera”, donde se relacionan con otros conjuntos de datos o con observaciones de segundo orden que producen información adicional sobre el conjunto, es decir, metadatos.

Luego, la literatura también caracteriza la producción del *big data* como un proceso iterativo, generalmente mediante la metáfora de los “ciclos de vida”. Leonelli (2020b) y Kitchin (2021), entre muchos otros, recurren a ella para remarcar que los datos son generados a partir de otros datos y que conforman “linajes” de *datasets* que comparten ciertos aires de familia -estructuras, campos, propiedades- y que se encuentran en constante reproducción y cambio. Además, la me-

táfora es acertada en la medida en que advierte el problema poco explorado del deceso de los datos y su reemplazo por nuevas generaciones. La utilidad de los datos en el futuro depende, en muchos casos, de introducir transformaciones para que no se desajusten respecto de las necesidades cambiantes de su entorno. Dicho de otra manera: la estabilidad de los datos es dinámica.

Hasta hace pocos años, la TSS reservaba estas caracterizaciones para la información y dejaba a los datos el lugar de simples medios de comunicación [medium]. Así Kallinikos (2006) argumentaba que la información se expande a través de un proceso autorreferente, en el que toda novedad remite a un estado actual que funciona como horizonte para su actualización, pero, debido a que dicha actualización es un proceso que pierde novedad constantemente, se ve forzada a regenerarse de manera continua, produciendo más a partir de sí misma. Esto se puede atribuir -sigue Kallinikos (2006: 100)- en gran medida a la disponibilidad de tecnologías de información con interoperabilidad, entre las que ubica las bases de datos de manera explícita. Más recientemente, Alaimo y Kallinikos (2024) parecen extender esa dinámica a la misma producción de datos:

A diferencia de las formas anteriores de datos, los datos digitales pueden ser editados, trasladados y recontextualizados de manera constante. En este sentido, el proceso de producción y uso de datos es un recorrido no lineal y recursivo, que permanece abierto, generando continuamente nuevos datos (y, por lo tanto, nuevos ciclos de vida de los datos) y adaptándose a diversos mecanismos de retroalimentación y optimización (Alaimo y Kallinikos, 2024: 63)

Para nosotros, la datificación puede ser entendida como un sistema autopoietico en el que sus elementos, al transformarse y vincularse, dan origen a nuevos elementos del mismo tipo. La datificación comienza con la construcción de los datos, pero continúa en su procesamiento y manipulación. En esas etapas, se generan nuevos datos sintéticos o derivados, resultantes de operaciones diversas, como la agregación, el modelado, el cruce con otras bases, la limpieza, y en vista de múltiples objetivos, como la clasificación, la reducción de dimensiones, la preservación de la privacidad, la interoperabilidad con otros entornos, etc. En cada momento, el sistema recrea, tanto su autorreferencia transformando los elementos de los que parte como su (hetero)referencia, lo “datificado”. El proceso podría ser descrito como la aplicación del código binario dato/no-dato, en el que se ponen en juego reglas y programas para asegurar que se pueda aceptar el dato como tal en comunicaciones subsiguientes.

DATIFICACIÓN EN LA COMUNICACIÓN

Entre las características destacadas de los datos digitales se encuentran su movilidad y su maleabilidad. La literatura sobre el *big data* remarca ambos aspectos para cuestionar suposiciones previas respecto al momento en que un dato se constituye como tal. A propósito de este debate, Leonelli (2015a, 2015b, 2020b) distingue dos perspectivas. Por un lado, una visión representacional, donde los datos son concebidos como registros confiables de la realidad o de ciertas propiedades de un fenómeno al menos. Desde ese enfoque, los datos son considerados objetos materiales y portables, generados en el momento de la medición o descripción y contendrían un significado inherente y estable. Así lo que cuenta como dato queda definido en su origen con independencia

de su utilización posterior. Esta perspectiva, dominante en la filosofía de la ciencia del siglo XX, entiende a los datos como fundamento objetivo para la construcción de conocimiento.

