

# La pandemia de los datos

## The Pandemic of Data

Daniel Innerarity

Catedrático de filosofía política,  
investigador Ikerbasque en la UPV/EHU  
[dinner@ikerbasque.org](mailto:dinner@ikerbasque.org)

ISSN 1989-7022

**RESUMEN:** Una crisis como la del coronavirus, que ha tenido lugar en un entorno digitalizado, ha puesto a prueba la capacidad y los límites del big data para proporcionarnos una imagen completa de la realidad e indicaciones para gestionar la pandemia. Uno de los aprendizajes que hemos de realizar es entender que los datos reflejan la desigualdad existente y, sin una correcta interpretación, invisibilizan a los grupos más vulnerables. Para ello lo más importante es dejar de concebir la realidad social a partir de los criterios estandarizados y pensarla como compuesta de individuos y grupos.

**ABSTRACT:** A crisis such as the coronavirus, which has taken place in a digitalised environment, has tested the capacity and limits of big data to provide us with a complete picture of reality and indications for managing the pandemic. One of the lessons we need to learn is to understand that data reflect existing inequality and, without correct interpretation, make the most vulnerable groups invisible. To this end, the most important thing is to stop conceiving social reality on the basis of standardised criteria and to think of it as composed of individuals and groups.

**PALABRAS CLAVE:** big data, pandemia, digitalización, vulnerabilidad, posverdad

**KEYWORDS:** big data, pandemic, digitalisation, vulnerability, post-truth

La del coronavirus es la primera pandemia de la sociedad de los datos, de la era de la información. La gran cantidad de datos generados y analizados sobre el virus y su afectación hace de esta crisis la primera “data-driven pandemic”. La comunicación y el escrutinio de los datos de infección y fallecimientos por coronavirus se han convertido en un ritual diario. No es extraño que los números, las comparaciones y las categorizaciones hayan sido tan importantes en la gestión que los gobiernos han hecho de esta crisis. Los datos clarifican públicamente las situaciones reales y orientan las decisiones que deben adoptarse. El recurso a los datos permite reestablecer un mínimo de fiabilidad que justifique las medidas de control y limitación de las actividades. La cuantificación resulta especialmente seductora en tiempos de incertidumbre. Gracias a ella se organiza y simplifica el conocimiento, facilitando la toma de decisiones (Merry 2016). La medición promete gestionar la complejidad y reducir la incertidumbre.

A pesar de todo, tenemos la sensación de que se ha desmentido el mito de que la mera cantidad de datos era suficiente para hacerse cargo de la realidad porque los datos por sí mismos no nos han permitido hacernos cargo de la complejidad del fenómeno. La pandemia ha puesto de manifiesto la inadecuación de nuestras infraestructuras de datos para resolver las crisis sociales. De entrada, ha habido un problema de insuficiencia o mala calidad de los datos. De hecho los datos de la pandemia han sido escasos y han



llegado muy fragmentados por las diferentes políticas nacionales de salud pública. Hasta los datos de fallecimientos han sido inciertos.

No solo tenemos un problema de escasez de datos, sino de errores en su interpretación o de la propia configuración de nuestro espacio informativo, en el que también se difunden las más extravagantes desinformaciones. Pero puede que nuestra principal torpeza provenga, paradójicamente, de un cierto exceso y confianza acrítica en los datos existentes. Desde los primeros días de la pandemia, con mayor o menor acierto, los gobiernos han informado del aumento de los contagios, el rastreo o la ocupación de los hospitales, y así era diariamente transmitido por los medios de comunicación. Lo que ha sido menos habitual es preguntarse de qué modo, en la cuantificación, “las condiciones de producción condicionan los tipos de conocimiento” (Davis / Fisher / Kingsbury 2012, 4) o cuáles son los efectos sociales que se siguen de esta cuantificación. Mi hipótesis es que el “dataísmo” (van Dick 2014), es decir, la creencia de que la cuantificación produce la verdad, privilegia una falsa idea de la objetividad y proporciona una certidumbre engañosa que impide un conocimiento cabal de la realidad, sobre el que deberían adoptarse las correspondientes de decisiones.

