

# CATALOGO DE DATOS COVID-19: TRANSPARENCIA, INNOVACIÓN Y COMPROMISO SOCIAL

## **Danielle Mendes Thame Denny**

Profesora licenciada de la Fundación Armando Alvares Penteado. Investigadora visitante en el University College London. Doctora en Derecho por la Universidad Católica de Santos. Magíster en Comunicación por la Faculdade Cásper Líbero. Postgrado en Diplomacia Económica por la Universidad Estadual de Campinas. Licenciada en Derecho por la Pontificia Universidad Católica. Orcid: <https://orcid.org/0000-0002-8964-5205>

## **Emre Kazim**

Research Fellow en la University College London con Doctorado y Máster en Filosofía por la King's College London.

## **Adriano Koshiyama**

Research Fellow en la University College London. Doutor en Ciencias de la Computación pela University College London. Magíster en Ingeniería Eléctrica de la Pontificia Universidade Católica do Rio de Janeiro. Licenciado en Economía por la Universidade Federal Rural do Rio de Janeiro.

---

**Resumen:** Por cuenta de la pandemia las firmas tecnológicas y gobiernos están en conversaciones sobre la creación de muchas formas de coletas de datos, incluso los georeferenciados e de salud para hacer monitoreo de lockouts bloqueos, cuarentenas, aislamiento y distancia social. Los datos de salud, localización, compras, transferencias de dinero, biometría facial, certificados digitales pueden ser un arma poderosa para combate del virus. Sin embargo, si los datos colectados y parametrizados no fueren abiertos, con respecto a privacidad y protección de los datos personales o sensibles, si no haber justicia, legalidad y transparencia de cómo los algoritmos están produciendo sus análisis y hasta que punto las decisiones políticas y administrativas están siendo fundadas en las sugerencias tecnológicas, podemos estar caminando para una crisis urgente de derechos digitales. En este texto el caso brasileño será detallado más a fondo a título de ejemplo. El estudio de caso en algunos sitios electrónicos brasileños fue hecho en marzo a junio de 2020.

**Palabras clave:** Covid-19. Abertura de datos. Catalogo de datos. Privacidad. Protección de datos. Inteligencia artificial.

**Sumario:** Introducción – I Planeación, identificación de beneficiarios, y preparo de procesos – II Los conjuntos de datos y el caso brasileño – III Los datos abiertos en Brasil: ley de transparencia y de acceso a información – IV Protección de los datos – V Vigilancia – VI Algoritmos – Conclusión – Bibliografía

---

## Introducción

Por cuenta de la pandemia las firmas tecnológicas y gobiernos están en conversaciones sobre la creación de muchas formas de coletas de datos, incluso los georeferenciados e de salud para hacer monitoreo de lockouts bloqueos, cuarentenas, aislamiento y distancia social. Hay también iniciativas para emisión de pasaportes de salud para atestar que algunos desarrollaron inmunidad por que ya han tenido Covid-19 e por lo tanto pueden regresar de al trabajo mediante pruebas de coronavirus exámenes y reconocimiento facial (Devlin y Proctor, 2020).

Los datos de salud, localización, compras, transferencias de dinero, biometría facial, certificados digitales pueden si ser un arma poderosa para combate del virus. Sin embargo si los datos colectados y parametrizados no fueren abiertos y no haber transparencia de cómo los algoritmos están produciendo sus análisis y hasta que punto las decisiones están siendo fundadas en las sugerencias tecnológicas podemos estar caminando para una crisis urgente de derechos digitales (Reventlow, 2020). La “implementación de una nueva y extensa tecnología electrónica y centrada en datos que claramente ayudará a salvar vidas. A medida que se implementa esta tecnología, también debemos tener cuidado con lo que perdemos” (Claypoole, 2020).

Las restricciones de derechos humanos pueden estar se extendiendo casi tan rápido como el virus mismo. Estados están ayudados por empresas privadas que utilizan datos del teléfono, imágenes de CCTV, puntos de control de temperatura, reservas de aerolíneas y ferrocarriles, información de tarjetas de crédito, registros de compras en línea, datos de redes sociales, reconocimiento facial y, a veces, incluso drones (Reventlow, 2020). Pero algunos cuestionamientos son esenciales ¿Cuál son los datos en uso? ¿Como están siendo colectados? ¿Qué tan seguros son nuestros datos? ¿Cuánto tiempo se guardará? ¿Hay transparencia sobre cómo se obtiene y procesa? ¿Se comparte o se reutiliza, y si es así, con quién? ¿Cuál son las formas de responsabilizar por el mal uso?

El objetivo de este trabajo es demostrar que en todos estos puntos los datos abiertos y la transparencia son fundamentales. Los gobiernos no pueden utilizar una emergencia sanitaria mundial para instaurar medidas de infracción de los derechos humanos sin ser auditados, cuestionados o eventualmente tener que revertir estas medidas y sin la abertura de los datos no es posible hacer verificaciones o responsabilizar los malos usos. También un catalogo de datos abiertos puede servir a la sociedad para que se pueda hacer análisis e la creación de aplicativos y por esto contribuí con la creatividad y con el desenvolvimiento de soluciones mejores para el problema colectivo.

La metodología elegida fue el análisis descriptivo normativo con estudio de caso en sitios web gubernamentales o relacionados con el tema de la pandemia durante el mes de marzo a junio de 2020. El primer parte trae los beneficios de se hacer planeación, identificación de beneficiarios, y preparo de procesos y no adoptar apresuradamente medidas que deban ser más estudiadas y debatidas para paliar los posibles daños. La segunda parte es sobre la gobernación de datos y trae ejemplos de iniciativas brasileñas de colecta, tratamiento e divulgación de datos relacionados con la pandemia. En el tercer parte el tópico son los datos abiertos en Brasil, y como las leyes de transparencia y de acceso a información regularlos. La cuarta parte completa el análisis legislativo con detalles de la Ley General de Protección de Datos Brasileña. En la quinta parte el tema es la posibilidad de vigilancia que las herramientas de control ponen en marcha por la pandemia sopesando la necesidad de prever su retiro luego de la situación de emergencia cesar y que durante su aplicación sean administradas de manera democrática. La ultima parte trae consideraciones a respecto de los algoritmos debieren ser verificados, certificados y regulados, principalmente en la solidez, la equidad y la capacidad de explicación del sistema. Hay la necesidad de transparencia y datos abiertos contribuyen para tanto.

## **I Planeación, identificación de beneficiarios, y preparo de procesos**

En medio a emergencia sanitaria la coleta y uso de los datos están siendo hechas sin planeación, con avances y retrocesos. Hay risco de se dejar iludir por una “herejía tecno-optimista” (Canales, 2020) que no da comta dela complejidad sistémica del problema pero nos ofrece aislar solo partes del problema y que nos ofrece, incluso, despolitizar la discusión, sin sopesar la combinación de pandemia con economías precarias, pobreza, hambre, y servicios de salud en los cuales se ha invertido escasamente, estos son los ingredientes para una “tormenta perfecta” (Canales, 2020. pág. 5).

(...) propongo clasificar los distintos roles que se le ha asignado a la tecnología en el en el marco del control de la pandemia, de acuerdo al grado de intrusividad en el ejercicio de los derechos de las personas, ordenadas de menor a mayor: Información de salud; Autodiagnóstico; Datos integrados para toma de decisiones de salud pública; Trazabilidad de contactos; Pasaportes de movilidad y trabajo; Vigilancia de confinamiento. (Canales, 2020. pág. 6)

Es necesaria una estrategia de apertura de estos datos que están siendo colectados, compartidos y alimentando las herramientas de inteligencia artificial. Solo con un planeamiento y una estrategia clara y mucho transparente se puede tener en cuenta las consideraciones necesarias y organizar las acciones requeridas para lograr los beneficios sociales de combate à Covid-19.

Data Privacy Brasil en su reporte “Privacidad y pandemia - recomendaciones para el uso legítimo de los datos en la lucha contra el COVID-19” elaboró un conjunto de directrices. Primer hay que se valuar la necesidad de desarrollar políticas de salud centradas en datos; después la definición clara del propósito y de la necesidad del procesamiento de datos; para con esto definir como y hasta cuando serán procesados y también cuando los datos pueden y deben ser eliminados; paralelamente hay de haber una definición jurídica clara en el ordenamiento con salvaguardias específicas para los derechos fundamentales envueltos e el proceso de trato de los datos; esto todo con garantía de publicidad, transparencia y participación (Bioni y otros, 2020. pág 6).