La mayor circulación y reutilización de los datos digitales lleva a Leonelli a oponer una perspectiva relacional, en la cual cualquier objeto puede ser considerado un dato en tanto y en cuanto cumpla dos condiciones: que pueda circular entre agentes sociales y que sea utilizado como evidencia en una argumentación para sostener reclamos en el conocimiento. Dejaremos en suspenso la cuestión de si el uso evidencial-argumentativo es el único posible. Lo que nos interesa remarcar de la propuesta es que convierte la condición de “ser un dato” en una atribución contextual. De este modo, los mismos objetos pueden ser tomados (o no) como datos en contextos distintos, a la vez que un mismo objeto puede ser tomado como fuente de datos para fenómenos distintos en otros contextos.

Leonelli explícitamente entiende a los datos como “herramientas para la comunicación” (2015b: 811). Podemos avanzar un poco más y ubicar la constitución de los datos en la unidad del acto comunicativo. De manera esquemática, la TSS entiende la comunicación como un proceso complejo que se constituye mediante la combinación de tres selecciones simultáneas. Primero, un recorte informativo particular (selección de información) es propuesto por un emisor, a través de alguna forma específica de alocución (selección del acto comunicativo). Luego, un receptor actúa en referencia a dicho acto, distinguiendo en ella una cierta información (comprensión). La comunicación emerge en este último momento: en ella se reconstituye la unidad de las tres selecciones, a la vez que se reinicia el ciclo con una nueva alocución que referencia la aceptación o el rechazo de la propuesta informativa. Los medios, las codificaciones, las convenciones y los símbolos propios ayudan a asegurar la propuesta. En ese marco, la comprensión no implica identidad entre las informaciones emitidas-recibidas; sino, más bien, la comunicación observada como un proceso de construcción iterativa de sentido a través de diferencias sucesivas.

La datificación que describe la concepción representativa se limita a las dos primeras selecciones del esquema comunicativo: primero, el recorte de cierta información mediante una observación sistemática; segundo, su codificación en algún formato que permita su resguardo y comunicación. Desde esta perspectiva, el dato *sobre algún fenómeno* está ya constituido. En cambio, la perspectiva relacional constituye el dato desde la tercera instancia: el momento en que un objeto portable y estructurado es abordado para su análisis en cierto contexto, persiguiendo determinados intereses y reconstruyendo la información codificada. En esa tercera selección se da unidad y dirección al proceso y se generan nuevos datos -derivados, sintéticos y/o metadatos- a partir de los objetos previos.

Si se acepta la descripción, el valor del dato es adjudicado en comunicaciones subsiguientes de manera análoga a la imputación temporal del carácter de verdad a ciertas construcciones propias por parte del sistema científico. Sin embargo, mientras que la ciencia opera con una presión interna hacia la coherencia, en el mundo de los datos tal exigencia no es tan apremiante. Podemos observar, asimismo, las diferencias entre los criterios y valores que orientan la práctica científica y la datificación en otros ámbitos sociales. Por ejemplo, mientras en la ciencia la transparencia metodológica es un criterio central de validez y el consentimiento informado es un requisito

ético que se evalúa en diversos momentos de la vida de los proyectos, en las plataformas digitales se construyen los datos mediante algoritmos inescrutables y procesos de extracción de datos que, en ocasiones, son forzados o coaccionados. Esa tensión entre valores epistémicos abre preguntas relevantes para la ciencia y, especialmente, para la investigación que utiliza *big data*: ¿qué problemas de validez y confiabilidad surgen cuando la producción de datos se realiza fuera del control comunicativo que el sistema científico exige para sus observaciones?

