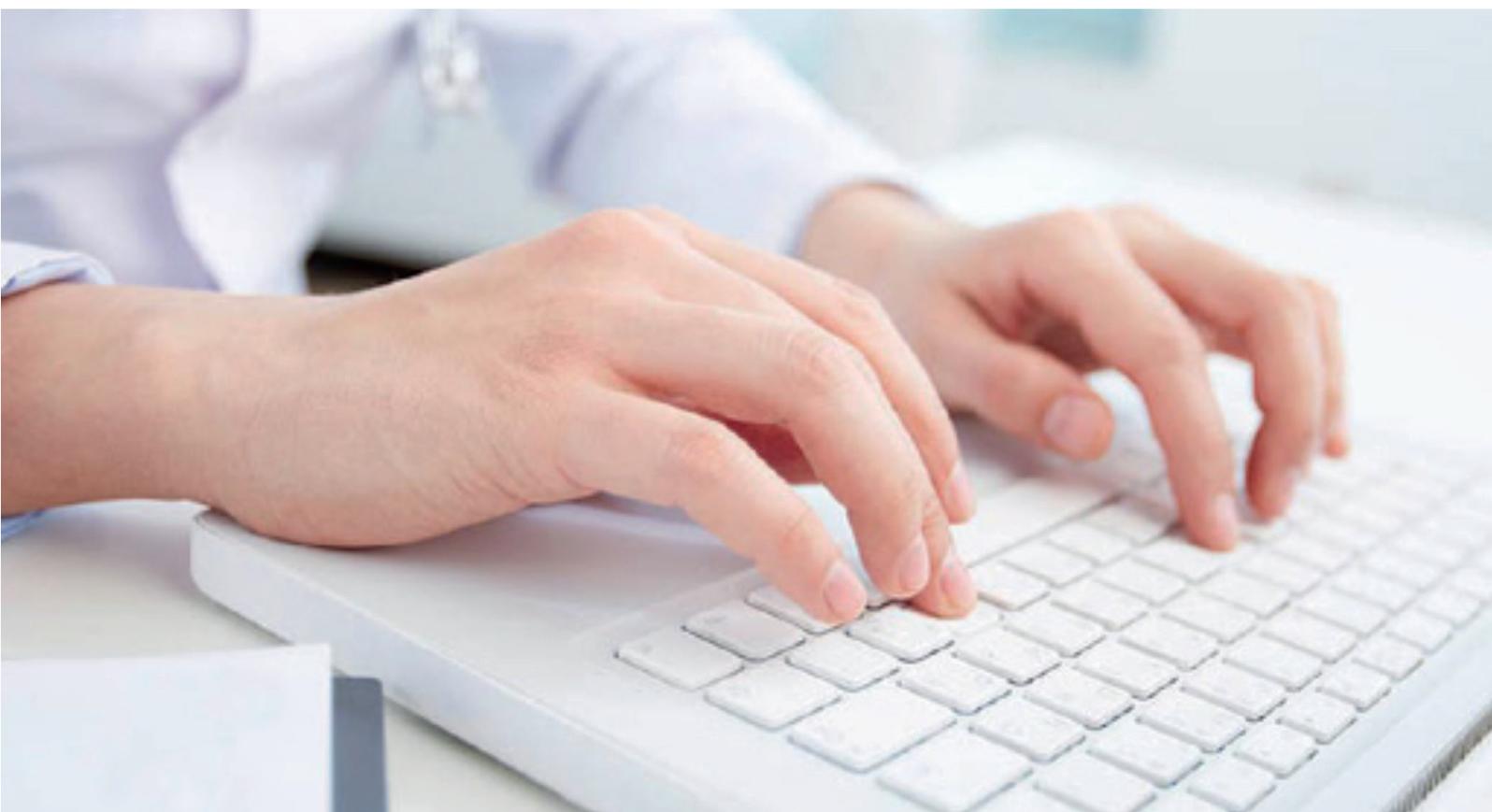


INFORME

**PROPUESTA DE PUBLICACIÓN DE DATOS
ABIERTOS SOBRE LA EPIDEMIA
DE COVID-19 EN URUGUAY**

Febrero de 2021





PROPUESTA DE PUBLICACIÓN DE DATOS ABIERTOS SOBRE LA EPIDEMIA DE COVID-19 EN URUGUAY

Bermolen, Paola; Etcheverry, Lorena; Fariello, María Inés; Mordecki, Ernesto

Introducción

Los datos abiertos son aquellos que se encuentran disponibles en formatos estándares, abiertos y están disponibles en la web para que cualquier persona pueda acceder a ellos, reutilizarlos, combinarlos y redistribuirlos para crear nuevos servicios, visualizaciones o realizar investigaciones a partir de ellos. Desde 2010, Uruguay ha desarrollado activamente una política de publicación de datos abiertos impulsada por AGESIC. En el Plan de Gobierno Digital 2020¹ se incluye como uno de los objetivos de Uruguay ampliar y mejorar la oferta de datos abiertos de gobierno en el Catálogo Nacional de Datos Abiertos² y disponer herramientas para facilitar su acceso y visualización.

