

UNCovid: Hacia un Tablero COVID de la Universidad Nacional del Comahue

Luis Reynoso¹

¹ Facultad de Informática, Universidad Nacional del Comahue (UNCo),
Buenos Aires 1400, 8300, Neuquén Capital

Tel: (0299) 4490326

luis.reynoso@fi.uncoma.edu.ar

Resumen: Una de las principales variables para medir el impacto y alcance de los datos abiertos es su reutilización. Y en la cadena de valor de datos abiertos, contribuyen tanto los creadores de datos abiertos como desarrolladores y enriquecedores que los explotan. La agregación de diferentes conjuntos de datos, permite crear nuevos datos que pueden conducir a nuevos servicios o productos de datos. Tal es el caso del tablero UNCovid, el cual integra e agrega por un lado datos abiertos del Ministerio de Salud y del IGN, como así también del servicio ArgenMap, un servicio WMTS (estándar de OGC para servir y obtener teselas de mapas georeferenciadas por la red). La información mostrada en el tablero permite visibilizar datos sobre el progreso de la vacunación por provincias, incidencia de los contagios, cantidades de contagios a nivel de distintas jurisdicciones (provincial y departamental), cantidad de fallecidos, y otras estadísticas según género y franjas etáreas. El tablero visibiliza la información reportando los métodos aplicados. Permite una visualización clara en distintas jurisdicciones, y es un importante ejemplo de la importancia de transparencia de datos y la reutilización de información en la cadena de valor, los cuales son importantes objetivos finales en el acceso a la información pública.

Palabras Claves: COVID-19, contagios, vacunación, inoculación, datasets, datos abiertos, tablero.

1. INTRODUCCIÓN

Varios organismos estatales ponen a disposición cantidades cada vez mayores de datos abiertos en Internet. Esto se debe fundamentalmente garantizar el efectivo ejercicio del derecho de acceso a la información pública, promover la participación ciudadana y la transparencia de la gestión pública, en el marco de La Ley 27275 de Acceso a la Información Pública. Estos conjuntos de datos abiertos (datasets) están

creciendo día a día, Para acceder a Datos Abiertos se utilizan con mayor frecuencia los portales nacionales de Datos Abiertos, seguidos de los que acceden directamente a las propias administraciones públicas. Sin embargo, los conjuntos de datos abiertos, pueden ser tan grandes, que mirar la información en modo tabular no nos dice mucho. En cambio geolocalizar la información en mapas nos ayuda también a comprender la distribución de los datos, y es muy ilustrativa una vez reflejada en un mapa. Del mismo modo los tableros de mando, son una herramienta de visualización de los principales indicadores y datos que nos ayudan a facilitar y simplificar la comprensión el estado general de una temática. Tal es el caso de UNCovid (tablero Covid de la Universidad Nacional del Comahue (UNCo)) que utiliza como insumos datos abiertos del ministerio de salud, y combina a su vez el servicio WMTS del IGN ArgenMap.

La información de los datasets del Ministerio de Salud es actualizada por el Ministerio diariamente a las 21:30 hs. (si bien la fecha de actualización de la página se actualiza cerca de las 20 hs. es a las 21.30 hs el horario en el cual el dataset es distinto del dataset del día anterior) notificando CASOS COVID-19 registrados en el país con un corte del día a las 17:45 hs. Luego de esa hora, **UNCovid** genera los datos a partir del dataset publicado y publica el actual html indicando día y fecha de reporte. Este tablero está en desarrollo y actualmente una etapa de validación avanzada, comparando resultados obtenidos con aquellos publicados en el ministerio de Salud y otros tableros del país.

La presente ponencia se estructura de la siguiente forma: la sección 2 describe los datos abiertos y geoservicios utilizados en el tablero, la sección 3 detalla como el tablero muestra el progreso de la vacunación; la sección 4 se enfoca en la evolución de casos de contagio por jurisdicción (provincias y departamentos) en cantidad y de acuerdo a la incidencia por 100 mil habitantes. La sección 5 ejemplifica los gráficos incluidos y la sección 6 detalla la tecnología con la cual se desarrolló el tablero. Finalmente, la sección 7 enuncia conclusiones y trabajo futuro.

2. DATOS ABIERTOS Y GEOSERVICIOS UTILIZADOS POR UNCovid

Como mencionamos anteriormente el tablero UNCovid utiliza dos datasets: Vacunas contra COVID-19 (Ministerio de Salud, 2021a), y COVID-19. Casos registrados en la República Argentina (Ministerio de Salud, 2021b). La información de estos datos no esta geolocalizada pero se asocia a los datos los códigos de jurisdicciones (provincia y departamento) con la cual se pudo hacer un mapeo a la información geográfica al utilizar shp de provincias y departamentos, siendo estos dos datos abiertos del IGN (IGN, 2021b).