DATIFICACIÓN COMO COMUNICACIÓN CON VALORES Y ASIMETRÍAS

Un enfoque más amplio de la datificación la sitúa dentro de un entramado de relaciones sociales. Ello permite cuestionar otro supuesto central del dataísmo y de ciertas tradiciones filosóficas: la objetividad de los datos. En la retórica del *big data*, esa promesa se expresó en la idea de que el análisis de grandes volúmenes de datos permitiría que ellos “hablen por sí mismos” y “ofrezcan respuestas a preguntas no explicitadas”. La pretensión era desplazar el componente subjetivo de la interpretación, asumiendo que los datos no solo serían preanalíticos, sino también prepolíticos o valorativamente neutros, aunque se reconociera que sus usos posteriores sí pudieran adquirir sesgos o intencionalidades. En respuesta, los estudios críticos y feministas de los datos han enfatizado que estos son resultados de selecciones y observaciones realizadas desde un posicionamiento, y que no puede omitirse la distinción primaria de quién decide qué cuenta como dato y qué no, bajo qué criterios, mediante qué procedimientos y a partir de qué soporte material seleccionado (D’Ignazio y Klein, 2020). En términos de Dalton y Thatcher (2014: 5): “tanto en su producción como en su interpretación, todos los datos, incluidos los *big data*, son siempre el resultado de prácticas sociales contingentes y discutibles, que producen y ocultan interpretaciones específicas del mundo”.

El problema de la objetividad es un asunto tratado profundamente en la reflexión científica. Las mejores soluciones apuestan por multiplicar comunicaciones. Según la TSS, el sistema científico establece comunicaciones orientadas a neutralizar o “autopurificar” los valores sociales (Luhmann, 1997: 420), desplazándolos como condicionamientos externos o relegándolos al ámbito de la observación psíquica. En cualquier caso, el inevitable paso a observaciones de segundo orden permite tematizar o explicitar tales condicionamientos, perdiendo su carácter normativo. En consonancia, las estrategias contemporáneas que buscan mitigar la propagación de sesgos en el tratamiento de datos apuntan a explicitar las decisiones tomadas en el contexto de producción, circulación y reutilización. Una de las propuestas más difundidas en este sentido es la de “Datasheets for datasets” que propone construir hojas de datos para documentar de manera sistemática las decisiones, criterios y limitaciones asociadas a la creación de conjuntos de datos (Gebru et al., 2021). También, desde las ciencias sociales críticas de los datos se han propuesto varios programas metodológicos que buscan integrar investigación cualitativa en la producción de datos con el objetivo de lograr datos enriquecidos mediante la comprensión densa del contexto social de su producción y uso, es decir, transformar el *big data* en *thick data* (Jemielniak, 2020). En ambos casos, se trata de incorporar datos sobre los datos en la comunicación.

La mirada comunicativa también ofrece la oportunidad de considerar asimetrías socialmente generadas y dinamizadas mediante la forma inclusión/exclusión en al menos dos sentidos. En primer lugar, como ocurre con otros medios digitales, en la participación en la comunicación de los datos se imponen condiciones que resultan en accesos desiguales (Pignuoli Ocampo, 2022). Aquí podemos mencionar al carácter propietario de las bases de datos, el uso de formatos cerrados o de encriptación, el costo de la infraestructura y del poder de cómputo, así como el tiempo y las competencias necesarias para su procesamiento. En segundo lugar, la comunicación de los datos reproduce inclusiones/exclusiones en su dimensión heterorreferencial, particularmente en la forma en que se recortan las personas y grupos sociales como unidades de datos. Particularmente, en el procesamiento de datos se introducen nuevos supuestos, como el nivel de agregación o agrupamiento que se asume para analizar, la selección de qué atributos utilizar en el modelado, entre otros, que pueden terminar por re-construir grupos y re-clasificar individuos. Ambas formas de asimetría se refuerzan en la brecha digital del *big data* o *big data divide* (Andrejevic, 2014): los datos comunican sobre sujetos que no acceden a esa comunicación, mientras permiten a otros procesarlos, clasificarlos y redefinirlos, asignándoles nuevas posiciones dentro de circuitos comunicativos que los afectan sin su participación directa. Luego, la capacidad de definir patrones, atribuir perfiles y organizar poblaciones se convierte así en una forma inclusión/exclusión, profundizando las desigualdades entre quienes producen datos y quienes pueden explotarlos.