Con el uso de la información que recolectan gobiernos, instituciones públicas, empresas privadas de telefonía, tecnologías y plataformas sociales, el ideal sería construir un catalogo de datos conteniendo estas informaciones de fuentes diversas e que pueden ser utilizadas por la sociedad a ejemplo del que hay en la provincia de Buenos Aires (PBA, 2020) no especifico para la pandemia mas que articula un conjunto variado de datos de diversas fuentes.

El objetivo de se crear un catalogo de datos es que el estoque masivo de datos pase de ser propiedad privada de una institución o de ser detenida por un órgano gubernamental y pase a ser información pública y de libre uso, que cumpla con las características técnicas adecuadas y sea de interés y utilidad para los usuarios. Menos susceptible entonces a las voluntades particulares de un grupo o de determinado gobierno.

Pero antes de la creación del catalogo es necesario conocer los intereses y necesidades que tienen los distintos usuarios potenciales de los datos, para establecer los parámetros de la demanda de información en términos de la cantidad de información, de cualidad do que se puede, y las ofertas temáticas que se pode satisfacer con a formación de esto catalogo de datos. Son varias las instituciones que trabajan con datos y el desafío mayor es voluntad política para envidar los esfuerzos necesarios para compilar una significativa mayoría.

El punto de partida sería identificar los beneficiarios con encuestas que solicita la participación de todos aquellos interesados directa o indirectamente en el catalogo de datos. En ellas se conocerían los intereses en relación con la información con preguntas de opción múltiple y algunas de respuestas abiertas. También el uso de indicadores automatizados indicando cuáles son los temas más

solicitados en un determinado momento que pueden ser el objetivo intermediario del proyecto para suplir esta demanda de forma efectiva, transparente e con datos abiertos al público.

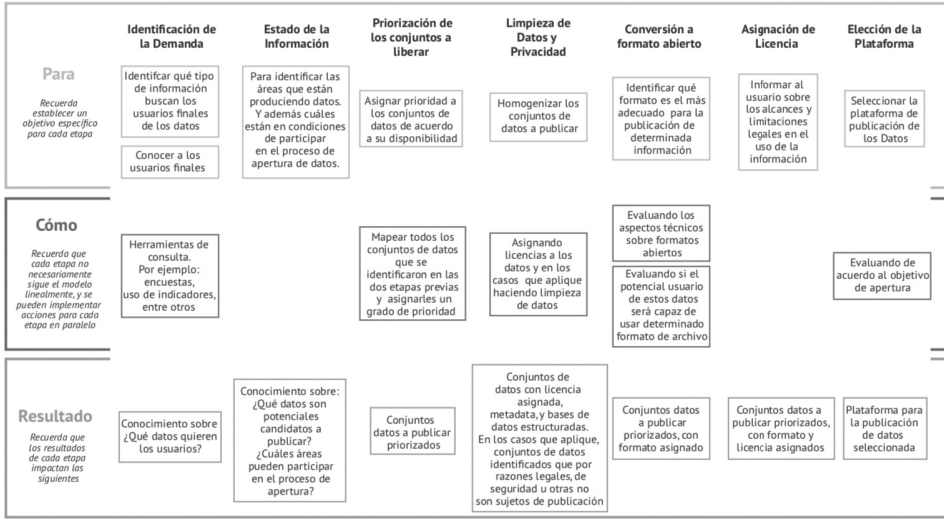
Para el desarrollo de la propuesta del catálogo de datos abiertos hay que seguir para mas a frente del proceso de identificación de demanda por los datos, para cual es el estado de la información, priorización de algunos conjuntos de datos frente a otros, limpieza e parametrización de los datos y incluso muy importante para el caso de covid-19 es la anonimización para que no se perjudique la privacidad de los titulares de los datos y no cause responsabilidad por mal uso de datos personales o sensibles a los controladores de los datos.

También puede ser necesario esfuerzos de conversión de lo que hay e ya esta siendo colectado en un formato que pueda ser divulgado de forma abierta con el determinada asignación de licencia de uso en una plataforma de publicación de datos que haga la divulgación facilitando el uso por la sociedad.

El modelo código abierto (*open source*) es o más útil para desarrollar programas de forma colaborativa siendo importante en la co-creación de tecnologías. “Colaboración, libertad y desarrollo continuo son algunas de las principales ideas” (IRIS, 2020). Deriva de los ideales de otro movimiento anterior –el software libre–, que predicó “la libertad de ejecutar, estudiar y cambiar, y redistribuir copias con o sin cambios” de un programa (IRIS, 2020).

La directriz código abierto acepta la existencia de software propietario y no siempre son gratuitos. Además de mantener abierto el código, hay algunos parámetros: distribución gratuita, código fuente, obras derivadas, integridad del autor del código fuente, no discriminación de personas y grupos, no discriminación por campos de actuación, distribución de licencia, la licencia no debe ser específica de un producto, la licencia no debe restringir otro software, y la licencia debe ser tecnológicamente neutra (Fok y Picon, 2016). Pero en las etapas de apertura de datos pode se percibir que no siempre es posible adoptar la modalidad de código abierto.

**Figura 1 – Etapas proceso apertura de datos**



Fuente: OEA, 2016

También es de fundamental importancia conocer el ecosistema de datos que estarán involucrados en el catalogo de datos de Covid-19 esto facilita identificar la demanda, estimar el potencial de uso de los datos y planear estrategias de apertura acordes a esta realidad. Identificar a los usuarios de datos por sector (ej. distintos órganos del gobierno, hospitales públicos o privados, farmacéuticas, empresas de telefonía, redes sociales, periodistas, sociedad civil, científicos de datos, academia, etc.) es útil para fines de categorización y de interacción con dichas comunidades ya que permite reconocer las actividades específicas que realizan y también se puede identificar el interés que tienen detrás de los datos.

**Figura 2 – Ecosistema de datos**



Fuente: OEA, 2016

Conforme se verifica en el diseño, integran el ecosistema de datos productores que crían los datos y los publican; intermediarios que almacenan y transforman; habilitadores que capacitan para que los usuarios de los datos puedan conocer las herramientas e las posibilidades de usabilidad de los datos; infomediarios que transforman los datos en información consumible para los usuarios finales y los consumidores que leen, se informan e contribuyen para la comunicación de estos datos.

## II Los conjuntos de datos y el caso brasileño

Desde antes de la pandemia está en marcha una revolución de datos (Kitchin, 2014), remodelando la forma en que se produce el conocimiento, se llevan a cabo los negocios y se promulga la gobernanza, se hace vigilancia, se protege la privacidad y seguridad. Los datos fluyen torrencialmente, a costos progresivamente bajos y dependen de infraestructuras cada vez más confiables y accesibles. Esta datificación trae nuevas vulnerabilidades que demandan estrategias de mitigación, en un conjunto amplio de intervenciones sistémicas (Kitchin y Dodge, 2019).

Son varios los conjuntos de datos actualmente siendo colectados, pero aun esparzo, de diversas fuentes y non uniformes. Y para el proyecto de apertura de datos pertinentes a el control de la pandemia hay que hacer una distinción

importante entre los datos que serán compartidos y los datos personales y sensibles, porque muchos datos de salud encontrasen en estas categorías.

En Brasil la ley de protección de datos personales está con sus sanciones postergadas hasta 2021 pero esto no quiere decir que los datos personales o sensibles podrían ser compartidos libremente. La ley general de protección de datos – Lei 13.709/2018 (Brasil, 2018) brasileña exige la anonimización para que sean compartidos. En Brasil los datos del sector público deben ser los primeros en ser procesados y colocados en el dominio público en el catalogo de datos del coronavirus, ya que la ley de transparencia - LC 131/2009 (Brasil, 2009) y la ley de acceso a la información - Lei 12527/2011 (C. N. Brasil, 2011) ya exigen que este sea el caso en Brasil.

Algunos datos públicos ya estaban siendo divulgados, pero el gobierno retrocedió. Un ejemplo es el Portal COVID-19 (SVS/MS, 2020) en cual sitio electrónico, diariamente, el Ministerio de Salud, a través de la Secretaría de Vigilancia Sanitaria (SVS/MS) publica datos consolidados sobre COVID-19 sobre la situación epidemiológica de COVID-19 en Brasil. El proceso de actualización de datos sobre casos y muertes confirmados se lleva a cabo a través de información oficial proporcionada por las Secretarías de Salud de los Estados (hay 27 Unidades Federativas estatales en Brasil).