Todos los mapas del tablero utilizan el mapa base Argenmap, el WMTS definido por IGN (IGN, 2021), el cual incluye la denominación correcta de provincias y de las Islas Malvinas. Un WMTS es un estándar de OGC para servir y obtener teselas de mapas georeferenciadas por la red. ArgenMap es y será un importante recurso que puede ser utilizado por distintos agregadores de valor.

El tablero al estar construido en lenguaje R con el paquete leaflet permiten hacer zoom-in (al hacer doble-click en el mapa, o al utilizar el + del selector superior izquierdo) y zoom-out (al utilizar el - del selector superior izquierdo).

3. VACUNACIÓN

Al abrir el tablero lo que inmediatamente se visualiza es el mapa de cantidad de vacunas aplicadas por jurisdicción. El mapa muestra la cantidad de dosis aplicadas por provincia. Estos datos son reportados en el dataset “*Vacunas contra COVID-19. Dosis Aplicadas en la República Argentina*” (Ministerio de Salud, 2021a). Estas cantidades se corresponden al total de dosis aplicadas e informadas por las 24 jurisdicciones al Registro Federal de Vacunación Nominalizado **NoMiVac**. El tipo de vacuna aplicada, número y tipo de dosis registradas proviene de los datos aportados por cada jurisdicción y esta información es actualizada una vez al día, en el dataset del Ministerio de Nación.

La figura 1 muestra el mapa de vacunación en el cual se han definido cuatro grupos de marcadores circulares. Cada grupo representa vacunas distintas que se están aplicando: Sputnik V, Covishield, Sinopharm y Astrazeneca.

La etiqueta del marcador del círculo muestra el nombre de la vacuna y la cantidad de dosis aplicadas en la provincia, esto es, suma de la cantidad de la primera dosis, y la segunda dosis. Esta etiqueta es visible sin hacer click en el marcador

El popup (este aparece al hacer click en el marcador) muestra el nombre la vacuna, el nombre de la jurisdicción, la cantidad de primera dosis, y la segunda dosis.

Los círculos representan la cantidad de vacunas aplicadas en su primera dosis, para ello se utilizó la variable `primera_dosis_cantidad`. La cantidad total de dosis podría reflejar una cantidad mayor a la cantidad real de personas vacunadas por esta razón se decidió mostrar la cantidad de personas vacunadas con la primera dosis.

Los rangos utilizados para la paleta de colores se corresponden con los cuartiles de la variable `primera_dosis_cantidad`.

Los centros de cada círculo se corresponden con el baricentro de cada jurisdicción a excepción de Jujuy cuya posición fue modificada debido a que su disposición geográfica es similar a un polígono cóncavo.

Los radios de los círculos utilizan la fórmula $\sqrt{\text{cantidadPrimeraDosis}} * 0.05$