DATIFICACIÓN ENTRE ACOPLAMIENTOS Y EMERGENCIA DE UN NUEVO SISTEMA

Una última línea se abre con aquellos autores que han caracterizado la datificación como un proceso complejo, implicando que la producción, circulación y reutilización de datos en la sociedad está marcada por la dificultad de vincular sitios múltiples y actores diversos. Así Leonelli habla de los “viajes de los datos” [*data journeys*], a los que define como “movimiento de datos desde su sitio de producción hacia muchos otros sitios donde se procesan, movilizan y reutilizan [*re-purposed*]” (2020a: 9). El mismo concepto fue antes usado por Bates, Lin y Goodale (2016) a efectos de visibilizar la interconexión entre infraestructuras en las que se dan las prácticas de datos. Al igual que Leonelli, los autores argumentan a favor de un entendimiento de los datos como objetos socio-materiales móviles y sujetos a mutaciones en distintos lugares, en última instancia, debido a relaciones de poder. Por su parte, Kitchin se refiere a “ensambles de datos” [*data assemblages*] definiéndolos como “un sistema socio-técnico complejo, compuesto por muchos aparatos y elementos que están profundamente entrelazados” y aclara que se trata no de mera infraestructura, sino de “aparatos tecnológicos, políticos, sociales y económicos que enmarcan su naturaleza, operación y funcionamiento” (Kitchin y Lauriault, 2014: 8). La noción remite a la idea foucaultiana de dispositivo y subraya que los datos funcionan como un punto de articulación entre poder y conocimiento.

Ya argumentamos que resulta una tesis demasiado extrema concebir dichos vínculos como un suerte de nuevo principio ordenador de la sociedad capaz de constituir un centro de poder. Desde el marco de la TSS, basta con mapear las múltiples formas en que se producen los acopla-

mientos estructurales entre las lógicas de distintos sistemas (Neri y Cordeiro, 2025). No obstante, también se impone la pregunta acerca de si, a partir de estos acoplamientos, podrían emerger sistemas sociales orientados a funciones aún no contempladas por la teoría que se desarrollen en paralelo a las existentes.

Jan Overwijk (2025) explora esa posibilidad en “Cybernetic capitalism”. Su referencia son los sistemas metrológicos que se orientan hacia procesos de medición, clasificación, comparación y estandarización. Según el autor, esos sistemas generan sus propios elementos, ya que las mediciones y estándares remiten a convenciones previas, y no a una realidad externa. La clausura operativa es lograda mediante el establecimiento de conmensurabilidades. En sentido estricto, se trataría de sistemas sociales, pues operan en el medio de la comunicación, estabilizando expectativas a través de la estandarización. Lo más importante es que se trata de sistemas funcionales que han coevolucionado a partir de acoplamientos con organizaciones y otros sistemas funcionales. Así la metrología se extiende desde las equivalencias necesarias para el manejo estadístico de poblaciones hasta los estándares de calidad y otros dispositivos organizativos³. Comparativamente hablando, las operaciones de los sistemas metrológicos nos resultan más difusas y fragmentadas que las de datificación, ya que en el fenómeno que nos interesa hay un medio más claro y unívoco. Por esta razón, creemos que puede haber elementos para pensar la autonomización de un problema/solución en ciernes. Se trataría de un movimiento que, a la vez, desacopla el dato de la lógica propia de la ciencia, en la cual se dieron los primeros desarrollos sobre datos, y lo generaliza a otros espacios en una suerte de movimiento del “dato por el dato”.

Por el momento, sólo hay algunos indicios a nivel semántico que podemos comenzar a explorar. Reconstrucciones acerca de la conceptualización del dato en las ciencias, como las de Gitelman (2013) o Rheinberger (2011), muestran que han atravesado una suerte de inversión histórica. En sus orígenes, remiten a lo dado *-datum* o *given-*, a las premisas aceptadas de un argumento, sin referencia externa ni evaluación de veracidad. Recién a partir del siglo XX, comenzaron a vincularse con fenómenos empíricos, cuando el término pasó a asociarse con los hechos *-facts-*. Los datos son la forma aceptada de la evidencia no sólo en la ciencia, sino también en la política y otras prácticas. Pero también, más recientemente, son los impulsores de otras prácticas “basadas sobre datos” [*data-driven*] (Beaulieu y Leonelli, 2022). En ellas, los requisitos comunicativos de las evidencias pueden resultar excesivos. De hecho, en varios sectores de la sociedad se adoptan algoritmos opacos que explotan modelos -datos sobre las estructuras de los datos- que desafían la explicabilidad y la razonabilidad (Esposito, 2022).