En el modelo sin censura de datos ya había limitaciones con relación con los datos municipales porque el proceso de actualización de información en municipios de forma consolidada en cada estado es dinámico y complejo además teniendo en cuenta la pluralidad de cada municipio brasileño con respecto al tamaño de la población, la infraestructura y la organización de los servicios de salud, además de todos los desafíos que impone la pandemia, falta de control, errores o demoras en la transferencia de información eran problemas esperados.

Además de los datos sobre casos y muertes confirmados el mismo sitio electrónico informa el número de casos nuevos confirmados que fueron registrados por los departamentos de salud municipales y estatales en relación con el día anterior; los casos acumulados (número total de casos confirmados); las nuevas muertes que fueron registradas por los departamentos de salud municipales y estatales en relación con el día anterior; las muertes consolidadas; y el coeficiente de incidencia, que es el número de casos confirmados de COVID-19 por 100.000 habitantes, en la población que reside en un espacio geográfico determinado, en el período considerado.

Para el análisis de casos y muertes por fecha de ocurrencia, se debe cambiar el sitio electrónico y utilizar los datos registrados en el Sistema de Vigilancia Epidemiológica de la Influenza (SIVEP-Gripe, sigla en portugués), donde los profesionales de salud deben informar todos los casos de Síndrome Respiratorio



Agudo Severo Hospitalizado (SRAG) o muertes debidas a SRAG, independiente de la hospitalización.

En termos de costos el Portal de Transparencia del Departamento Federal de Asuntos Internos, muestra cuanto hay destinado a ayuda de los estados y municipios. En 25 de marzo de 2020, se había comprometido aproximadamente R\$2.01 mil millones, con énfasis en las transferencias de R\$1,21 mil millones a fondos municipales de salud y R\$424 millones a fondos estatales, este último con fondos liquidados y pagados (C. G. da U. Brasil, 2020). El sitio específico de este tema genera datos solo día a día y no es posible descargar el archivo en ningún formato, solo ver los compromisos y pagos en línea a la junta ejecutiva del fondo nacional de salud para transferirlos a fondos municipales. No hay API especifica para busca en los datos relacionados con la pandemia.

**Figura 3 – Ejemplo del portal de transparencia brasileño**

The screenshot displays the 'Portal da Transparência' interface. At the top, there is a search bar with the text 'Busque por órgão, cidade, CNPJ, servidor...'. Below the search bar is a navigation menu with items like 'Sobre o Portal', 'Painéis', 'Consultas Detalhadas', 'Controle social', 'Rede de Transparência', 'Receba Notificações', and 'Aprenda mais'. The main heading is 'Documentos de execução da despesa pública' with a button 'ORIGEM DOS DADOS'. On the left, there is a 'FILTRO' sidebar with categories: PERÍODO, NOME FAVORECIDO, CPF/CNPJ FAVORECIDO, DOCUMENTO, ÓRGÃO, UNIDADE ORÇAMENTÁRIA, and Localidade. The main area shows 'FILTROS APLICADOS:' with the following selected filters: Período de: 01/03/2020, Período até: 31/03/2020, Órgão: DIRETORIA EXECUTIVA DO FUNDO NAC. DE SAUDE (Unidade Gestora), Subordinado(a): Ministério da Saúde - Unidades com vínculo direto, Órgão superior: Ministério da Saúde, Função: 10 - Saúde, Subfunção: 122 - Administração geral, Programa Orçamentário: 5018 - ATENCAO ESPECIALIZADA A SAUDE, Ação Orçamentária: 21CO - ENFRENTAMENTO DA EMERGENCIA DE SAUDE PUBLICA DE IMPORTANCIA INTERNACIONAL DECORRENTE DO CORONAVIRUS, Grupo de Despesa: 3 - Outras Despesas Correntes, Elemento de Despesa: 41 - Contribuições, Modalidade de Aplicação: 41 - TRANSFERENCIAS A MUNICIPIOS - FUNDO A FUNDO. A 'LIMPAR' button is at the bottom right. At the bottom left, it says 'Dados atualizados até: 09/06/2020'.

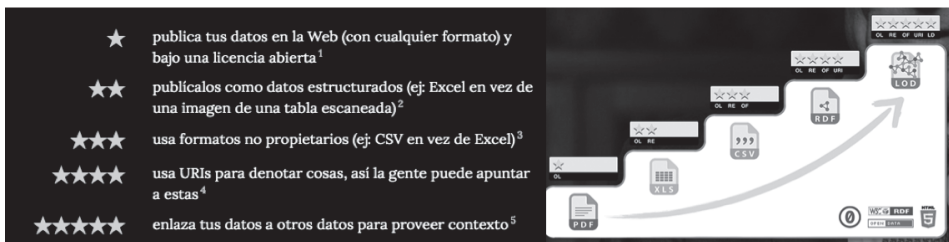
Fuente: Portal da transparencia (C. G. da U. Brasil, 2020).

Los APIs disponibles son para otros tipos de recorte: acuerdos de *leniência*, datos del beca familia con cantidad de beneficiarios por municipalidad, programa de erradicación de trabajo infantil, garantía zafra, seguro defeso, catastro de expulsos da administración federal, catastro de empresas inidóneas o suspensas, catastro de empresas municipales, contratos con el ejecutivo federal, convenios

con el poder ejecutivo federal, gastos públicos, cartones de pago, entidades privadas sien fines lucrativos impedidas, licitaciones del poder ejecutivo federal, servidores do poder ejecutivo federal, viajes a servicio, agencias catastradas en Sicafi y Siape, sistemas de administración integrada das finanzas y de personal respectivamente. Sin embargo, los datos completos se pueden descargar en otro enlace en formato CSV de acuerdo con temas más generales (sin recortar a la pandemia) en el portal brasileño de datos abiertos (S. de T. da I. Brasil Ministério do Planejamento, Desenvolvimento e Gestão, 2020).

El análisis de estos dos enlaces plantea otro desafío: el volumen, los detalles de la nomenclatura, la complejidad de los datos y, a veces, los campos que se dejan sin información o se rellenan solo con números en las tablas lo hacen incluso cuando hay una clara intención de hacer que los datos se abran con las 5 estrellas de Lee, será necesario que el usuario tenga muchos conocimientos técnicos o aplique inteligencia artificial para procesar este volumen de datos de una manera que los haga sentido (Lee, 2012).

**Figura 4 – 5 estrellas por Lee**



Fuente: 5stardata (Lee, 2012)

Relacionados con los contratos, debido a la pandemia, se promulgó una ley en Brasil (L13.979/2020) que exige a la administración pública de seguir los procedimientos normales de contratación (licitación) para comprar bienes, servicios y suministros de salud. Según esta ley, la exención de licitación es temporal y se aplica solo mientras dure la emergencia de salud pública y todos los contratos o adquisiciones realizados deben estar disponibles de inmediato en Internet y ten sido hecho en uno sitio web de lo Ministerio de la Salud (M. da S. Brasil, 2020). En él, la información se describe para poder copiarla, pero no hay posibilidad de descargarla en ningún formato, por lo que no cumple con todos los niveles de apertura, siendo no máximo dos estrellas se no una (Lee, 2012).

Otra iniciativa es la aplicación desarrollada por el Sistema Único de Salud, llamado Coronavirus – SUS (S. Ú. de S. Brasil, 2020) tiene como objetivo sensibilizar a la población sobre el virus, brindando información sobre diversos temas, como síntomas, cómo prevenirlo, qué hacer en caso de sospecha e

infección, un mapa que indique las unidades de salud cercanas, etc. Sin embargo, los datos recopilados no están disponibles en transparencia activa. Probablemente serán informados a petición de los usuarios que se realizarán en el sitio (E. F. Brasil, 2011) para pedidos de aseso a las informaciones gubernamentales en atendimento a ley (C. N. Brasil, 2011).

Otra iniciativa del Ministerio de Salud es el sitio electrónico para luchar contra las noticias falsas (Brasil, 2019), listando los resultados de analices hechas por el Ministerio tiendo puesto a disposición un número de WhatsApp para la población enviar mensajes con sospechas de ser falsas, para que el personal técnico del Ministerio los revise y responda oficialmente si son verdaderos o falsos. Es importante tener en cuenta que WhatsApp es la principal forma de acceso de la población brasileña al contenido en línea, ya que los operadores telefónicos no aplican tarifa por su uso (el uso de la aplicación no consume el paquete de datos del usuario) y el teléfonos móviles son ampliamente diseminados en la población brasileña, 93% de los domicilios brasileños tiene los (Cetic.br, 2019).

