

Ciencia de Datos para la Salud Pública: apuntes de una conversación en desarrollo

RENÉ LAGOS BARRIOS, ANDRÉS GONZÁLEZ-SANTA CRUZ, FELIPE A. MEDINA MARÍN,
SANDRA FLORES-ALVARADO, AMARU AGÜERO JIMÉNEZ

GRUPO DE CIENCIA DE DATOS Y SALUD PÚBLICA, ESCUELA DE SALUD PÚBLICA, UNIVERSIDAD DE CHILE*

Por iniciativa del Grupo de Ciencia de Datos y Salud Pública (CDSP) se convocó a un conjunto de estudiantes, profesionales y académicos/as que se desempeñan en esta área, quienes se reunieron el pasado jueves 17 de agosto en la Escuela de Salud Pública. El objetivo fue definir y elaborar las líneas temáticas de las 1^{as} Jornadas de Ciencia de Datos y Salud Pública a realizarse en enero de 2024 (ver el listado de asistentes en los agradecimientos).

La discusión se llevó a cabo en tres mesas, en las que se discutieron: 1) los alcances de cada línea, 2) los actores involucrados y 3) el rol de la academia, el gobierno y las empresas. Al final de la sesión se realizó una plenaria en donde se expusieron y discutieron los principales puntos abordados en cada mesa. A continuación se presenta un resumen de la discusión en cada una de ellas.

Mesa 1: Soluciones de ciencia de datos para la salud pública

1) ¿Cuáles son los alcances de la línea temática?

Se identificaron tres ejes que delimitan el alcance de esta línea. Como primer eje se abordaron las posibles soluciones para la captación y procesamiento de datos. Lo anterior, considerando que uno de los principales propósitos tras la aplicación de la ciencia de datos a la salud pública proviene de la necesidad de asegurar la calidad e integridad de los datos de salud, frente a su reciente proliferación. Para esto se hace necesario elaborar y mantener *catálogos abiertos de bases de datos* actualizados y aptos para su uso en salud pública. Por otra parte, se requiere la definición y aplicación de *estándares y arquitecturas* para el manejo de tales datos. Se identificó la *interoperabilidad* de los sistemas de información, tanto de salud como de los distintos organismos del Estado, como un atributo indispensable para poder compartirlos. Además se requieren arquitecturas que permitan integrar datos intersectoriales con datos de salud locales para que las soluciones sean adecuadas para cada territorio. Este eje también considera los nuevos sistemas de registro y captación de datos a partir del uso de plataformas web, dispositivos móviles y redes sociales.

El segundo eje fue el modelamiento de datos para la salud pública. Éste comprende los modelos que abordan la salud desde el enfoque de los determi-

tes sociales de la salud, incluyendo el modelamiento de los fenómenos epidemiológicos, así como de la oferta del sistema de salud y de los determinantes sociales de la salud. Considera también la implementación de modelos que sirvan para la prevención y la promoción de la salud; así como modelos predictivos que busquen un tratamiento integral de las personas enfermas, es decir, desde el enfoque biopsicosocial. Resultan de particular interés los modelos que incorporen datos geoespaciales o de texto libre por su capacidad para aportar nueva información.

Finalmente, como tercer eje de esta línea se encuentran las prácticas de visualización, interpretación y comunicación de datos, que permitan acercar los datos y la salud pública a las personas. En particular, las formas en que la ciencia de datos sea centrada en las personas para que pueda apoyar la comunicación de riesgo a la ciudadanía. En este contexto, las y los participantes destacaron la necesidad de conocer e identificar los visores de datos que se construyen en las instituciones de salud, muchos de los cuales son descargables.

2) ¿Quiénes son los actores relevantes involucrados?

En primer lugar, el Ministerio de Salud y las reparticiones y organismos que dependen de él. En particular el Departamento de Epidemiología y el de Salud Digital que impulsan políticas de salud y de digitalización en salud, así como el Instituto de Salud Pública (ISP) que genera políticas regulatorias de bienes y servicios en salud. Por otra parte, son relevantes otras instituciones del Estado que proveen datos indispensables para salud pública, como el Instituto Nacional de Estadística (INE), el Ministerio de Desarrollo Social y Familia (principalmente a través de la encuesta CAsEN), el Ministerio del Medio Ambiente, el Instituto Meteorológico y, finalmente, el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, que impulsa y financia proyectos de ciencia de datos para la salud pública a través de la Agencia Nacional de Investigación y Desarrollo (ANID), y la Corporación de Fomento de la Producción (CORFO).