Otros indicios posibles de clausura son la emergencia de roles profesionales específicos y de nuevas formas organizacionales centradas en la manipulación de datos. Destacamos la figura del científico de datos [*data scientist*], definido como “un profesional de alto nivel, con la forma-

³ Por ejemplo, en relación con la emergencia de un sistema de gestión de calidad [*quality management*] refiere que: “señala el volverse-eficiente de la salud y el volverse-promotora-de-salud de la eficiencia. Es un tercer sistema —un ejemplo de lo que denomino un ‘sistema metrológico’— que opera sobre un código simbiótico de calidad/no-calidad, estructuralmente acoplado a una multiplicidad de organizaciones (hospitales, universidades, escuelas) y sistemas funcionales (economía, salud, educación) (Overwijk, 2025, p. 37, nuestra traducción)

ción y la curiosidad necesarias para realizar descubrimientos en la era del Big Data” (Davenport y Patil, 2012: 72). Dorschel (2021) subraya dos características distintivas del rol. Por un lado, su hibridez: “los científicos de datos son contruidos simultáneamente como generalistas y especialistas; técnicos y comunicadores; explotadores de datos y éticos de los datos”; por otro lado, la necesidad permanente de justificar su propia función o necesidad ante las múltiples lógicas y demandas de las organizaciones. A nivel organizacional, Matei, Jullien y Goggins (2017) describen la proliferación de estructuras a las que denominan “fábricas de grandes datos” [*big data factories*], orientadas a la producción, transformación y circulación de datos bajo lógicas colaborativas. Estas entidades se caracterizan por su apuesta a la interoperabilidad, el uso de estándares abiertos y estructuras distribuidas. Se diferencian tanto de los centros de investigación tradicionales, como de las agencias estadísticas o los *data warehouses* empresariales, cuyo fin prioritario no es la producción de conocimiento académico, ni la toma de decisiones estratégicas, ni tampoco la generación directa de valor económico, sino más bien su función parece centrarse en la propia gestión y circulación de los datos.

¿Qué pueden hacer las ciencias frente a estos indicios de diferenciación interna que plantea la datificación? Como mencionamos, las respuestas al *hype* del *big data* -y en particular a la provocación de Anderson (2008) sobre el supuesto “fin de la teoría”- han llevado a revalorizar el conocimiento como producto reflexivo en muchos casos. Precisamente, el impulso explica el interés de las ciencias sociales por teorizar la datificación como un proceso social, resistiendo su reducción a una cuestión técnica o instrumental. Al mismo tiempo, cabe esperar que las prácticas basadas sobre datos sigan expandiéndose dentro de la ciencia, sin por ello comprometer necesariamente su carácter autopoietico, garantizado por la aplicación del código verdadero/falso (Luhmann, 1997). De hecho, tal es una de las propuestas de Kitchin (2014a), quien refiere a ciencias “basadas en datos” [*data-driven sciences*] para reivindicar un uso crítico y políticamente progresivo del *big data* en el marco de una epistemología situada, reflexiva y contextual, capaz de evitar los excesos del dataísmo, pero incorporando técnicas computacionales. Si el diagnóstico es correcto, podría esperarse que la evaluación de lo que constituye evidencia en las ciencias empiece a incorporar cada vez más la producción activa de nuevos datos a partir de datos, favoreciendo criterios epistémicos como la replicabilidad y la transferencia. Luego, otras novedades tienen una proyección menos previsible. Por ejemplo, nos preguntamos cuáles podrían ser los efectos en la ciencia si sigue creciendo el valor de los *dataset* como producto y formato de comunicación en lugar de los *papers*, en aspectos importantes como la evaluación interna del sistema científico, o como aporte demandado por la sociedad.

CONCLUSIONES

El interés por el *big data* y los procesos de datificación responde a la centralidad que los datos digitales han adquirido en las diversas prácticas contemporáneas de reducción de complejidad. Adoptando algunos lineamientos de la TSS, argumentamos que el *big data* puede entenderse

como una autodescripción de la sociedad. Ella proyecta una imagen de unidad bajo un cierto imaginario de control y nuevo orden. Si bien dicha narrativa presenta elementos típicos de una inflación de expectativas en favor de la tecnología, sostenemos que no debe ser desatendida sociológicamente, porque revela posibles cambios en la manera en que socialmente entendemos los datos en la sociedad. Así una sociología de los datos podría contribuir a indagar las condiciones de emergencia de fenómenos sociotecnológicos novedosos que han hecho propia la promesa del *big data*, especialmente la inteligencia artificial.