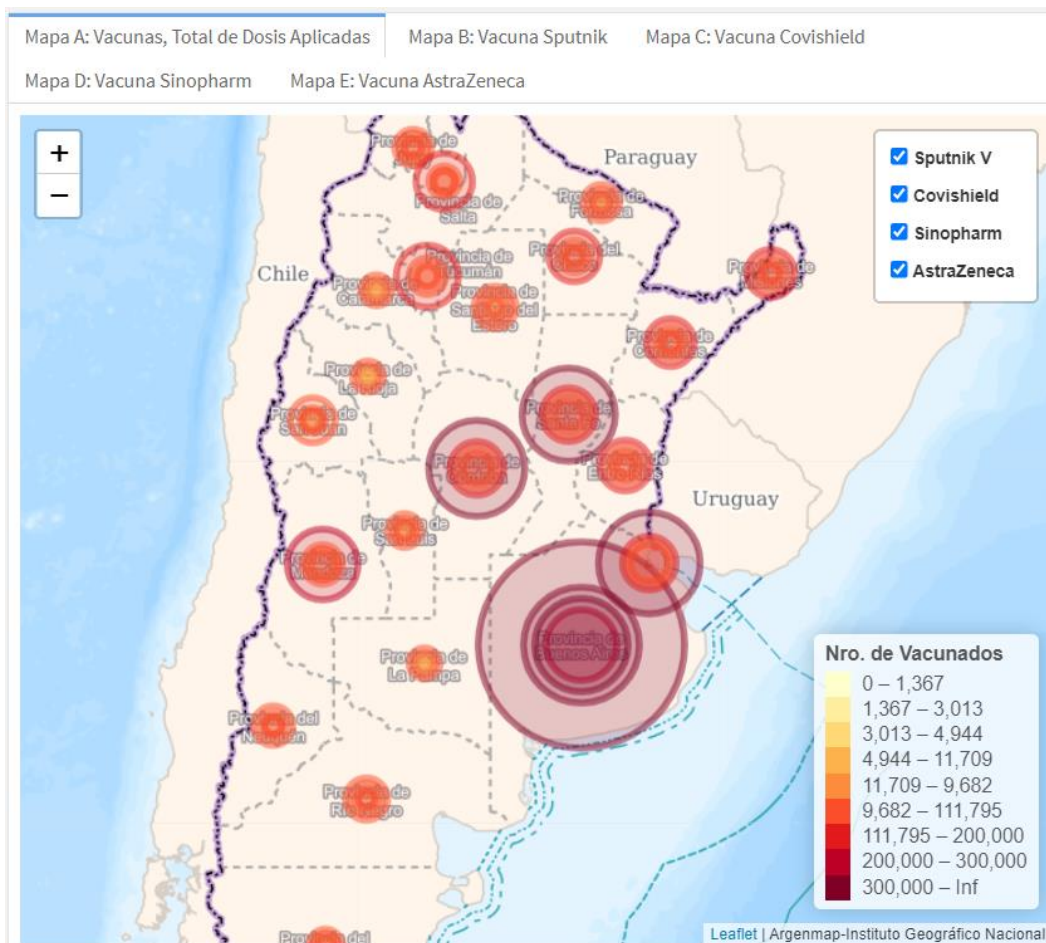


Figura 1: Cantidad de Vacunados por Provincia

A medida que el plan de vacunación aumenta es necesario realizar ajustes en la fórmula para mejorar la visualización del plan de vacunación. Lo mismo sucedió durante el año 2021 cuando en distintos meses se fueron incorporando distintas vacunas en el plan de vacunación.

Se construyó un mapa principal o general en el cual se pueden superponer los marcadores de distintas vacunas seleccionadas a partir de un check-botton. Y por otro lado un conjunto de solapas permite visualizar un mapa del país para cada vacuna.

El reporte del Tablero que se puede acceder a partir del menú superior a la derecha incluye una explicación del tablero. El apartado Cantidad de Vacunados del reporte incluye una explicación similar a la incluida en esta sección y cuatro tablas de datos, uno por cada vacuna en el cual cada tabla contiene cinco

columnas: el nombre de la jurisdicción, el nombre técnico de la vacuna, la cantidad de primera dosis, la cantidad de segunda dosis, y la suma de primera y segunda dosis, similar al tabla mostrada en figura 2

Cantidad de Dosis Sputnik por Jurisdicción

Jurisdiccion	vacuna_nombre	primera_dosis_cantidad	segunda_dosis_cantidad	cantidadTotal
CABA	Sputnik V COVID19 Instituto Gamaleya	521158	112141	633299
Neuquén	Sputnik V COVID19 Instituto Gamaleya	79243	17879	97122
San Luis	Sputnik V COVID19 Instituto Gamaleya	61269	13880	75149
Santa Fe	Sputnik V COVID19 Instituto Gamaleya	438712	76543	515255
La Rioja	Sputnik V COVID19 Instituto Gamaleya	45441	9711	55152
Catamarca	Sputnik V COVID19 Instituto Gamaleya	47710	10383	58093
Tucumán	Sputnik V COVID19 Instituto Gamaleya	195986	35737	231723
Chaco	Sputnik V COVID19 Instituto Gamaleya	124239	25024	149263
Formosa	Sputnik V COVID19 Instituto Gamaleya	64124	9019	73143
Santa Cruz	Sputnik V COVID19 Instituto Gamaleya	41932	8858	50790
Chubut	Sputnik V COVID19 Instituto Gamaleya	62214	11792	74006
Mendoza	Sputnik V COVID19 Instituto Gamaleya	240121	42518	282639
Córdoba	Sputnik V COVID19 Instituto Gamaleya	450061	71245	521306
Entre Ríos	Sputnik V COVID19 Instituto Gamaleya	134683	28551	163234
San Juan	Sputnik V COVID19 Instituto Gamaleya	96545	13239	109784
Jujuy	Sputnik V COVID19 Instituto Gamaleya	70508	17178	87686
Santiago del Estero	Sputnik V COVID19 Instituto Gamaleya	94601	16711	111312
Río Negro	Sputnik V COVID19 Instituto Gamaleya	85193	16967	102160
Buenos Aires	Sputnik V COVID19 Instituto Gamaleya	2158005	314293	2472298
Corrientes	Sputnik V COVID19 Instituto Gamaleya	118094	15812	133906
Misiones	Sputnik V COVID19 Instituto Gamaleya	116188	19076	135264
Salta	Sputnik V COVID19 Instituto Gamaleya	152986	26647	179633
La Pampa	Sputnik V COVID19 Instituto Gamaleya	45473	9005	54478
Tierra del Fuego	Sputnik V COVID19 Instituto Gamaleya	20255	4491	24746