Por lo tanto, en este sitio web, aparecen mensajes con imágenes o textos que se publicaron en las redes sociales y se enviaron al ministerio para confirmar que la información es correcta. Se analizan varios casos, todos aparecen en el mismo formato y los usuarios pueden consultarlos caso por caso. Sin embargo, los metadatos relacionados con estos análisis no están disponibles, ni es posible descargar archivos en ninguna forma

Además, de estas iniciativas publicas hay datos que están siendo recopilados por convenios de periodistas, colectivos y universidades. Por ejemplo, después del “apagón” de la información del gobierno federal brasileño, los vehículos de comunicación concurrentes (O Estado de S. Paulo, Folha de S.Paulo, O Globo, G1 e UOL) se unieron para informar datos sobre la pandemia.

En el sitio Brasil.io (Justen, 2020) un grupo de trabajo de 40 voluntarios compilan diariamente boletines epidemiológicos de los 27 departamentos de salud estatales, proporcionamos una base de datos con serie histórica de casos confirmados y muertes por municipio, proporcionando también datos útiles relacionados con la pandemia, como la población de los municipios afectados, presuntas muertes registradas.

También una estrategia mucho útil para el detalles de las políticas de enfrentamiento es a de la territorización de la pandemia (LabCidade, 2020b) para saber al cierto donde los hospitales están mas sobrecargados, porque la región metropolitana de São Paulo, que es el recorte metodológico aplicado, no es homogénea, hay localidades con muchos casos y otras con bien menos. Los gráficos interactivos fueron hechos con base en datos del Ministerio de la Salud (D. Brasil, 2020) pero esto pasó a censurar los códigos postales que permitían la elaboración de la analice localizada do los impactos de la pandemia y la

difusión espacial de Covid-19 (identificación del código postal de residencia de las personas hospitalizadas con Síndrome Respiratorio Agudo Severo (SRAS), que incluí Covid-19).

También hay agentes privados compartiendo datos con el gobierno como el para monitorear los desplazamientos hechos por las plataformas digitales y podrían/deberían ser abiertos para que toda la sociedad pueda aprovechar estos beneficios de analizar tales datos y para minimizar los riesgos de exacerbar el control de los individuos por parte de los gobernadores o de uso de datos de usuario en beneficio de empresas privadas.

Son diversos los ejemplos que están siendo ya producidos y pueden ser consolidados para complementar o conjunto de datos públicos que deban ser obligatoriamente colocados en dominio público de forma consolidada como un ejemplo a ser seguido de la Rede InfoAmazonia (Mãe d'Água) entre tecnólogos, periodistas e ambientalistas para producir datos abiertos sobre la Amazonia Legal (una región bien mayor que la floresta). Un de los estudios que esta red produce es un mapa interactivo con los números de respiradores por municipio, informando cuantos hay e cuantos están en uso (InfoAmazonia, 2020).

### **III Los datos abiertos en Brasil: ley de transparencia y de acceso a información**

En Brasil, la Unión, los Estados, el Distrito Federal y los municipios deben observar la ley de transparencia – LC 131/2009 (Brasil, 2009) y la ley de acceso a la información – Lei 12527/2011 (LAI) (C. N. Brasil, 2011), que incluyen los poderes ejecutivo, legislativo, tribunales de cuentas, el poder judicial, el ministerio público; autarquías, fundaciones públicas, empresas públicas, empresas de capital mixto y entidades controladas directa o indirectamente por las entidades de la federación; y entidades privadas sin fines de lucro que reciben fondos públicos.

La ley de transparencia fue creada para divulgar los ingresos y gastos de todas las entidades públicas en tiempo real, con un período máximo de 24 horas, en algún sitio electrónico, siguiendo algunos requisitos técnicos que validarán que esta información sea correcta bajo las inspecciones del Ministerio Público.

En relación con los ingresos, deben publicarse los valores de todos los ingresos de la unidad de gestión, incluida al menos su naturaleza, en relación con el pronóstico; lanzamiento; forma de colección, incluida la referencia a cualesquiera recursos extraordinarios.

De los gastos debe estar disponible: el monto del compromiso, liquidación y pago; el número del proceso de ejecución correspondiente, cuando corresponda; la

clasificación del presupuesto, especificando la unidad presupuestaria, la función, la subfunción, la naturaleza del gasto y la fuente de los fondos que financiaron el gasto; la persona física o jurídica que recibe el pago, incluidos los desembolsos para operaciones independientes, excepto en el caso de la nómina de personal y los beneficios de la seguridad social; el procedimiento de licitación realizado, así como su desestimación o imposibilidad de ejecución, cuando corresponda, con el número de proceso correspondiente; y el bien o servicio prestado.

La LAI regula los procedimientos para el derecho a la información garantizados por la Constitución Federal, requiriendo que la publicidad sea la regla y la excepción del secreto, pero no se exige que la información esté disponible de forma inmediata en forma de datos abiertos.

En la Constitución brasileña, las regulaciones con respecto a los datos abiertos son las siguientes:

Artículo 5º - XXXIII - toda persona tiene derecho a recibir de los organismos públicos información de su interés particular, o de interés colectivo o general, que se proporcionará dentro del plazo de la ley, bajo pena de responsabilidad, excepto aquellos cuyo secreto es esencial para la seguridad de la sociedad y del Estado. (CF, 1988)

Artículo 37 - La administración pública directa e indirecta de cualquiera de los Poderes de la Unión, los Estados, el Distrito Federal y los Municipios deberá obedecer los principios de legalidad, impersonalidad, moralidad, publicidad y eficiencia (...). §3º - La ley regulará las formas de participación de los usuarios en la administración pública directa e indirecta, regulando en particular: II - el acceso de los usuarios a registros administrativos e información sobre actos gubernamentales. (CF, 1988)

Artículo 216 - §2º La administración pública, bajo los términos de la ley, es responsable de la gestión de la documentación gubernamental y de las medidas para abrir su consulta a todos los que la necesiten. (CF, 1988) (Traducción realizada por los autores)

La información de interés público y colectivo relevante producida o mantenida por entidades públicas debe publicarse independientemente de los requisitos de forma inmediata en forma de datos abiertos. En otras palabras, los sitios web de los organismos sujetos a LAI deben proporcionar una lista de información espontáneamente, sin tener que hacer una solicitud de información. Son ellas:

Registro de contactos, un organigrama (“quién es quién”) y habilidades; Respuestas a preguntas frecuentes de la sociedad; Registro de gastos y transacciones financieras; Información sobre ubicaciones de servicio y servicios

disponibles; Información sobre ofertas, procedimientos de licitación, contratos y aditivos; datos generales para el seguimiento de políticas y obras públicas; La lista de documentos que ya no son confidenciales y la lista de documentos que son confidenciales, con una indicación del tema, grado de confidencialidad, autoridad y fecha de clasificación; Informe estadístico periódico sobre las solicitudes de información recibidas, servidas y rechazadas; Motor de búsqueda y enlace al portal de transparencia; Los datos deben estar en un formato abierto (Logarezzi, 2016)

Para otros asuntos, no hay obligación de que la información esté disponible inmediatamente en forma de datos abiertos. Sin embargo, si no hay acceso inmediato en línea, aún es obligatorio que los ciudadanos puedan enviar solicitudes de información en línea. Las agencias gubernamentales tienen hasta 20 días para responder a estas solicitudes, el período se puede extender por otros 10 días, previa justificación. La extensión de la solicitud debe hacerse y justificarse antes de que finalice el período inicial de 20 días.

## **IV Protección de los datos**

El Brasil ya promulgó su Ley General de Protección de Datos – Lei nº 13.709/2018 (Brasil, 2018) pero sanciones comenzarán a aplicarse solo en agosto de 2021. Su disposición está en gran medida alineada con la Ley General de Protección de Datos de la UE (E. P. and C. EU, 2016). Como en la ley europea, se hace protección de datos no solamente de la privacidad. En la Carta de los Derechos Fundamentales de la UE la protección de datos es un derecho fundamental (artículo 8) que se establece de manera distinta del derecho fundamental al respeto de la vida privada y familiar (artículo 7) (EU, 2016).

Y esto hace una gran diferencia, cuando gobiernos, empresas y autoridades de salud pública están considerando la recopilación de datos personales y el seguimiento digital de las personas como una forma efectiva de detener la propagación de la pandemia. El derecho a la protección de datos personales ha sido concebido particularmente para ayudar a las sociedades a enfrentar la realidad de la automatización masiva de sistemas alimentados con datos sobre individuos. Tiene naturaleza procedimental regulando las condiciones para recopilar y usar datos personales, con necesidades de justificación, minimización, transparencia, equidad, control de los interesados, concesión de derecho de acceso, corrección y eliminación, creación una autoridad supervisora, y de responsabilidad de quien sea responsable de la recopilación y uso de datos personales.