Para combatir eficazmente a una pandemia hace falta conocer sus modos de propagación y hasta qué punto afecta a los distintos tipos de personas. Me pregunto si la pretensión de neutralidad con que se presentan los datos no nos ha seducido con la idea de que eran exactos y no hacía falta interrogarse por su contexto. Para sortear ese riesgo deberíamos practicar con los datos y sus tecnologías lo que alguien ha llamado una “duda post-cartesiana” (Amoore 2019, 149), es decir, una duda sobre aquello que supuestamente nos saca de ella, esa que hay que ejercer en un momento en el que los algoritmos son el modo principal de proporcionar evidencias. Hay muchos sesgos inherentes a toda la producción, análisis y visualización de datos, pero el más perturbador de todos es el supuesto de que los datos son algo neutro, una especie de árbitros apolíticos de la verdad. Las mismas prácticas de recolección, análisis y visualización de los datos llevan a ignorar determinados aspectos de la realidad. Pensemos en la sorpresa que supuso el contagio masivo de ciertos ámbitos de población como los temporeros, los presos o las personas mayores en las residencias, en la limitada eficacia de ciertas recomendaciones generales debida precisamente a que no se tomaban en suficiente consideración las distintas realidades familiares, domiciliarias o laborales.

Las técnicas usadas para la obtención de datos hacen que sean invisibles las tasas de contagio entre ciertos grupos sociales. Solemos olvidar que las medidas para reducir el riesgo reflejan y fomentan determinados comportamientos, asignan las personas a categorías y construyen un “humano estandarizado” que responde a una visión muy limitada y excluyente de la sociedad (Epstein 2009, 36). Los grupos menos visibles tienden a ser aquellos que se adaptan menos a las normas de comportamiento a partir de los cuales se realiza la analítica de datos y con frecuencia son quienes resultan más contagiados y contagiosos. Las estrategias sanitarias para encauzar el comportamiento de la población son ineficaces, por ejemplo, frente a quienes desobedecen las recomendaciones por pura necesidad económica o los migrantes que se desplazan hacia sitios de mayor riesgo. Es muy importante tener esto en cuenta cuando lleguen las vacunas, o sea, cuando haya que poner en marcha todas las decisiones que tienen que ver con su distribución, las prioridades o la correspondiente política de comunicación.

El problema de “traducir la vida social a categorías conmensurables de modo que diferentes eventos se conviertan en casos de la misma cosa” (Merry 2016, 27) puede ser algo peligroso

cuando existen formas de desigualdad o vulnerabilidad. Hay sujetos sobre-representados y otros a los que el sistema de gobernanza no detecta, como los migrantes sin documentación o quienes trabajan sin contrato. Si los datos se generan por el consumo, la movilidad o el activismo en las redes sociales, la representación favorecerá a quienes produzcan más datos en esos ámbitos. Los sistemas de datificación tienen aquí un ángulo ciego que excluye a ciertas personas, precisamente a la más vulnerables, de las estrategias de mitigación. Esta ceguera ha sido corregida, tal vez parcialmente y demasiado tarde. El abordaje de la pandemia cambió en muchos países cuando se puso de manifiesto que ciertos grupos de población (minorías étnicas, determinados trabajadores o ancianos en residencias) tenían unas tasas de contagio desproporcionadamente más elevadas. En Gran Bretaña, por ejemplo, comenzó a mencionarse la etnia de los fallecidos, lo que permitía en adelante disponer de unos datos más útiles para llevar a cabo las políticas de prevención. Esta atención a lo particular es una de las asignaturas pendientes de nuestros sistemas de cuantificación. La recolección de datos tendría que incluir a quienes no están asegurados, no tienen permiso de residencia, ni acceso a los servicios de salud, que no obstante son los más contagiados y, por tanto, más contagiosos.

La incertidumbre que los datos no logran disolver está muchas veces relacionada con grupos olvidados, sectores de la sociedad que por su identidad o trabajo son menos visibles y cuya autonomía es menor. Esto tiene implicaciones en las medidas tecnológicas que se aplican para hacer frente a la pandemia con la mediación de los datos, como por ejemplo las apps de rastreo. De entrada, las estrategias de rastreo tienden a priorizar, generalmente de un modo no intencional, a determinados tipos de personas y marginalizar a otras: los migrantes sin papeles pueden tener miedo de ser denunciados a las autoridades; muchas personas llevarán la enfermedad en la soledad de sus domicilios o en la calle y nadie recogerá datos sobre ellos; los diseñadores de las apps solo tienen en cuenta un cierto tipo de usuario, principalmente alguien con competencias digitales y capacidad económica para tener un *smartphone* con sistemas operativos puestos al día.