Por otra parte, se encuentran los centros de investigación que trabajan en aplicaciones de la ciencia de datos a problemas de salud pública como, por ejemplo, el Centro Nacional en Sistemas de Información en Sa-

* cienciadatos.med@uchile.cl

lud (CENS), el Centro Nacional de Inteligencia Artificial (CENIA), el Instituto Milenio en Ingeniería e Inteligencia Artificial para la Salud (I-Health), el Centro de Microdatos (CMD), el Instituto Milenio Fundamentos de los Datos (IMFD), el Centro de Modelamiento Matemático (CMM) y el Instituto de Sistemas Complejos de Ingeniería (ISCI). Asimismo, participan las empresas de registros clínicos electrónicos y tecnologías de información y comunicación en salud.

3) ¿Cuál es el rol de la academia, el gobierno y el sector privado?

Para desarrollar el campo de ciencia de datos para la salud pública, se requiere que los organismos públicos puedan compartir datos fidedignos tanto de la salud como de los contextos socioeconómicos, educacionales y medioambientales de la población. La academia debe colaborar en mapear las fuentes de datos relevantes para la salud pública y definir estándares científicos de los datos y aplicaciones que aseguren calidad de los datos y su uso en las aplicaciones para la salud pública. El sector privado, en tanto, tiene el rol de impulsar el desarrollo de soluciones de ciencia de datos, experiencia que también es importante conocer y relevar.

Mesa 2: Ética y gobernanza de datos para la salud pública

1) ¿Cuáles son los alcances de la línea temática?

Abarca aspectos legislativos (por ejemplo, marcos normativos paradigmáticos como el europeo y el estadounidense) e institucionales (comités de ética de la investigación), pero les trasciende al abarcar dimensiones prácticas y colectivas, que se expresan en la forma de abordar desafíos y hacernos cargo de nuestras capacidades como sociedad, pero también de los riesgos a los que nos exponemos. Este aspecto fue relevado principalmente en relación a la naturaleza sensible que tienen los datos que comprometen la salud. La mala utilización plantea un problema razonable de “seguridad crítica”, o como un mecanismo que permitiría profundizar desigualdades (ej., usuarios del sistema público vs. privado) o estigmatizar a grupos de personas (ej., con VIH). Por otra parte, adquieren una materialidad, una dimensión concreta, la cual se ve reflejada en una infraestructura y marco jurídico, destinados a hacerse cargo de estos problemas. Se reconoce una complejidad de los alcances éticos, ya que al solo delimitarla a un área ética-judicial, se deja de lado el rol de involucrar a la sociedad civil en la responsabilidad social de la gobernanza, como un ideal de estado (por ejemplo, levantando datos que son valorados por los pacientes y comunidades, o promoviendo la “apropiación” tecnológica en un sentido amplio). Por último, un elemento distintivo de esta dimensión es la importancia que se le otorga a socializar y a co-construir un propósito de las iniciativas y políticas públicas por parte del Estado

(por ejemplo, de apertura de datos). El Estado debería tener claridad cuáles son las políticas públicas que quiere generar, encaminar y facilitar el uso responsable de los datos para la salud pública.

2) ¿Quiénes son los actores relevantes involucrados?

El Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, en específico la ANID. La Secretaría General de la Presidencia y su División de Gobierno Digital (que pasará al Ministerio de Hacienda). El Ministerio de Salud, particularmente la Oficina Nacional de Bioética y los centros de salud -quienes levantan los datos- y el Departamento de Estadísticas e Información de Salud (DEIS), quienes los sistematizan y que a su vez se han visto limitados en su facultad de publicar algunos datos. El Ministerio del Medio Ambiente y la Subsecretaría de Telecomunicaciones son mencionados en relación con la medición del rango de señal comunicacionales y la cercanía con hospitales, como ejemplo de infraestructura y la materialidad del dato en salud pública. La integración y confluencia en la gobernanza de datos de organizaciones con distinta estructura, dependencia administrativa y naturaleza jurídica, puede ser problemática. Finalmente, se destaca la importancia de que la sociedad civil comprenda el uso de los datos y su propósito, y se promueva el desarrollo de una reflexividad crítica respecto a ellos, más que un mero uso de las tecnologías, entre ellas, las que involucran la salud.

3) ¿Cuál es el rol de la academia, el gobierno y el sector privado?