Es importante destacar que el objetivo de la protección de datos es garantizar que la información relacionada con las personas se recopile y utilice de tal manera que todos sus otros derechos fundamentales estén protegidos. Esto incluye la

libertad de expresión, el derecho a la vida privada / privacidad, el derecho a la vida, el derecho a la seguridad, el derecho a la no discriminación, etc. E por eso que la protección de datos está equipada para permitir combatir la pandemia utilizando datos personales.

Son cinco los ejes principales: i) unidad y aplicación general de la ley brasileña; ii) legitimación para el tratamiento de datos (autorizando hipótesis); iii) principios y derechos del titular; iv) obligaciones de los agentes de procesamiento de datos; v) responsabilidad de los agentes (Mendes y Doneda, 2019). Además, la Autoridad Nacional de Protección de Datos es un pilar fundamental de la ley porque tiene funciones relevantes, como supervisar el procesamiento de datos y sancionar el incumplimiento de la legislación, regular hipótesis no especificadas en la legislación, orientar a la sociedad sobre la aplicación de la ley y recibir denuncias sobre violaciones de las normas de protección de datos (Mendes y Doneda, 2019).

Los imperativos de protección de los datos personales o sensibles se aplican a cualquier respuesta tecnológica o digital a la pandemia que se base en datos personales: el flujo de datos debe ser claro, incluidas todas las categorías de datos que se recopilarán y utilizarán, los propósitos deben ser claros, específicos, granulares, bien definidos, la red de responsabilidad debe ser clara (¿quiénes son los controladores y los procesadores?), a los datos personales no se deben compartirse ni tener acceso a ellos, más allá de la red de responsabilidad definida (por ejemplo, a través de acuerdos controlador-procesador).

Debe haber transparencia de manera inteligible para las personas (titulares de los datos) sobre cuyos datos personales se colectan y recopilan y como, atendiendo a la minimización (¿puede el proyecto prescindir de algunos de ellos y lograr el mismo propósito?). Todos los datos personales deben ser precisos, se no, el titular puede solicitar su corrección, o que sea borrado mediante justificado. Los datos hay que estar seguros y conservarse solo durante el tiempo que sea necesario para lograr el propósito (luego, deben eliminarse después del uso o hacer anonimización).

El artículo 7 de la ley brasileña autoriza expresamente el tratamiento de datos personales (información relacionada con una persona física identificada o identificable) para la protección de la salud, por autoridades de salud o sanitaria. Y el artículo 11 autoriza el uso de los datos sensibles (como los datos relacionados con la salud o biométricos, cuando están vinculados a una persona natural) independientemente de autorización del titular en casos donde sea indispensable para la protección de la salud, por autoridades de salud o sanitaria.

Pero esto no es tan simples puede haber discusión sobre la indispensabilidad de los datos y también sobre la capacidad técnica de anonimización (Veale, Binns y

Edwards, 2018) o sobre cuales serian los medios técnicos razonables disponibles para tanto en el momento de tratamiento de los datos.

Un ejemplo de esta discusión es el uso del CEP (código postal brasileño) nos dados disponibles en el sitio del Ministerio de la Salud. Esta información fuera suprimida para algunos esta decisión fuera acertada en virtud del CEP poder servir para identificar personas que moran en cayes pequeñas (Moura, 2020) pero para otros tratase de mas una tentativa de obstaculizar el monitoreo por parte de activistas, gerentes e investigadores, y con eso confundir a la población y minimizar la gravedad de la pandemia (LabCidade, 2020a). Algunas herramientas respectando a privacidad están ahora disponibles para proteger entre otros aspectos esta localización (Ferraz y Moura, 2020).

Es importante señalar que además de la Ley General de Protección de Datos en la legislación brasileña, existen normas relacionadas con la apertura de datos para investigación científica, pero aún no existe una regulación vigente que regule específicamente el acceso, tratamiento y uso de datos para investigación en salud y hay otras cuestiones además de la protección de datos que también están vinculadas a la apertura de datos para la investigación en salud, como la propiedad intelectual, la seguridad de la información, la reutilización y el intercambio de datos. Gran parte de la regulación original se basa también en directrices, códigos de conducta y límites derivados de la autorregulación (Souza y otros, 2018).

Esta naturaleza no consensual y dinámica de la protección de datos faz reforzar aún más la necesidad de articulación de un catalogo de datos con gobernanza transparente e *multistakeholder*, colaborando con la necesaria interoperabilidad de los conjuntos de datos que contienen informaciones críticas para enfrentar los desafíos de salud pública. El catalogo facilita también la conducción evaluación de impacto de protección de datos cuando es probable que surja un alto riesgo para los derechos del individuo desde la naturaleza, el alcance, el contexto o los propósitos del procesamiento de datos.

## V Vigilancia

La vigilancia tecnológica hace parte del dia-a-dia, y hay errores regulatorios que impiden la aplicación de las debidas limitaciones a lo “capitalismo de vigilancia” (Zuboff, 2019). Y esto no es por acaso, en sus interacciones con los órganos regulatorios “os capitalistas de la vigilancia han fomentado modos de interacción regulatoria que marginan el controle general en favor de la autorregulación y la autocertificación” (Cohen, 2019).



Particularmente preocupante es el uso de tecnologías de vigilancia bajo estatutos de emergencia que apelan a las lógicas de guerra, haciendo una “securitización de la salud” (Castro, 2017; Castro y Santos, 2020).

(...) la CIDH ha observado que se han suspendido y restringido algunos derechos, y en otros casos se han declarado ‘estados de emergencia’, ‘estados de excepción’, ‘estados de catástrofe por calamidad pública’, o ‘emergencia sanitaria’, a través de decretos presidenciales y normativa de diversa naturaleza jurídica con el fin de proteger la salud pública y evitar el incremento de contagios. Asimismo, se han establecido medidas de distinta naturaleza que restringen los derechos de la libertad de expresión, el derecho de acceso a la información pública, la libertad personal, la inviolabilidad del domicilio, el derecho a la propiedad privada; y se ha recurrido al uso de tecnología de vigilancia para rastrear la propagación del coronavirus, y al almacenamiento de datos de forma masiva. (OEA, 2020)

Tal securitización se da trasladando los problemas de salud del ámbito de la prevención o cura de enfermedades para al ámbito de la seguridad, lo que justifica la adopción de medidas extraordinarias y urgentes que en situaciones normales no sería aceptado por la población, incluyendo relativización de los derechos fundamentales, excepcionalidad en los balances y controles de aquellos en el poder público, y sin necesidad de rendición de cuentas, transparencia y supervisión exigibles en otro contexto.

La tecnología de vigilancia permite decidir quién está autorizado a participar de la vida pública, quién puede trabajar, quién debe quedarse en casa, quién recibe la ayuda económica que le permita sobrevivir y quién no. Nada de eso es propio de una sociedad movida por principios democráticos ni el respeto de los derechos humanos. Más bien, parece un eco de las mejores novelas de ficción que vaticinaban un futuro de autoritarismo y control en nombre del bien público. (Canales, 2020. pág 23)

Para mitigar estos efectos orwellianos (1949) es importante siempre considerar el dicotomía en la adopción de opciones tecnológicas que pueden ser tanto usadas para el bien común como traer aun otros problemas quica peores. Así es con el uso de datos de teléfonos móviles en los esfuerzos analíticos para controlar la pandemia sin puede ofrecer contribución a varias áreas amplias de

investigación, pero hay de haber un modelo de divulgación de datos que favorezca los beneficios e mitigue los danos.

Son varios los beneficios primero la conciencia situacional se beneficiaría de un mayor acceso a estimaciones de población e información de movilidad previamente no disponibles para permitir a los interesados en todos los sectores mejor comprender las tendencias de contaminación y la distribución geográfica. Segundo los casos de uso de causa y efecto pueden ayudar a las partes interesadas a identificar los factores clave y las consecuencias de implementar diferentes medidas para contener la propagación del virus. Su objetivo es establecer qué variables marcan la diferencia. Tercero las tareas de predicción aprovecharían los recuentos de población en tiempo real y los datos de movilidad para permitir nuevas capacidades predictivas y permitir a las partes interesadas evaluar riesgos, necesidades y oportunidades futuras. Cuarto la evaluación de impacto tiene como objetivo determinar cuál y cómo varias intervenciones afectan la propagación del virus y requiere datos para identificar los obstáculos que impiden el logro de ciertos objetivos o el éxito de intervenciones particulares (Oliver y otros, 2020).