El objetivo de este documento es contribuir en ese sentido, proponiendo ampliar los datos que actualmente se publican sobre la epidemia de COVID-19, evitando la publicación de datos que permitan identificar a las personas (nombre y apellido, documento, etc.) y adhiriendo a buenas prácticas para reducir el riesgo de la reidentificación de las personas involucradas al combinar los datos publicados con otras fuentes de información.

Motivación

La comunidad científica uruguaya jugó y juega un rol importante en la comprensión y el control de la pandemia. La disponibilidad de datos permitirá consolidar esta experiencia a través de estudios científicos, de los cuales puedan resultar tanto asesoramiento en los más diversos ámbitos, como publicaciones científicas que

1 Plan de Gobierno Digital 2020, Uruguay, https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/sites/agencia-gobierno-electronico-sociedad-informacion-conocimiento/files/2019-05/Descargar%20Plan%20de%20Gobierno%20Digital%20Uruguay%202020_0_0.pdf

2 Catálogo de Datos Abiertos, Uruguay, <https://catalogodatos.gub.uy/>



colaboren con el conocimiento de esta nueva enfermedad. Como ejemplos relevantes, el conocimiento de la incidencia en grupos diversos de la población (escolares, liceales, adultos mayores, trabajadores de la salud, etc.) permitirá comprender los riesgos asociados a las diversas actividades que se están gradualmente retomando. Cuestiones biológicas como la duración de las distintas etapas de la enfermedad en las personas según sus características, duración de las etapas de internación, factores de riesgo como la edad, permitirán profundizar el manejo responsable de la pandemia. Capítulo aparte merece el tema de posibles futuras pandemias de otros virus similares, así como de este mismo virus, dado que aún no es conocido si las posibles mutaciones del virus permitirán una vacunación efectiva de largo plazo.

Experiencia internacional

Muchos países publican diariamente diversos datos sobre la epidemia de COVID-19 en formatos abiertos, desde datos sobre los casos a estadísticas de acatamiento a las medidas de distanciamiento social³.

En la región cabe destacar el caso del Gobierno de Río Grande do Sul⁴, que, además de publicar visualizaciones y datos resumidos sobre la epidemia, publica diariamente datos detallados sobre cada uno de los casos detectados en formato CSV. Brasil también publica a nivel nacional datos abiertos y detallados sobre cada uno de los casos sospechosos, a través de su Portal de Datos de Salud⁵. El Gobierno de la Ciudad de Buenos Aires también publica datos abiertos sobre la epidemia de COVID-19. A través del sitio COVID-19 en datos⁶ es posible visualizar datos resumidos, así como acceder a datos detallados de cada caso como sexo, edad, provincia, clasificación (confirmado, sospechoso, descartado, negativizado), barrio, fecha de toma de muestra, tipo de contagio, alta_médica y fecha de fallecimiento⁷.

A nivel internacional podemos mencionar países similares en población al nuestro, como Nueva Zelanda, que publica datos sobre casos sospechosos y casos confirmados, brotes y su estado, con detalles que permiten una visualización clara de edades, sexo, y

3 En el sitio Open Data Watch es posible acceder a un relevamiento exhaustivo por países y tipo de datos <https://opendatawatch.com/what-is-being-said/data-in-the-time-of-covid-19/>

4 Pánel Coronavirus RS <https://ti.saude.rs.gov.br/covid19/>

5 Open DATASUS: Datos sobre notificaciones de casos sospechosos <https://opendatasus.saude.gov.br/dataset/casos-nacionais>

6 COVID-19 en Datos, Ciudad de Buenos Aires <https://www.buenosaires.gob.ar/coronavirus/datos>

7 Los metadatos del conjunto de datos publicado puede accederse aquí <https://data.buenosaires.gob.ar/dataset/casos-covid-19/archivo/df05ae09-975f-40fc-9011-c5ccff0024ab>



localización⁸. Finlandia publica también datos detallados de los casos como edad, sexo, y cantidad de internaciones en los principales hospitales, detallando la ubicación de los casos hasta el nivel municipal cuando hay más de 4 casos por municipio⁹. Ambos países, con características distintas, presentan un importante control de la pandemia.

Experiencia nacional

En Uruguay, el SINAIE publica diariamente datos estadísticos mediante su Visualizador de casos¹⁰. Esta información es recopilada y publicada diariamente en formato abierto (CSV) por el Grupo Uruguayo Interdisciplinario de Análisis de Datos de COVID-19¹¹ desde comienzos de la epidemia, y el SINAIE publica los indicadores del visualizador en formato CSV desde mediados de junio en el Catálogo de Datos Abiertos¹². Si bien esta información es útil, carece del nivel de detalle necesario para realizar siquiera un análisis descriptivo completo, y es publicada en general días después de publicada la información, lo que reduce mucho su utilidad. Los datos que se propone dejar disponibles permitirían por ejemplo tener una estimación del porcentaje de reporte de casos ajustada a la realidad uruguaya sin necesidad de utilizar parámetros obtenidos de la literatura internacional (e.g. distribución de tiempo entre fecha de inicio de síntomas e ingreso a CI/CTI o muerte, así como su relación con la edad, lo cual podría determinar el uso de distribuciones mixtas)¹³.