La academia es la responsable de generar investigación y conocimiento sobre las mejores prácticas y estándares en ética y gobernanza de datos. Igual de importante es la educación y capacitación de profesionales y estudiantes en ética, privacidad y manejo de datos en salud, promoviendo la alfabetización en datos entre la comunidad. La academia proporciona un espacio para innovación y explorar nuevas técnicas, herramientas y tecnologías para mejorar la gestión de datos y la seguridad. Finalmente, la academia debe poner énfasis en la calidad y el estándar ético y jurídico, en determinados casos (por ejemplo, en la recopilación de datos antropométricos por parte dispositivos móviles). Junto con ello, impulsar la reflexión académica, considerando marcos normativos, estándares de comparación/*benchmark* (OECD), posiciones e interdisciplina. De igual manera, debe adoptar un espíritu crítico hacia la legislación y las prácticas actuales. Para ello, se plantea necesario rescatar a los comités de ética en las discusiones sobre las implicancias éticas de las investigaciones que involucran el trabajo con datos que comprometen la salud.

El gobierno es el responsable de la gobernanza de datos y de decidir qué datos se hacen públicos y cuáles no, lo que debiese hacer en función de una finalidad.

Además, es el responsable de establecer la regulación y el marco normativo. Debe ejercer supervisión y considerar la transparencia y la seguridad en la gestión de datos debido a su carácter sensible, confidencial y a su vez empoderador (por ejemplo, no sólo usar la tecnología, sino reconocer sus riesgos como también para lo que le favorece). El gobierno realiza promoción y apoya y financia iniciativas de salud pública y proyectos que involucran la ética y gobernanza de datos. De igual forma, tiene el deber de retribuir a la población que proporcionó los datos. De ello se desprende que es el responsable de la gobernanza de datos y de decidir qué datos se hacen públicos y cuáles no, lo que debiese hacer en función de una finalidad.

El sector privado juega un papel importante en la gobernanza de datos. Desarrolla y ofrece soluciones tecnológicas y herramientas para el almacenamiento, análisis y protección de datos. El sector privado puede proveer financiamiento e invertir en investigación y desarrollo en áreas relevantes para la protección de datos. Se centra en la colaboración y trabajo en conjunto con el sector público y la academia en el marco de políticas de innovación tecnológica. Se reconoce que existe riesgo de profundizar desigualdades a través del uso de datos y restringir su acceso para su aprovechamiento público, pero la colaboración academia-gobierno-privado debe apuntar a minimizar estos riesgos.

Mesa 3: Cultura de transformación digital para la salud pública

1) ¿Cuál es el alcance de la línea?

El alcance de la línea es amplio, transversal y abarca diversos aspectos que van más allá del mero registro de datos. Por un lado, implica cambios en la forma en que fundamentamos la toma de decisiones, por ejemplo, reemplazando la intuición por información basada en datos concretos. Esto, a su vez requiere que haya cambios en la accesibilidad a registros de datos de buena calidad. Es posible que la falta de digitalización o un buen registro de los datos en la actualidad se deba a que ambos se entienden de forma limitada como la simple recopilación de información. Lo anterior está lejos de la realidad, ya que la accesibilidad a registros de buena calidad es lo que da valor a los datos en su uso en múltiples contextos, desde la investigación hasta la toma de decisiones, tanto por instituciones y empresas como por individuos de la sociedad civil. Para cambiar la actual y generalizada concepción limitada sobre la importancia del acceso y la calidad de los datos, esta línea busca fomentar una mentalidad que promueva la eficiencia en la captura de datos (“una buena cultura de registro”) y su uso, mediante la educación y comunicación de la importancia de la interdisciplinariedad y cooperación de todas las partes involucradas en la transformación digital.

2) ¿Quiénes son los actores relevantes involucrados?

Los actores relevantes involucrados en la cultura de transformación digital son varios. En la escala individual encontramos a las personas que son la fuente de información para quienes la registran y la convierten en datos. Luego, en una escala superior, están las instituciones y empresas que desempeñan un papel esencial en la generación, recopilación, almacenamiento y uso de datos. Entre las empresas destacan las empresas tecnológicas que proporcionan herramientas y tecnologías para la captura, gestión y análisis de datos. El Gobierno está presente en el establecimiento de regulaciones y estándares, que deben promover la transparencia y uso adecuado de los datos. Otro actor importante es la academia, la cual contribuye con la investigación, la educación y la promoción de mejores prácticas en la transformación digital. La sociedad civil desempeña un rol en la concientización, la protección de los derechos de los individuos y la supervisión respecto al uso de los datos.

3) ¿Cuál debería ser el rol de la academia, gobierno y empresas?