Para que los beneficios sopesen los riesgos el hecho de que estos datos estén disponibles abiertamente después de ser tratados por anonimato puede generar un aumento significativo en la calidad y cantidad de posibles análisis. Por ejemplo, los datos abiertos puestos a disposición por el gobierno federal como resultado de la ley de transparencia ya permiten varias otras analices. Otros ejemplos son la infografía e reportos hechos por periodistas, académicos y otros infomediarios de datos con los datos puestos a disposición. El trabajo sería aún más completo si más datos estuviesen disponibles.

Construir un catalogo de datos abiertos relacionados con Covid-19 tiene el objetivo avanzar en el acceso y uso de información de alta calidad, que forman la base del trabajo de analice científica e tomada de decisiones fundamentadas para la formulación de políticas basadas en evidencia. Y también contribuir para construcción de una cultura organizacional de transparencia y un entorno político de apoyo para que las organizaciones y la sociedad proporcionen los datos, las capacidades para satisfacer esta demanda y usos creativos para estos datos.

A medida que los esfuerzos anteriores han hecho que los datos gubernamentales sean más accesibles y abiertos, articular el valor económico y práctico de esa información para el gobierno y la que es producida pelo sector privado y compartida con el gobierno es un elemento esencial del éxito futuro de políticas eficaces de controle de la pandemia sin comprometer los derechos de los individuos, incluso los derechos humanos de privacidad y libertad de expresión. La investigación, producción del catalogo y el proceso de abertura permite a los responsables políticos, a los ciudadanos y a las empresas privadas comprender

mejor “por qué los datos son importantes” para la sociedad se puede aplicar para comprender las políticas y opciones actuales para futuras reformas (Data Foundation, 2020).

La creación de la capacidad técnica para mejorar la calidad de los datos, la dinámica cultural y la comprensión de la propuesta de valor para el uso de datos solo pueden tener éxito en el cumplimiento de los objetivos establecidos si las capacidades técnicas existen para permitir que la sociedad utilice datos de alta calidad. La aplicación de estándares y requisitos de datos específicos sobre cómo se publican los datos en formatos particulares basados en tecnología moderna puede mejorar la calidad de los datos. Fomentar políticas y prácticas que permitan eficientemente la recopilación y revisión de datos puede garantizar de manera similar que los datos del gobierno e los compartidos pelas empresas privadas sean más útiles para la sociedad.

La transparencia contribuí para que el uso de datos sea hecho de manera mas responsable, el efectivo valor de los datos solo puede realizarse si los datos se usan realmente en la práctica, y existen múltiples formas y estrategias para extraer valor de los datos, incluida la investigación de políticas, estadísticas, evaluación de programas y ciencia de datos, entre otros campos. En cada uno de estos dominios, la capacidad de vincular, combinar y compartir datos es cada vez más relevante. Este pilar se centra en las estrategias para acceder y compartir datos gubernamentales, así como para garantizar la confianza pública en la protección de datos personales y sensibles.

## VI Algoritmos

La inteligencia artificial como la internet de las cosas y *machine learning* van tener cada vez mas como materia prima datos planificados con *standards* internacionales de preferencia. Las nuevas tecnologías que agregan, limpian, recopilan, administran, acceden y utilizan los datos precisan que haya padrones y que haya abertura permitiendo el flujo de datos. Hasta mismo las tecnologías de preservación de la privacidad o enfoques para mejorar la seguridad en las transacciones de datos, dependen cada vez más del desarrollo y la implementación de flujo de datos, sin ellos es difícil desarrollar nuevas soluciones al adoptar nuevos enfoques para los datos y políticas gubernamentales tan necesarios en época de pandemia e emergencia publica.

Dos desarrollos en este contexto son: las tecnologías de mejora de la privacidad (PET – *Privacy Enhancing Technologies*) junto con enfoques de minimización, anonimización, cifrado de datos; y tecnologías que se están desarrollando para manejar las propias transacciones de datos (Engin y otros, 2020, pág. 12). Que

los datos generen nuevas relaciones de poder y políticas es evidente, incluso en varias de las recientes controversias políticas (Ruppert. Isin y Bigo, 2017).

Los algoritmos deben ser verificados, certificados y regulados. Esto ya ocurre con las aplicaciones financieras, en las cuales los desarrolladores de software necesitan demostrar que sus algoritmos de negociación han sido probados exhaustivamente, demuestran el mejor rendimiento y no resultan en la manipulación del mercado (Stucke y Grunes, 2016).

La evaluación del impacto del algoritmo se centra en la toma de decisiones, principalmente en la robustez, la equidad y la capacidad de explicación del sistema. Esto corresponde a definir límites de uso y una fecha de vencimiento para generar confianza entre las partes interesadas, responsabilizar a los creadores del sistema por los resultados de sus cronogramas y proporcionar más transparencia, la posibilidad de rendir cuentas y recibir auditorías.

La robustez significa que los sistemas deben ser seguros y estar protegidos contra fallas, intrusiones o comprometer los datos sobre los que están capacitados. Equidad, que los sistemas deberían utilizar datos y modelos de capacitación sin sesgos, para evitar el trato injusto de ciertos grupos. Y explicabilidad, que los sistemas deben proporcionar decisiones o sugerencias que puedan ser entendidas por sus usuarios y desarrolladores. Es posible tener sistemas que, por diseño, sean capaces de satisfacer estas tres demandas (Koshiyama, 2019).

Para este propósito, específicamente con respecto a la explicabilidad, el desarrollador debe ser capaz de responder el siguiente cuestionario: ¿Ha evaluado en qué medida se pueden entender las decisiones y, por lo tanto, el resultado obtenido por el sistema de IA? ¿Garantizó una explicación de por qué el sistema hizo una determinada elección, lo que resultó en un cierto resultado que todos los usuarios pueden entender? ¿Diseñó el sistema de IA con la capacidad de interpretación en mente desde el principio? ¿Ha investigado e intentado utilizar el modelo más simple e interpretable posible para la aplicación en cuestión? ¿Ha evaluado si puede examinar la interpretabilidad después de la capacitación y el desarrollo del modelo o si tiene acceso al flujo de trabajo interno del modelo? (Koshiyama, 2019).

La inteligencia artificial tiene indudablemente un enorme potencial para traer beneficios al Gobierno y a la sociedad, especialmente si existe colaboración entre diferentes sectores, incluso a través de asociaciones público-privadas, desarrollo e incentivos para centros y laboratorios de innovación; creación de consejos, redes y comunidades que involucran diferentes áreas del gobierno y la sociedad; y una gestión estratégica de la transparencia, colaboración y apertura de los datos del gobierno.

Sin embargo, se deben tomar medidas constantes para evitar impactos sociales negativos. El principal riesgo es el uso de datos asimétricos para procesar

datos que se utilizarán en la provisión de servicios públicos. Esto puede conducir a la discriminación, la manipulación o la falta de control sobre lo que subyace a las decisiones de las agencias públicas cuando se automatiza.

Un cierto bias en los datos es inevitable incluso en modelos no automatizados. Las decisiones siempre están sujetas a un cierto grado de subjetividad e imperfecciones. Esto continuará ocurriendo durante la captura y el procesamiento de datos electrónicamente. Los estereotipos y prejuicios en la sociedad también tienden a ser reproducidos y amplificados por los sistemas de inteligencia artificial. Finalmente, los límites resultantes de la formulación de los problemas, la selección de parámetros y los criterios que serán utilizados por los algoritmos también representan un sesgo.

Lo importante al usar la tecnología en el sector público es identificar constantemente estas posibilidades y actuar de manera efectiva para mitigar estos sesgos tanto como sea posible. Para hacer este análisis, generalmente se divide en grupos de comparación y si uno no tiene una tasa de aprobación, por ejemplo, igual o superior al 80% de la tasa de aprobación de grupo más alta, esto se considera evidencia de que hay un impacto adverso, no se alcanzó el umbral de 4/5 al probar la equidad algorítmica (Lay, Koshiyama y Tarman, 2019). Equilibrar varios intereses a menudo en conflicto es un problema ético y técnico extremadamente desafiante, pero debe abordarse.

Por lo tanto, no solo en la administración pública sino principalmente allí, se deben desarrollar, aplicar y exigir estándares y certificaciones. Basado en el supuesto de que la toma de decisiones utilizando algoritmos nunca es neutral o naturalmente ética, equilibrada y justa, se debe desarrollar un proceso de estandarización y certificación antes de que estos sistemas se pongan en uso. Esto debe incluir una declaración de riesgos potenciales (incluidos los sesgos identificados) para los casos de uso según lo previsto, y también de forma no intencionada y con y sin actores maliciosos (Lay, Koshiyama y Tarman, 2019).