Figura 2: Ejemplo de Tabla de Cantidad de dosis (primera y segunda) por cada vacuna incluida en el reporte del tablero.

4. CONFIRMADOS

Se construyeron distintos mapas sobre cantidad de confirmados. Ellos son:

1. Mapa de cantidad de casos confirmados por provincia.
2. Mapa de tasa de casos confirmados por 100.000 habitantes.
3. Mapa de confirmados por Departamento.

4.1 CONFIRMADOS POR JURISDICCIÓN

El Mapa de Confirmados por provincia reúne las siguientes características:

1. Muestra un círculo por provincia, en el que su radio está en proporción con la cantidad total de confirmados.
2. El centro del círculo coincide con el baricentro de la parcela.
3. Al hacer clic en el círculo se muestra información sobre el nombre de la provincia seguido de la cantidad total de casos confirmados al día del reporte.
4. El color del círculo está categorizado utilizando una paleta de colores de amarillo a rojo, según los siguientes rangos:
 - cantidad de confirmados < 500
 - cantidad de confirmados ≥ 500 y < 1000
 - cantidad de confirmados ≥ 1000 y < 10000
 - cantidad de confirmados ≥ 10000 y < 50000
 - cantidad de confirmados ≥ 50000 y < 100000
 - cantidad de confirmados ≥ 100000 y < 150000
 - cantidad de confirmados ≥ 150000

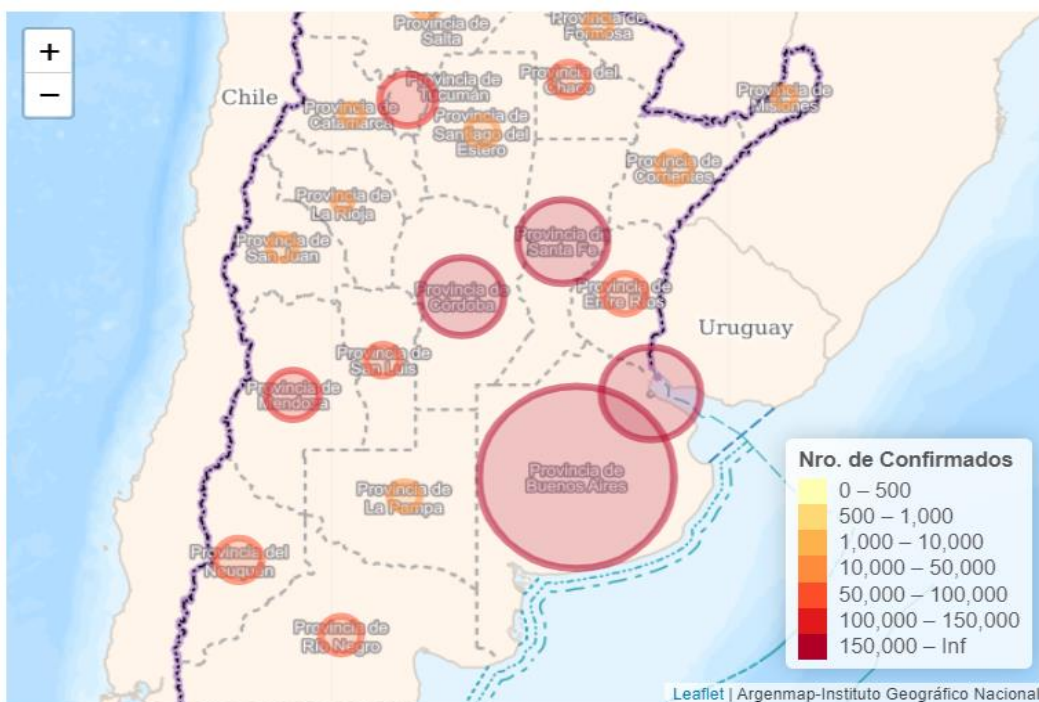


Figura 3: Número de Confirmados por Jurisdicción.