Propuesta de publicación de datos sobre casos

Se propone la publicación diaria de una tabla en formato CSV, donde cada fila representa un caso infectado de COVID-19 y las columnas representan las variables que se indican a continuación:

8 Datos sobre COVID-19, Ministerio de Salud de Nueva Zelanda, <https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus>

9 Datos sobre COVID-19, Ministerio de Salud de Finlandia , <https://thl.fi/en/web/infectious-diseases-and-vaccinations/what-s-new/>

10 Visualizador de casos de coronavirus COVID-19, <https://www.gub.uy/sistema-nacional-emergencias/pagina-embebida/visualizador-casos-coronavirus-covid-19-uruguay>

11 Grupo Uruguayo Interdisciplinario de Análisis de Datos de COVID-19, <https://guiad-covid.github.io/>

12 Datos estadísticos de evolución de casos de COVID-19, <https://catalogodatos.gub.uy/dataset/sinae01-covid19>

13 Herrera,D; Bermolen,P.;Fariello, M. Estimación del porcentaje de reporte de casos de COVID-19 en Uruguay, Abril 2020, <https://guiad-covid.github.io/publication/nota3/>



Variable/s:	Nº caso
Tipo de dato:	número entero
Concepto / Justificación	Permite identificar registros que corresponden al mismo caso e identificar modificaciones en el estado de las variables a medida que el caso evoluciona

Variable/s:	Edad
Tipo de dato:	Literal, Valores posibles 0-5, 6-11, 12-17, 18-24, 25-30, 30-34, 35-39, etc.
Concepto / Justificación	Permitiría el análisis por franjas etarias

Variable/s:	Sexo biológico
Tipo de dato:	Literal, Valores posibles: Femenino, Masculino
Concepto / Justificación	Permite el análisis por sexo

Variable/s:	Ubicación geográfica
Tipo de dato:	Departamento, Localidad, Barrio o Departamento, Sección Censal, Segmento Censal, Zona Censal.
Concepto / Justificación	<p>Información sobre la ubicación del caso, lo más granular posible. Podría usarse tanto la jerarquía Departamento, Localidad, Barrio como la definida por INE¹⁴. Es necesario definir cuál se usará y usar una sola.</p> <p>Para reducir la posibilidad de reidentificación, sea cual sea la jerarquía que se defina usar, se recomienda refinar la ubicación mientras exista cierta cantidad de casos en ese nivel (por ejemplo: si en el Barrio X del Departamento Y hay 5 o más casos, se indica Barrio, si no sólo se indica Departamento)</p>

14 Unidades Geoestadísticas (UGeo) - Uruguay, INE <http://www.ine.gub.uy/documents/10181/18006/definiciones+para+web.pdf/896410b7-f7c2-40f0-b5c3-4d7b7326f51c>



Variable/s:	Fecha de última exposición
Tipo de dato:	Fecha
Concepto / Justificación	Fecha de última exposición, si el caso estuvo en contacto con otro caso

Variable/s:	Caso asintomático
Tipo de dato:	Booleano
Concepto / Justificación	Indica si es o no un caso asintomático

Variable/s:	Fecha de inicio de síntomas
Tipo de dato:	Fecha
Concepto / Justificación	Fecha inicio de síntomas, en caso de no ser asintomático

Variable/s:	Fecha de aislamiento
Tipo de dato:	Fecha
Concepto / Justificación	Esta fecha permite conocer cuánto tiempo estuvo posiblemente contagiando

Variable/s:	Fecha de resultado de test
Tipo de dato:	Fecha
Concepto / Justificación	Fecha de confirmación de test positivo

Variable/s:	¿Estuvo en otros países de circulación de COVID-19?
Tipo de dato:	Booleano
Concepto / Justificación	Se busca identificar si es un contagio en el exterior

Variable/s:	¿Estuvo en contacto cercano con caso confirmado de COVID-19?
Tipo de dato:	Booleano



Concepto / Justificación	Se busca identificar si el contagio es por contacto con caso conocido
---------------------------------	---

Variable/s:	¿Requiere hospitalización?
Tipo de dato:	Booleano
Concepto / Justificación	Si en caso de que requiera ingreso hospitalario de cualquier nivel

Variable/s:	Cantidad de días en CI
Tipo de dato:	Entero
Concepto / Justificación	

Variable/s:	Cantidad de días en CTI
Tipo de dato:	Entero
Concepto / Justificación	

Variable/s:	Fecha de recuperación
Tipo de dato:	Fecha
Concepto / Justificación	Si no fallece, fecha de recuperación de acuerdo a los criterios de DEVISA.

Variable/s:	¿Fallece?
Tipo de dato:	Booleano
Concepto / Justificación	Indica si fallece o no

Variable/s:	Fecha de fallecimiento
Tipo de dato:	Fecha
Concepto / Justificación	Si fallece, fecha de fallecimiento