También es importante que haya un monitoreo constante de los resultados y el impacto, ya que el bias se puede reintroducir, crear y/o amplificar cuando se introducen nuevos datos en el sistema y también a partir de la retroalimentación humana. Esta tarea depende en gran medida del desempeño de las autoridades reguladoras, pero también requiere una comunidad técnica consciente, calificada y comprometida para desarrollar la tecnología de manera ética.

El paso a paso para verificar si los algoritmos están teniendo un impacto adverso generalmente incluye: a) auditar los datos de rendimiento que se utilizan para capacitar al modelo para garantizar que la muestra sea representativa; b) probar con un alto nivel de atención la presencia de impacto adverso en las predicciones del modelo; c) una vez encontrados, todos los factores que han

demostrado contribuir a los resultados sesgados deben eliminarse; d) para que el modelo sea entrenado nuevamente; y e) monitoreado hasta que el impacto adverso se resuelva y se elimine por completo (Koshiyama, 2019).

## **Conclusión**

Este texto empezó resaltando la importancia de la planeación, identificación de beneficiarios, y preparo de procesos antes de se aplicar tecnologías de coleta y uso de los datos. Aun en tiempos de emergencia sanitaria es necesaria una estrategia de apertura de estos datos que están siendo colectados, compartidos y alimentando las herramientas de inteligencia artificial.

En seguida se describió las etapas de un proceso de apertura de datos y el ecosistema de datos. Actualmente son varios los conjuntos de datos actualmente siendo colectados, pero aun esparzo, de diversas fuentes y non uniformes. Y cualquier proyecto de abertura de datos hay que hacer una distinción entre los datos que serán compartidos y los datos personales y sensibles.

Para ilustrar la problemática fuera hecho un estudio de caso en sitios electrónicos brasileños que han divulgado tales datos pertinentes a la pandemia. El análisis de estos evidenció otro desafío: el volumen, los detalles de la nomenclatura, la complejidad de los datos y, a veces, los campos que se dejan sin información o se rellenan solo con números en las tablas lo hacen incluso cuando hay una clara intención de hacer que los datos sean abiertos, será necesario que el usuario tenga muchos conocimientos técnicos o aplique inteligencia artificial para procesar este volumen de datos de una manera que los haga sentido.

Con la finalidad de complementar el estudio de caso en los sitios electrónicos brasileños fue hecha también una analice normativa brasileña a respecto de los datos abiertos analizando la ley de transparencia, la ley de acceso a información y la ley general de protección de los datos.

En la quinta parte la vigilancia tecnológica y la “securitización de la salud” fueron temas centrales para resaltar la importancia de siempre considerar la dicotomía en la adopción de opciones tecnológicas. El punto central de esta parte es que son varios los beneficios posibles y un catalogo de datos abiertos pude favorecer los beneficios y mitigar los maleficios.

La ultima parte fue sobre la necesidad de se tener en comta los aspectos técnicos de los algoritmos que serán utilizados para coleta e tratamiento de los datos. El punto principal de esta parte es que los algoritmos deben ser verificados, certificados y regulados, principalmente en la robustez, la equidad y la capacidad de explicación del sistema.

Construir un catalogo de datos abiertos relacionados con covid19 avanzar en el acceso y uso de información de alta calidad, contribuir para construcción de una cultura organizacional de transparencia, y crea demanda y usos creativos para estos datos entre otras posibilidades. Con un catalogo de datos abiertos el estoque masivo de datos que ya están siendo colectados y los que serán en breve pasan de la propiedad privada de una institución o de ser detenida por un órgano gubernamental y pasa a ser información pública y de libre uso, que cumpla con las características técnicas adecuadas y sea de interés y utilidad para los usuarios. Menos susceptible entonces a las voluntades particulares de un grupo o de determinado gobierno.

Los datos abiertos son la oportunidad para que gobierno empresas y ciudadanos cooperen para generar más valor a la transparencia. En una realidad de emergencia como la pandemia de covid19 un flujo continuo de intercambio de información e una economía colaborativa motivada por los datos abiertos, poden ser materia prima para crear valerosas iniciativas en términos económicos y en gaño sociales. Por esto la creación del catalogo de datos abiertos relacionados con pandemia es tan esencial. Librando siempre que el monitoreo y evaluación también son importantes, para que haya continuidad, y continuo mejoramiento del catalogo de datos abiertos. Las iniciativas no pueden ser como un monólogo del gobierno hacia los ciudadanos (y viceversa) sino como una forma de interacción permanente entre ellos.

---

### **Cataloging Covid-19 data: transparency, innovation and public good**

**Abstract:** Due to the Covid-19 pandemic technology firms and governmental bodies are in discussion about the creation and organisation of many forms of data collection including georeferenced and health data to monitor lockdowns, quarantines, isolation policies and social distancing. Health, location, shopping behaviour, money transfers, facial biometrics, digital certificates etc. can be used and serve as a powerful weapon for fighting the virus. However an urgent crisis of digital rights will emerge if the data collected and parameterized is not maintained in an open manner, respecting both data regulation about privacy and protection of the personal and sensitive information, as well as reassuring trustworthiness (fairness, lawfulness and transparency), of the algorithmic suggestions that are producing analysis and an array of decisions. States are increasingly assisted by private companies by using their collected telephone data, CCTV images, temperature control points, airline and rail reservations, credit card information, online shopping records, social networking data, facial recognition and in some cases, drone data. In this context transparency and openness of data is essential. We case studied Brazil in order to illustrate the ideas of this paper, the analyses were made over several Brazilian websites during March and June of 2020.

**Keywords:** Covid-19. Open data. Data catalogue. Privacy. Data protection. Artificial intelligence.

**Summary:** Introduction – I Planning, identification of beneficiaries, and preparation of processes – II Data sets and the Brazilian case – III Open data in Brazil: law on transparency and access to information – IV Data protection – V Surveillance – VI Algorithms – Conclusion – References

### **Catálogo de Dados Covid-19: transparência, inovação e compromisso social**

**Resumo:** Em razão da pandemia as empresas de tecnologia e governos estão em discussões sobre a criação de muitas formas de coleta de dados, incluindo os georreferenciados e de saúde, para realizar o monitoramento de locautes, bloqueios, quarentenas, isolamentos e distanciamento social.

Os dados de saúde, localização, compras, transferência de dinheiro, biometria facial e certificados digitais podem de fato ser uma arma poderosa no combate ao vírus. Entretanto, se os dados coletados e parametrizados não forem abertos, com respeito à privacidade e proteção de dados pessoais ou sensíveis, se não houver justiça, legalidade e transparência sobre como os algoritmos estão produzindo sua análise e até que ponto suas decisões políticas e administrativas estão sendo fundadas em ingerências tecnológicas, podemos estar caminhando para uma urgente crise de direitos digitais. Neste texto o caso brasileiro será detalhado mais a fundo a título de exemplo. O estudo de caso em alguns sites brasileiros foi feito entre março e junho de 2020.

**Palavras-chave:** Covid-19. Abertura de dados. Catálogo de dados. Privacidade. Proteção de dados. Inteligência artificial.

**Sumário:** Introdução – I Planejamento, identificação de beneficiários e preparo de processos – II Os conjuntos de dados e o caso brasileiro – III Os dados abertos no Brasil: Lei de transparência e de acesso à informação – IV Proteção de Dados – V Vigilância – VI Algoritmos – Conclusão – Bibliografia

## Bibliografia

Bioni Bruno y otros (2020). “Privacidade e pandemia: recomendações para o uso legítimo de dados no combate à COVID-19. Conciliando o combate à COVID-19 com o uso legítimo de dados pessoais e o respeito aos direitos fundamentais”. Disponible en: [https://www.dataprivacybr.org/wp-content/uploads/2020/04/relatorio\\_privacidade\\_e\\_pandemia\\_final.pdf](https://www.dataprivacybr.org/wp-content/uploads/2020/04/relatorio_privacidade_e_pandemia_final.pdf)

Brasil. Congresso Nacional (2018). *Lei Geral de Proteção de Dados Pessoais* (LGPD) L13709/2018.

Brasil. Congresso Nacional (2011). LAI – Lei de acesso à informação L12527/2011.