La academia debería fomentar la colaboración interdisciplinaria y enseñar la importancia de abordar problemas en equipo mediante el uso de datos. Además, debería fomentar la difusión efectiva de sus resultados. También podría ayudar a estandarizar procesos y facilitar la transferencia de conocimiento desde la investigación hacia su aplicación práctica, prestando además asesoría tanto a instituciones como empresas y gobierno. El gobierno debería establecer regulaciones claras para proteger la privacidad y seguridad de los datos, así como promover estándares de calidad en la recopilación y uso de éstos. Con estas regulaciones y estándares, el gobierno puede agilizar la disponibilidad de datos y promover la transparencia en su uso. Por otra parte, las empresas tienen la responsabilidad de ser conscientes de la relevancia de los datos y su impacto en la sociedad. Además, podrían contribuir a la educación y concientización sobre la importancia de la transformación digital. Las empresas tecnológicas, por otro lado, pueden proporcionar herramientas y soluciones para la gestión de datos. Cualquiera de estos actores –academia, gobierno y empresas– deben además buscar formas de beneficiar a las comunidades que son fuente de los datos con información relevante para ellas, logrando con esto mostrar que ser fuente de información les beneficia.

Conclusiones

A través de estas mesas de discusión se evidenciaron una serie de puntos de convergencia que subrayan la importancia de abordar estas líneas temáticas asumiendo sus intersecciones y considerando algunos aspectos clave. Particularmente, se destacó que la calidad

e integridad de los datos, así como su interoperabilidad, son esenciales para desarrollar soluciones efectivas de ciencia de datos en salud pública. Además, se resaltó la importancia de la visualización y comunicación de datos para acercar la información a la ciudadanía. Con esto, la transformación digital no se limita a la captura de datos, sino que implica una mentalidad que promueva la eficiencia en su adquisición y uso, fomentando la colaboración interdisciplinaria y la conciencia sobre la importancia de los datos, bajo un marco ético y de gobernanza sólida. Este último punto resulta fundamental para garantizar un uso adecuado y responsable de la información sensible en salud pública.

La academia, el gobierno y el sector privado desempeñan roles complementarios en estas líneas temáticas, desde la investigación y regulación hasta la provisión de herramientas y soluciones tecnológicas. Asimismo, se destaca una idea común a todas las líneas: la necesidad de garantizar que las personas que proporcionan datos sean retribuidas y protegidas por los actores involucrados en su recolección y uso.

Es importante destacar que estas líneas no deben interpretarse como límites rígidos, sino como áreas de interacción y colaboración. La discusión continúa y confiamos en que estas jornadas sirvan como punto de partida para un diálogo continuo y para la generación de soluciones innovadoras en el campo de la ciencia de datos para la salud pública.

Agradecimientos

Queremos agradecer encarecidamente a quienes aceptaron nuestra invitación y participaron en esta actividad:

Mesa 1: Tomás Bralic, Sandra de la Fuente, Janepsy

Díaz, Karen Oróstica, Rodrigo Salas, María Teresa Valenzuela, Adriana Villa-Murillo (secretaria) y René Lagos (moderador).

Mesa 2: Pablo Godoy, Diego A. Martínez, Gonzalo Reyes, Ivonne Vargas, Cristian Flores (secretario) y Andrés González (moderador).

Mesa 3: Marcela Aguirre, Tania Alfaro, Rodrigo Assar, Mauricio Fuentes, Nicolás Montalva, Paulina Pino, Felipe Medina Marín (secretario) y Cecilia Meneses (moderadora).

Moderadora y coordinadora de las mesas: Sandra Flores Alvarado.

También agradecemos al Comité 80 Años de la Escuela de Salud Pública por el apoyo para realizar esta actividad, en particular a Verónica Zúñiga, Paola Videla y Catherine Sanhueza, por su apoyo en la producción.

Financiamiento

La Agencia Nacional de Investigación y Desarrollo de Chile (ANID) financia los estudios de doctorado de Sandra Flores Alvarado, René Lagos y Andrés González a través de la Beca de Doctorado Nacional (CONICYT-PFCHA/ Doctorado Nacional/ 2020-21200398, Subdirección de Capital Humano/Doctorado Nacional/2022-21220971 y Subdirección de Capital Humano/Doctorado Nacional/2023-21230172). Las 1as Jornadas de Ciencia de Datos y Salud Pública cuentan con el financiamiento del Departamento de Postgrado y Postítulo de la Vicerrectoría de Asuntos Académicos de la Universidad de Chile a través del concurso “Estímulo para proyectos académicos de estudiantes de postgrado de la Universidad de Chile”, convocatoria 2022 – 2023.



Foto grupal de participantes presenciales en la actividad.