En ese marco, desarrollamos una propuesta programática de conceptualización de la datificación que recoge los principales debates de la literatura: su caracterización como un proceso, el problema de su constitución, su carácter valorativo y una suerte de autonomización en ciernes. En cada discusión propusimos una clave interpretativa que diera cuenta del fenómeno desde la TSS. Así en relación a la primera pregunta caracterizamos la datificación como un proceso social autorreferente y autopoietico, en el que los datos se producen a partir de otros datos. En segundo lugar, ubicamos la constitución de los datos en la comunicación, logrando su unidad mediante las tres selecciones que definen este proceso según la TSS. En tercer lugar, atendemos a las críticas de la objetividad del dato remitiendo a las inclusiones/exclusiones de la comunicación. Finalmente, argumentamos que hay indicios para pensar la emergencia de un nuevo sistema social funcional orientado a la producción de datos.

Una vez establecido el diálogo entre las epistemologías y estudios críticos de los datos y la TSS, nuestra propuesta para una sociología sistémica de los datos ahora tiene pendiente la tarea de llevar adelante una discusión “puertas adentro” de la teoría. Hasta aquí, sólo hemos mencionado unos pocos autores como posibles antecedentes para pensar la iteratividad de los datos o para vislumbrar cómo se pueden pensar nuevos sistemas funcionales. Los desarrollos de la TSS que tratan con las nuevas tecnologías son numerosos. Sin embargo, lamentablemente, se carece de un mapa o una sistematización que permita esclarecer, al menos, si sus recortes empíricos son equivalentes. Además, en línea con la tradición cibernética, la literatura suele privilegiar el interés por la información, obstaculizando la pregunta por los datos⁴.

REFERENCIAS

- Alaimo, C. y Kallinikos, J. (2024). *Data rules: reinventing the market economy*. The MIT Press.
- Anderson, C. (2008). The end of theory. The data deluge makes the scientific method obsolete. *Wired*. <https://www.wired.com/2008/06/pb-theory/>
- Andrejevic, M. (2014). The big data divide. *International Journal of Communication*, 8(1), 1673–1689.

⁴ Como mencionamos, Kallinikos (2006) —pero también Marton (2009)— argumentan la emergencia de como sistemas autorreferentes, pero en relación a las “tecnologías de información”, apenas elaborando sobre los datos. Esto recién empieza a complejizarse con Alaimo y Kallinikos (2020), aunque con un anclaje en las organizaciones como sistemas sociales. Otros, como Reichel (2011), van más lejos para hablar de la tecnología —en general, y no sólo la digital— como un sistema autopoietico, aunque discutiendo su carácter social. También podríamos mencionar a otros tantos autores que interpelan las tecnologías digitales y a los datos desde perspectivas teóricas alternativas que no pretenden ver en ellos sistemas, sino que los caracterizan como medios de difusión o propagación.