Siendo unos dispositivos que nos han proporcionado datos muy valiosos, las apps de rastreo tienen algunas limitaciones, como, por ejemplo, la falsa seguridad que pueden dar a sus usuarios o el hecho de que leen la localización más que el tipo de comportamiento. El virus no se transmite solo por las vías que pueden ser detectados usando el app. El rastreo que realiza el móvil es útil para sustituir el más costoso pero más exacto procedimiento de preguntar a la gente acerca de sus contactos. Por si fuera poco, para que estos mecanismos funcionen hace falta que las personas confíen en las autoridades y en la gestión centralizada de los datos, algo que está muy lejos de suceder en una sociedad cuya creciente desconfianza la pandemia no hace más que aumentar.

La pandemia ha irrumpido en un mundo en el que hay, al mismo tiempo, acceso al conocimiento científico, un entorno informativo digital caótico y una desconfianza hacia los expertos y hacia los gobiernos. Este entorno plantea dificultades especiales, también en lo que se refiere a los datos, a su fiabilidad para la gestión de la pandemia.

Un factor que puede explicar nuestro relativo fracaso para gobernar esta crisis es la instalación de una actitud "*post-truth*" en la vida social contemporánea, donde los hechos objetivos parecen influir menos en la configuración de nuestra opinión, personal y pública, que las apelaciones a la emoción y las creencias personales (Shelton 2020, 1). Una parte de este desprecio a la verdad es atribuible a la acción de algunos gobiernos, que han ocultado los datos o

los han manipulado. Más preocupante, sin embargo, es la desorientación y los errores que proceden de datos verdaderos, pero que no han sido situados en su contexto o analizados correctamente. Se pone así de manifiesto que los datos son tan concluyentes como maleables y que cualquiera puede presentarlos de modo que favorezcan lo que uno quiere decir. La beatería de los datos tiene a defenderlos como si nos aseguraran frente a la ideologización. Ahora bien, los datos no son necesariamente lo opuesto de la ofuscación ideológica; pueden favorecer la objetividad pero también ser puestos al servicio de cualquier ideología. Se trata de la parte más grosera pero menos inquietante de nuestra confusión porque lo más problemático de esta distorsión de la realidad es aquello que tiene razones estructurales y que no se deben a la intención deliberada de esconder o mentir. Me refiero a la ambigua relación con la verdad que tiene nuestro actual entorno informativo, en el que conviven posibilidades inéditas de acceso al conocimiento con la libre difusión de los errores, sean en forma de desinformación o de extravagantes teorías de la conspiración. En esta “infodemia” las noticias falsas se expanden más rápidamente que el virus (United Nations Department of Global Communications 2020).

Hay un tipo de desinformación muy vinculada a la propia naturaleza de las redes sociales y que contrasta con la potencialidad que se les había asignado a la hora de responder a estas crisis de una manera eficiente. Una de las cosas que esta pandemia pone en cuestión es aquella opinión tan extendida de que las redes sociales podrían ser sistemas de vigilancia temprana para alertar el desarrollo de las enfermedades y que las huellas digitales harían visibles amenazas como el coronavirus antes que los gobiernos o los científicos. Los datos que circulan en las redes no están exentos de sesgos y conviven con la propagación de las noticias falsas. La desinformación en torno a la pandemia se debe a la existencia de bots —parece ser que lo eran más de la mitad de las cuentas de Twitter que emitían opiniones sobre la pandemia (Hao 2020)—, pero es más inquietante aún constatar que en su propagación participan cantidad de personas. Esta desinformación ha debilitado la confianza ciudadana en las autoridades y ha reducido el efecto de las medidas sanitarias que pretendían motivar comportamientos de prudencia en la ciudadanía, como las mascarillas, la distancia social o el confinamiento.

A esta sociedad pos-verdad ha podido contribuir el datocentrismo de los últimos años, es decir, un entorno poblado de datos sin contexto y sin una narrativa coherente que diera cuenta de lo que estaba pasando. Nuestra propia gestión de los datos puede estar generando más perplejidad que comprensión. No hace falta voluntad expresa de confundir para que todos estemos en buena medida confundidos. Es cierto que periodistas y sociólogos han hecho un gran trabajo para comunicar y visualizar los datos de la pandemia. No juzgo sus intenciones sino que trato de llamar la atención sobre un efecto no pretendido de cierta gestión de los datos para lo que no hemos desarrollado todavía una cultura apropiada. La redundancia de datos que se nos ofrecen cada día en mapas, números y gráficos apenas nos permite distinguir una cifra de otra (la mortalidad de la letalidad, el contagio de la infección o las razones de que aumenten los fallecimientos cuando hay menos contagiados) y comprender el sentido de lo que está pasando. Otro ejemplo de ello es cómo el énfasis en una representación continua y actualizada de los datos puede limitar nuestra percepción a lo más urgente y hacer incomprensible los modos como este tipo de crisis resultan de procesos que actúan en una mayor escala temporal. En este contexto no es de extrañar que las teorías de la conspiración resulten más atractivas.