4.2 INCIDENCIAS: CONFIRMADOS POR 100.000 HABITANTES

El Mapa de tasa de confirmados por 100.000 habitantes reúne las siguientes características:

1. Muestra un círculo por provincia, en el que su radio está en proporción con la cantidad de confirmados.
2. El color del círculo esta categorizado según rangos.
3. Al pasar con el mouse en el círculo se muestra una etiqueta con información sobre el nombre de la provincia seguido de la cantidad tasa de casos confirmados por 100.000 habitantes, al día del reporte.
4. Al hacer click en el círculo se muestra un popup con la siguiente información:
 - Nombre de la provincia
 - Tasa de casos confirmados por 100 mil habitantes
 - Cantidad de casos total para la provincia
 - Población de referencia
5. El cálculo de la tasa se obtiene aplicando la siguiente fórmula: $T=(NT/PT)*100000$ $T=(NT/PT)*100000$ Donde:
 - T: es la tasa por 100 mil habitantes;
 - NT: es el número total de eventos de interés;
 - PT: es la población total a mitad de periodo.

4.3 CONFIRMADOS POR DEPARTAMENTO

Para la elaboración de este mapa se utilizó el shp de departamentos del IGN, disponible como dato abierto en (IGN, 2021b). El mapa de departamentos se puede descargar en distintos formatos desde la página del IGN (shp, kml, geojson, csv) e incluso se puede descargar un pdf que describe sus metadatos.

Para muchos casos confirmados, el dataset no dispone de información del departamento de procedencia, por ello en el extremo superior derecho del mapa similar al siguiente del porta del Ministerio de Salud se indica: *Los casos representados en este mapa son sólo aquellos en los que el departamento de residencia fue correctamente consignado en el Sistema Nacional de Vigilancia, por lo que la cantidad de casos mapeados puede ser menor al total de casos confirmados.*

La siguiente tabla describe para cada provincia, la cantidad de casos confirmados COVID en los cuales no se registró información de cual es el departamento asociado (esto es, muchas observaciones tienen un valor nulo en el campo departamento):

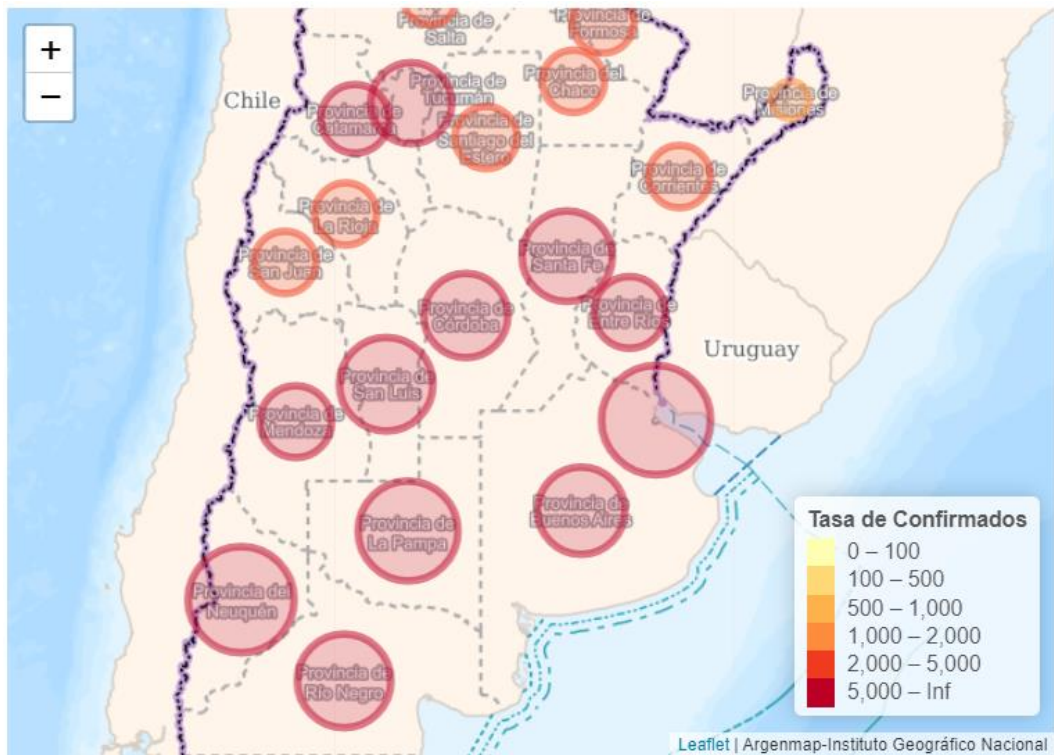


Figura 4: Tasa de confirmados cada 100.000 habitantes.