- <https://doi.org/1932-8036/20140005>
- Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data and Society*, 2(2), 1–11.
<https://doi.org/10.1177/2053951715594634>
- Bates, J., Lin, Y. W., & Goodale, P. (2016). Data journeys: Capturing the socio-material constitution of data objects and flows. *Big Data and Society*, 3(2), 1–12. <https://doi.org/10.1177/2053951716654502>
- Beaulieu, A., y Leonelli, S. (2022). *Data and Society. A Critical Introduction*. Sage Publications
- Becerra, G. (2021). The promise and the premise: How digital media present big data. *First Monday*, 26(9).
<https://doi.org/https://doi.org/10.5210/fm.v26i9.10539>
- Becerra, G. (2022). Representaciones sociales del big data en la prensa digital argentina. *Questión*, 3(72).
<https://doi.org/https://doi.org/10.24215/16696581e726>
- Becerra, G., Mezzadra, J., & Gambino, B. (2025). Mapeando la datificación: Un análisis bibliométrico de perspectivas y patrones en la comunicación científica. *Revista de Ciencias Empresariales y Sociales*, 12(9), 1–28.
<https://publicacionescientificas.uces.edu.ar/index.php/empresarialesysociales/article/view/1878>
- boyd, D., y Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Byung-Chul, H. (2014). *Psicopolítica. Neoliberalismo y nuevas técnicas de poder*. Herder
- Castells, M. (1998). *La era de la información: Economía, sociedad y cultura*. Alianza Editorial
- D'Ignazio, C. y Klein, L. (2020). *Data feminism*. The MIT Press
- Dalton, J. y Thatcher, J. (2014). What does a critical data studies look like, and why do we care? Seven points for a critical approach to 'big data.' *Society and Space*, 1–12. <http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/>
- Davenport, T. H. y Patil, D. J. (2012). Data scientist. *Harvard Business Review*, 90(October 2012), 70–76.
http://128.255.244.58/strategic/articles/data_scientist-the_sexiest_job_of_the_21st_century.pdf
- Diebold, F. X. (2012). The Origin(s) and development of “Big Data”: the phenomenon, the term, and the discipline. *Penn Economics Working Paper*, 12. <https://doi.org/10.2139/ssrn.2202843>
- Dorschel, R. (2021). Discovering needs for digital capitalism: The hybrid profession of data science. *Big Data and Society*, 8(2). <https://doi.org/10.1177/20539517211040760>
- Esposito, E. (2022). *Artificial communication. How algorithms produce social intelligence*. The MIT Press.
- Flensburg, S. y Lomborg, S. (2023). Datafication research: Mapping the field for a future agenda. *New Media and Society*, 25(6), 1451–1469. <https://doi.org/10.1177/14614448211046616>
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Iii, H. D., y Crawford, K. (2021). Datasets for datasets. *Communications of the ACM*, 64(12), 86–92. <https://doi.org/10.1145/3458723>
- Gitelman, L. (2013). *“Raw Data” Is an Oxymoron*. The MIT Press.
<https://doi.org/10.1080/1369118X.2014.920042>
- Jemielniak, D. (2020). *Thick big data. Doing digital social sciences*. Oxford University Press.
- Kallinikos, J. (2006). Information out of information: On the self-referential dynamics of information growth. *Information Technology and People*, 19(1), 98–115. <https://doi.org/10.1108/09593840610649989>
- Kitchin, R. (2014a). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1). <http://journals.sagepub.com/doi/10.1177/2053951714528481>
- Kitchin, R. (2014b). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Sage. <https://doi.org/10.4135/9781473909472>
- Kitchin, R. (2014c). The real-time city? Big data and smart urbanism. *GeoJournal*, 79(1), 1–14.
<https://doi.org/10.1007/s10708-013-9516-8>
- Kitchin, R. (2021). *Data Lives. How data are made and shape our world*. Bristol University Press.
- Kitchin, R., & Lauriault, T. P. (2014). Towards critical data studies : Charting and unpacking data assemblages and their work. *Geoweb and Big Data*, 1–19. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112
- Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26