La salida de esta crisis sanitaria debería llevarnos a distintos modos de ser y conocer, también en relación con la manera de entender y estar en el mundo que implica la tecnología del análisis de datos. La sociedad construída sobre los datos tiene una gran dificultad para integrar en su infraestructura y gobernanza otros modos de conocimiento y existencia alternativos de los estandarizados. Medir y trazar ha sido más importante para los gobiernos que entender exactamente qué debía ser medido y trazado. Habría que invertir los términos y preguntarse no por los datos a partir de los cuales se llevarían a cabo determinadas políticas sino por qué datos requieren las decisiones políticas que hemos de tomar.

La ampliación de la mirada hacia quienes no suelen ser objeto de atención podría contribuir a que entendiéramos la sociedad desde la lógica del colectivo y no desde la noción de la mayoría. Para entender y gestionar una sociedad contagiosa es incomparablemente más útil hacerlo desde la categoría de lo común que a partir de lo mayoritario. Tenemos que desarrollar un nuevo tipo de atención hacia la realidad social que se interesa por lo común y por las situaciones particulares. Un cambio en la línea del cuidado requiere un cambio también en el modo de entender los datos. “La ciencia y la política serían capaces de controlar mejor la pandemia si consideraran las fuentes de la incertidumbre y los datos olvidados no como vacíos en el paisaje informativo sino como individuos que seguramente son miembros de grupos menos visibles y menos poderosos” (Taylor 2020, 1). Se trataría de ver a la sociedad como personas y grupos, no poblaciones, lo que permitiría tomar en cuenta las vulnerabilidades particulares y, por tanto, cuidar especialmente esos espacios de infección. Podríamos hablar entonces de una democracia de los datos, no tanto desde la perspectiva habitual que se pregunta por su propietario, sino si esos datos representan a toda la sociedad, a lo común, a todos y a todas. Esto requiere una diferente concepción de los datos, pues supondría no poner el foco en la mayoría sino en la diversidad, en experiencias concretas como las de los económicamente desaventajados o socialmente excluidos. En vez de un rebaño (*herd*) regido por la normalidad estadística, tendríamos un mosaico de diferentes vulnerabilidades. La salida de la pandemia exige un cambio en nuestra concepción de los datos y por tanto un cambio en nuestra manera de entender la sociedad.

---

#### Referencias:

- Amoore, Louise (2019), “Doubt and the algorithm: On the partial accounts of machine learning”, *Theory, Culture & Society* 36 / 6, 147–169.
- Davis, Kevin / Fisher, Angelina / Kingsbury, Benedict et al. (eds.) (2012), *Governance by Indicators. Global Power through Quantification and Rankings*, Oxford: Oxford University Press.
- Epstein, Steven (2009), “Beyond the Standard Human”, en Lampland, Martha / Star, Susan Leigh (eds.), *Standards and Their Stories. How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*, Ithaca & London: Cornell University Press, 35–53.
- Hao, Karen (2020), “Nearly half of Twitter accounts pushing to reopen America may be bots”, MIT Technology Review, 21 May. Available at: <https://www.technologyreview.com/2020/05/21/1002105/covid-bot-twitter-accounts-push-to-reopen-america/>
- Merry, Sally Engle (2016), *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking*, Chicago and London: University of Chicago Press.

Shelton, Taylor 2020, "A post-truth pandemic?", *Big Data & Society* 1 / 6, 1-10.

Taylor, Linnet (2020), "The price of certainty: How the politics of pandemic data demand an ethics of care", *Big Data & Society* 1 / 7, 1-10.

United Nations Department of Global Communications (2020), UN tackles 'infodemic' of misinformation and cybercrime in COVID-19 crisis. United Nations COVID-19 Response. 31 March. Available at: <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%98infodemic%E2%80%99-misinformation-andcybercrime-covid-19>

Van Dijck, José (2014), "Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology", *Surveillance & Society* 12 / 2, 197-208.