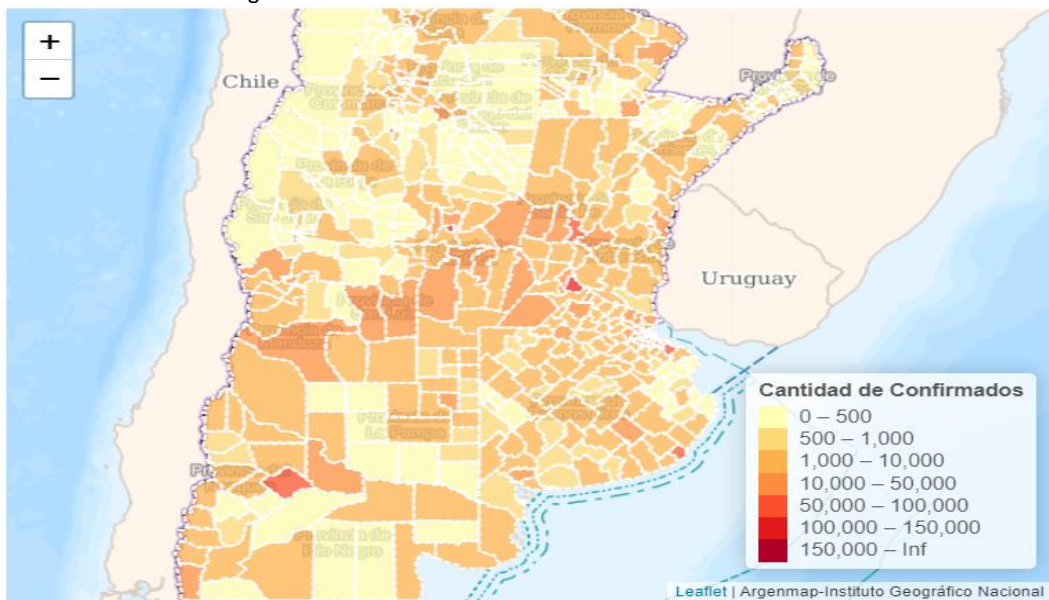


Figura 5 Cantidad de Confirmados por Departamento

Todos los datos mostrados en el mapa son tomados del dataset del Ministerio de Salud. Una observación importante es que en el caso de CABA los valores mostrados no son informados por comunas, sino que muchos de ellos son datos agregados a nivel ciudad. Por esta razón las cantidades para cada comuna pueden no ser relevantes. Pero se optó por mostrar la información provista por el dataset entendiendo que en algún momento esta información será normalizada.

5. GRAFICOS

Se incluyó un gráfico de barras muestra cantidad de casos confirmados por provincia, ordenado de mayor a menor.

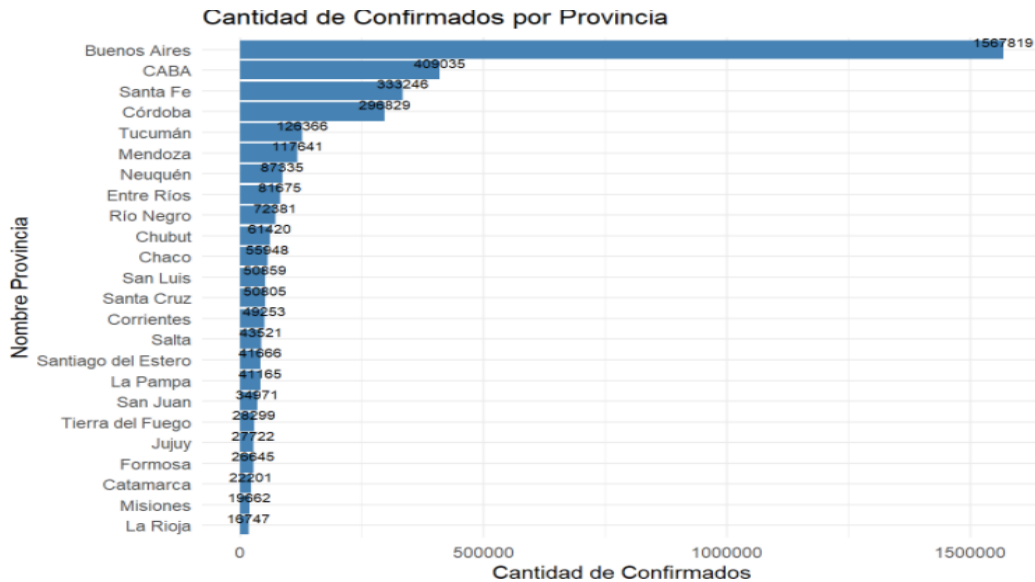


Figura 6: Casos confirmados por sexo: La categoría OTROS refiere a casos donde no se ha registrado el dato de sexo.