- datasets. *Big Data and Society*, 3(1), 1–10. <https://doi.org/10.1177/2053951716631130>
- Leonelli, S. (2015a). The philosophy of data. In L. Floridi (Ed.), *The Handbook for the Philosophy of Information*. Routledge.
- Leonelli, S. (2015b). What Counts as Scientific Data? A Relational Framework. *Philosophy of Science*, 82(5), 810–821. <https://doi.org/10.1086/684083>
- Leonelli, S. (2020a). Learning from data journeys. In S. Leonelli & N. Tempini (Eds.), *Data Journeys in the Sciences* (pp. 1–24). https://doi.org/10.1007/978-3-030-37177-7_1
- Leonelli, S. (2020b). Scientific Research and Big Data. *Stanford Encyclopedia of Philosophy*.
- Luhmann, N. (1997). *La ciencia de la sociedad*. Anthropos
- Luhmann, N. (1998). *Sistemas sociales. Lineamientos para una teoría general*. Anthropos
- Luhmann, N. (2000). *La realidad de los medios de masas*. Anthropos
- Luhmann, N. (2007). *La sociedad de la sociedad*. Herder
- Lupton, D. (2016). *The quantified self*. Polity Press
- Lycett, M. (2013). “Datafication”: Making sense of (big) data in a complex world. *European Journal of Information Systems*, 22(4), 381–386. <https://doi.org/10.1057/ejis.2013.10>
- Lyytinen, K. y Yoo, Y. (2002). Issues and Challenges in Ubiquitous Computing. *Communications of the ACM*, 45(12), 3–12. <https://doi.org/10.1016/b978-0-12-600955-2.50021-6>
- Matei, S. A., Jullien, N. y Goggins, S. P. (2017). *Big Data Factories*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-59186-5>
- Marton, A. (2009). Self-Referential Technology and the Growth of Information: From Techniques to Technology to the Technology of Technology. *Soziale Systeme*, 15(1), 138–159. <https://doi.org/10.1515/sosys-2009-0109>
- Mayer-Schönberger, V. y Cukier, K. (2013). *Big data. A revolution that will transform how we live, work, and think*. Eamon Dolan/Houghton Mifflin Harcourt.
- Neri, H. y Cordeiro, V. (2025). Reimagining Sociality in the Digital Age: Transcending the Interaction/Society Dichotomy. *Systems Research and Behavioral Science*, 1–15. <https://doi.org/10.1002/sres.3142>
- Pignuoli Ocampo, S. (2022). Comunicación digital: Definición operativa y aproximación a la participación bajo la forma inclusión / exclusión digital. *MAD*, 46, 70–83. <https://doi.org/10.5354/0719-0527.2022.68542>
- Overwijk, J. (2025). *Cybernetic Capitalism*. Fordham University Press
- Pentzold, C. y Knorr, C. (2024). When data became big: revisiting the rise of an obsolete keyword. *Information Communication and Society*, 27(3), 600–617. <https://doi.org/10.1080/1369118X.2023.2227673>
- Portmess, L. y Tower, S. (2015). Data barns, ambient intelligence and cloud computing: the tacit epistemology and linguistic representation of Big Data. *Ethics and Information Technology*, 17(1), 1–9. <https://doi.org/10.1007/s10676-014-9357-2>
- Qvortrup, L. (2006). Understanding new digital media: Medium theory or complexity theory? *European Journal of Communication*, 21(3), 345–356. <https://doi.org/10.1177/0267323106066639>
- Rheinberger, H. J. (2011). Infra-experimentality: From traces to data, from data to patterning facts. *History of Science*, 49(3), 337–348. <https://doi.org/10.1177/007327531104900306>
- Rijmenam, M. van. (2014). *Think bigger*. American Management Association.
- Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data and Society*, 6(1), 1–12. <https://doi.org/10.1177/2053951718820549>
- Taekke, J. (2022). Algorithmic Differentiation of Society – a Luhmann Perspective on the Societal Impact of Digital Media. *Journal of Sociocybernetics*, 18(1), 2–23. <https://papiro.unizar.es/ojs/index.php/rc51-jos/article/view/6225>
- Taekke, J. (2024). From media evolution to the Anthropocene: Unpacking sociotechnical autopoiesis. *Systems Research and Behavioral Science*, February 2024, 383–395. <https://doi.org/10.1002/sres.3009>
- Taekke, J. (2025). Sociological Perspectives on AI, Intelligence and Communication. *Systems Research and Behavioral Science*, 1–11. <https://doi.org/10.1002/sres.3123>

- The Economist (2017). Data is giving rise to a new economy. 6/5/2017. <https://www.economist.com/briefing/2017/05/06/data-is-giving-rise-to-a-new-economy>
- The Guardian. (2018). How Cambridge Analytica turned Facebook “likes” into a lucrative political tool. 17/03/2018. <https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm>
- UNESCO. (2021). *Recomendación sobre la Ética de la Inteligencia Artificial*. UNESCO <https://www.unesco.org/es/legal-affairs/recommendation-ethics-artificial-intelligence>
- UNESCO. (2024). Guía para el uso de IA generativa en educación e investigación. UNESCO. <https://www.unesco.org/es/articles/guia-para-el-uso-de-ia-generativa-en-educacion-e-investigacion>
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance and Society*, 12(2), 197–208.