Brasil. Congresso Nacional (2009). “Lei da Transparência LC 131/2009”. Disponible en: [http://www.planalto.gov.br/ccivil\\_03/Leis/LCP/Lcp131.htm](http://www.planalto.gov.br/ccivil_03/Leis/LCP/Lcp131.htm)

Brasil. Controladoria-Geral da União (2020). “Documentos de execução da despesa pública – Covid19 – Portal da transparência”. Disponible en: <http://www.portaltransparencia.gov.br/despesas/favorecido?de=01/03/2020&ate=31/03/2020&funcaoSubfuncao=FN10&funcaoSubfuncao=SB122&programa=5018&acao=21C0&programaGoverno=1&grupo=3&elemento=41&modalidade=41&orgaos=UG257001&ordenarPor=valor&direcao=desc&minifiedPath=%2Fminified&projectVersion=1.20.0>

Brasil. DATASUS (2020). “DataSUS – Sistemas e Aplicativos”, DATASUS.

Brasil. Executivo Federal (2011). “E-SIC. Site sobre a Lei de Acesso à Informação (LAI)”. Acesso à Informação. Disponible en: <https://www.gov.br/acessoainformacao/pt-br/pagina-inicial>

Brasil. Ministério da Saúde (2020). “Contratos coronavírus”. Disponible en: <https://www.saude.gov.br/contratos-coronavirus>

Brasil. Ministério da Saúde (2019). “Saúde sem Fake News”. Disponible en: <https://www.saude.gov.br/fakenews>

Brasil. Secretaria de Tecnologia da Informação. Ministério do Planejamento, Desenvolvimento e Gestão (2020). “Portal Brasileiro de Dados Abertos”. Disponible en: <http://www.dados.gov.br/>

Brasil. Sistema Único de Saúde (2020). “Coronavírus - SUS”. Disponible en: <https://www.gov.br/pt-br/apps/coronavirus-sus>

Canales, María Paz (2020). “La herejía tecno-optimista florece en pandemia”. *Derechos Digitales America Latina*.

Castro, Douglas (2017). “As Dimensões Internacional E Nacional Da Securitização Da Saúde: Um Estudo Empírico Sobre O Vírus Zika (The International and National Dimensions of the Health Securitization – An Empirical Study of the Zika Virus)”. No. ID 3051188. Rochester, NY. Social Science Research Network. octubre.



- Castro, Douglas y James Oliveira Santos (2020). "Securitization of the Health and Economy in the COVID Times", AfronomicsLaw.
- Cetic.br – Centro Regional para o Desenvolvimento da Sociedade da Informação (2019). "Domicílios que possuem equipamento TIC 2019". Disponible en: <https://cetic.br>
- CF (1988). Constituição da República Federativa do Brasil.
- Claypoole, Theodore (2020). "COVID-19 and Data Privacy: Health vs. Privacy", Business Law Today from ABA. Disponible en: <https://businesslawtoday.org/2020/03/covid-19-data-privacy-health-vs-privacy/>
- Cohen, Julie (2019). "Cohen reviews The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power". Surveillance & Society. vol. 17. No. 1/2.
- Data Foundation (2020). "The Data Foundation in 2020: Advancing Pillars for Access and Use of High-Quality Information". Disponible en: <https://www.datafoundation.org/blog-list/the-data-foundation-in-2020-advancing-pillars-for-access-and-use-of-high-quality-information/2020>
- Devlin, Hannah y Kate Proctor (2020). "Coronavirus UK: health passports 'possible in months' | Politics | The Guardian". Disponible en: <https://www.theguardian.com/politics/2020/may/03/coronavirus-health-passports-for-uk-possible-in-months>
- Engin, Zeynep y otros (2020). "Data-driven urban management: Mapping the landscape". Journal of Urban Management. vol. 9. No. 2. 1 de junio.
- EU (2016). "Carta de los Derechos Fundamentales de la Unión Europea EUR-Lex - 12016P/TXT". Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex%3A12016P%2FTXT>
- EU, European Parliament and Council (2016). GDPR Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). OJ L, mayo.
- Ferraz, Lara y Raíssa Moura (2020). "InLoco – Meios de Controle à Pandemia da COVID-19 e a Inviolabilidade da Privacidade".
- Fok, Wendy W. y Antoine Picon (2016). Digital Property: Open-source Architecture. Oxford. Academy Press, noviembre.
- InfoAmazonia. Rede Infoamazonia (Mãe d'Água) (2020). "Distantes de UTIs e respiradores, indígenas da Amazônia tentam se blindar do vírus", InfoAmazonia.
- IRIS – Instituto de Referência em Internet e Sociedade (2020). "Código aberto – um caminho de união no combate à pandemia", IRIS-BH. Disponible en: <https://irisbh.com.br/codigo-aberto-um-caminho-de-uniao-no-combate-a-pandemia/>
- Justen, Álvaro (2020). "COVID-19 - Datasets - Painel Covid - Brasil.IO". Disponible en: <https://brasil.io/covid19/>
- Kitchin, Rob (2014). The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. SAGE Publications Ltd, de agosto de.
- Kitchin, Rob y Martin Dodge (2019). "The (In)Security of Smart Cities: Vulnerabilities, Risks, Mitigation, and Prevention". Journal of Urban Technology. vol. 26. No. 2, Routledge. 3 de abril.
- Koshiyama, Adriano (2019). "Algorithmic Impact Assessment: Fairness, Robustness and Explainability in Automated Decision-Making".
- LabCidade (2020a). "Denúncia: Ministério da Saúde retira dados que permitiam a territorialização da pandemia", LabCidade.
- LabCidade (2020b). "Simplificação da leitura do comportamento da epidemia no território dificulta seu enfrentamento".
- Lay, Aixe. Adriano Koshiyama y Zeynep Derya Tarman (2019). "Algorithmic bias in recruitment Response to CDEI Call for Evidence". UK.

- Lee, Tim Berners (2012). “5 estrelas de Dados Abertos”. Disponível em: <http://5stardata.info/es/>
- Logarezzi, Lia (2016). Guia prático da lei de acesso à informação. Artigo 19, Brasil.
- Mendes, Laura Schertel y Danilo Doneda (2019). “Um perfil da nova Lei Geral de Proteção de Dados brasileira”. Governança e regulações da Internet na América Latina: análise sobre infraestrutura, privacidade, cibersegurança e evoluções tecnológicas em homenagem aos dez anos da South School on Internet Governance. Rio de Janeiro, RJ.
- Moura, Camille (2020). “Datusus CEP”.
- OEA – Organização dos Estados Americanos (2020). “Resolucion 1/2020 – Pandemia y Derechos Humanos en las Américas”.
- OEA – Organização dos Estados Americanos (2016). “Material Guía Curso Introducción a los Datos Abiertos”.
- Oliver, Nuria y otros (2020). “Mobile phone data and COVID-19: Missing an opportunity?” arXiv:2003.12347 [cs] (arXiv: 2003.12347). 27 de marzo.
- Orwell, George (1949). 1984. Houghton Mifflin Harcourt.
- PBA (2020). “Catalogo de Datos Abiertos Provincia de Buenos Aires”. Disponível em: <https://catalogo.datos.gba.gov.ar/>
- Reventlow Nani Jansen (2020). “COVID-19 and the Digital Rights Crisis”, IntLawGrls.
- Ruppert, Evelyn. Engin Isin y Didier Bigo (2017). “Data politics:” Big Data & Society, SAGE PublicationsSage UK: London, England. 3 de julio.
- Souza, Allan Rocha de y otros (2018). Marcos legais nacionais em face da abertura de dados para pesquisa em saúde: dados pessoais, sensíveis ou sigilosos e propriedade intelectual. Rio de Janeiro, RJ. Fundação Oswaldo Cruz.
- Stucke, Maurice y Allen Grunes (2016). Big Data and Competition Policy. Oxford, United Kingdom. Oxford University Press, agosto.
- SVS / MS Secretaría de Vigilancia Sanitaria (2020). “Portal COVID-19 -Coronavírus Brasil”. Disponível em: <https://covid.saude.gov.br/>
- Veale, Michael. Reuben Binns y Lilian Edwards (2018). “Algorithms that remember: model inversion attacks and data protection law”. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. vol. 376. No. 2133. 28 de noviembre.
- Zuboff, Shoshana (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York. PublicAffairs, enero.

---

Informação bibliográfica deste texto, conforme a NBR 6023:2018 da Associação Brasileira de Normas Técnicas (ABNT):

DENNY, Danielle Mendes Thame; KAZIM, Emre; KOSHIYAMA, Adriano. Catalogo de datos Covid-19: transparencia, innovación y compromiso social. *Direitos Fundamentais & Justiça*, Belo Horizonte, ano 14, p. 43-68, nov. 2020. Número especial.

---

Recebido em: 15.07.2020

Aprovado em: 14.09.2020

Cota Convite