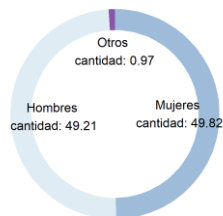


Figura 7: Confirmados por Grupo Etario: Se agrupan los casos confirmados en rangos según su edad.

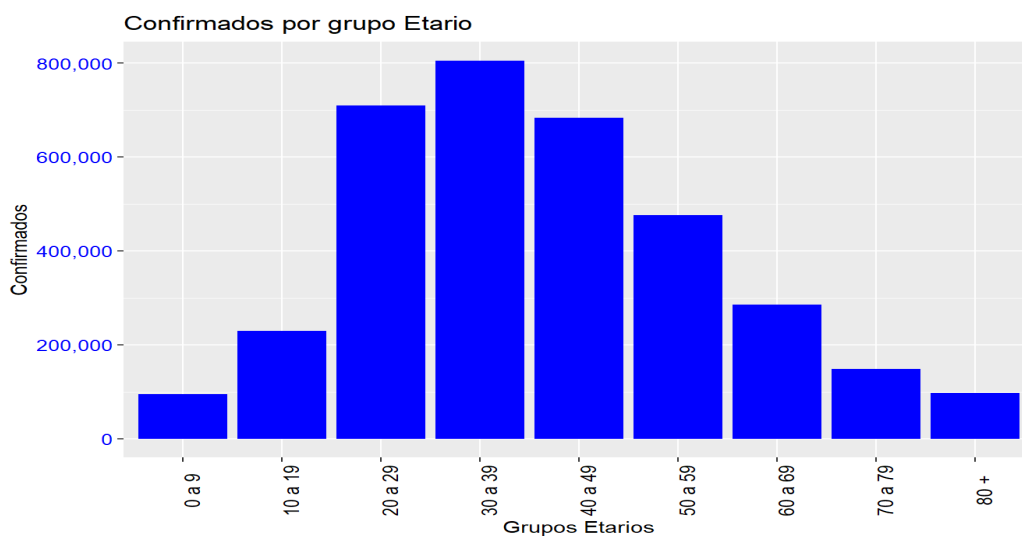


Figura 8: Confirmados por Grupo Etario.

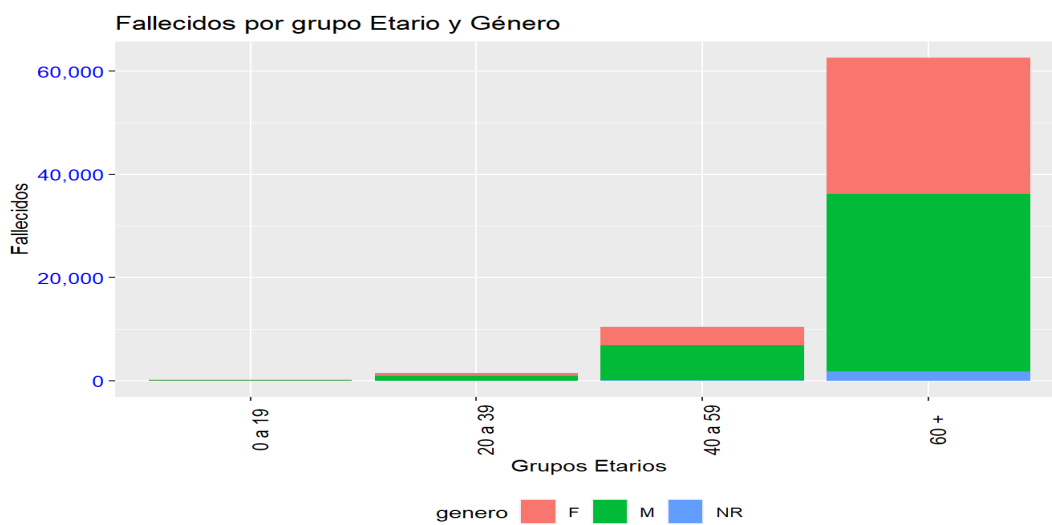


Figura 9: Fallecidos por Grupo y Genero

6. TECNOLOGIA UTILIZADA

El tablero UNCovid esta publicado en la siguiente página web:

<https://opendata.fi.uncoma.edu.ar/TableroUNCovid.html>

Y el reporte del tablero en la siguiente dirección:

<https://opendata.fi.uncoma.edu.ar/UNCovid.html>

El tablero, desarrollado utilizando lenguaje R, incluyó la instalación de los siguientes componentes: leaflet, flexdashboard, readr, tidyverse, dplyr, ggplot2, sp, sf y scales. El servidor web es un servidor Linux, instalado en una máquina virtual en el Facultad de Informática de la Universidad Nacional del Comahue.

7. CONCLUSIONES

En el tratamiento y uso de datos abiertos, no solo intervienen proveedores, que posibilitan la creación de datos, sino que se identifican varios actores involucrados en el proceso. Posteriormente a la creación de datos abiertos, los denominados “agregadores” recopilan y agregan los datos. Berends et al. (2020) identifican dos arquetipos que utilizan los datos en su trabajo: desarrolladores y enriquecedores. Los arquetipos utilizan los datos para el desarrollo de nuevas aplicaciones (desarrolladores) o para obtener nuevas y mejores conocimientos del análisis de los datos (enriquecimiento). De este modo la reutilización de datos abiertos y estándares OGC en la producción de nueva información es esencial en la cadena de valor en la producción de información.

En esta ponencia se describió un ejemplo en el cual fue posible construir un tablero COVID, a partir del uso de datos abiertos y estándares OGC (el WMTS como ArgenMap), permitiendo visibilizar el progreso de inoculación y casos confirmados de COVID, en distintos niveles de jurisdiccionales (provincias y departamentos). La construcción del mismo ha sido progresiva, incorporando nuevos mapas de vacunación a medida que el país incorporó nuevas vacunas, y se publicaban nuevos datasets. A modo de trabajo futuro, se planifica mantener el tablero, incorporar el inicio de inoculación con nuevas vacunas (ej. Moderna) y desarrollar estadísticos temporales, preservando información de vacunación y cantidad de casos confirmados por día o por semana, y mostrar los mismos a partir de gráficos interactivos. Se prevee la realización de estadísticas más precisas y propias de nuestra región o provincias en conjunto con la Facultad de Medicina de la UNCo.

8. AGRADECIMIENTOS

La presente publicación se realiza en el marco del proyecto 04/F010: “Visualización de Datos y Realidad Virtual” y del proyecto “Mapas web & Datos Abiertos” (opendata.fi.uncoma.edu.ar) de la Facultad de Informática de la Universidad Nacional del Comahue,

9. REFERENCIAS

Instituto Geográfico Nacional (2021a). Argenmap mapa base. <https://www.ign.gob.ar/AreaServicios/Argenmap/Introduccion> [accedido 1 de Junio de 2021].

Instituto Geográfico Nacional (2021b). Datos abiertos del IGN.
<https://www.ign.gob.ar/NuestrasActividades/InformacionGeoespacial/CapasSIG>
[accedido 1 de Junio de 2001].

Ministerio de Salud, Argentina (2021a). Vacunas contra COVID-19. Dosis Aplicadas en la República Argentina. Datos Abiertos del Ministerio de Salud. Dataset: <http://datos.salud.gob.ar/dataset/vacunas-contra-covid-19-dosis-aplicadas-en-la-republica-argentina> [accedido 1 de Junio de 2001].

Ministerio de Salud de la Nación, Argentina (2021b). COVID-19. Casos registrados en la República Argentina. Datos Abiertos del Ministerio de Salud. Dataset: <http://datos.salud.gob.ar/dataset/covid-19-casos-registrados-en-la-republica-argentina> [accedido 20 de Junio de 2001].

Ángeles Navarro (2016) ¿Cómo ayudar a los ciudadanos a comprender los datos abiertos? <https://blogs.iadb.org/conocimiento-abierto/es/como-ayudar-a-los-ciudadanos-a-comprender-los-datos-abiertos/> [accedido 20 de Junio de 2001].

Berends, J., Carrara, W., Engbers, W., Vollers, H. (2020) Re-using Open Data. A Study o Companies transforming Open Data into Economic & Social Value. European Data Portal, European Commission. ISBN: 978-92-78-41872-4 https://data.europa.eu/sites/default/files/re-using_open_data.pdf [accedido 20 de Junio de 2001].