

Resumen del Taller

Herramientas de análisis de información para la respuesta a crisis humanitarias en salud

Introducción

Ante la continua evolución de tecnologías de la información, y el aumento de eventos que amenazan poblaciones humanas como guerras, anomalías climáticas o crisis económicas, Crisis Ready, en asociación con la Universidad del Rosario y la Universidad Nacional de Colombia organizó un taller para discutir con los tomadores de decisiones, la importancia de la calidad, disponibilidad y uso oportuno de los datos para responder a emergencias humanitarias. En el taller participaron organizaciones multilaterales que realizan trabajo de campo como OIM, PNUD, OPS, Cruz Roja así como ONGs, grupos académicos y empresas privadas que realizan trabajo de análisis de datos para contribuir a la toma de decisiones con base en evidencia en contextos de crisis humanitarias con énfasis en salud. En especial organizaciones como iMMAP, Procalculo, entidades académicas y grupos que llevaron a cabo análisis de datos durante la pandemia de Covid-19 para entidades gubernamentales a nivel nacional y regional. Asimismo, participaron varias entidades del gobierno, como el Instituto Nacional de Salud, el Departamento Nacional de Estadística, entre otros. Durante los dos días del taller se discutieron aspectos relacionados con la coyuntura de crisis humanitaria desde la analítica de datos.

- Los objetivos del taller han sido:
 - Promover la creación de una red de actores importantes en atención a crisis humanitarias.
 - Identificar los diferentes momentos de una crisis, y el papel que puede cumplir la analítica de datos en cada potencial escenario.
 - Asociar la red de actores a las necesidades de información, análisis y distribución de datos, y herramientas para generar resiliencia frente a crisis humanitarias.

Panel 1: Uso y naturaleza de los datos

Independientemente de los datos que estén siendo usados, para poder responder preguntas sobre la representatividad, resolución, posibles sesgos, peligros en términos de privacidad y uso potencial, es fundamental tener en mente dos preguntas:

- (1) ¿De dónde provienen los datos?
- (2) ¿Cómo están siendo agregados?

¿De dónde provienen los datos?

Esta pregunta encapsula tanto el tipo de dato que se está capturando (ej. imágenes satelitales, datos de telefonía móvil, etc.), como las características a diferentes niveles de resolución o granularidad.

Si tomamos el ejemplo de datos de telefonía móvil que expuso Nishant Kishore (Harvard University), el sólo análisis de cómo se recolectan estos datos permite inferir la representatividad de estos y sus posibles sesgos. Por naturaleza, los datos que se recolectan usando anuncios publicitarios en aplicaciones (e.g. Rappi) son la representación de sus usuarios y acarrean consigo los sesgos que estos usuarios representan (e.g. ¿Qué subpoblación usa Rappi con más frecuencia?). Esta característica no es única de los datos obtenidos a partir de aplicaciones o telefonía móvil. Si pensamos en datos de sensores remotos (*Remote Sensing*) tales como satélites o drones (Hamish Patten, Oxford University), la forma en que se capturan estos datos también puede proporcionar indicios acerca de su representatividad. En este caso la representatividad puede entenderse en términos de resolución a diferentes niveles (espacial, temporal, espectral y radiométrica), pues inevitablemente al muestrear una área determinada, se están haciendo sacrificios en términos del objeto del muestreo y su frecuencia. Por ejemplo, los datos satelitales son notorios por tener mayor cantidad de fotos (ej. mejor resolución temporal) en lugares más desarrollados, mientras que regiones como Colombia sufren de problemas de nubosidad que indirectamente afectan el grado de representación de un área determinada.

La representatividad de los datos implica por lo tanto sesgos que deben ser tomados en cuenta al informar la manera en la que se usan estos datos.

¿Cómo están siendo agregados?

Esta pregunta aunque fundamental, no siempre es fácil de responder. En el caso de datos de CDR (Call Detailed Record), donde se registra la localización de la torre celular más usada por algún dispositivo durante el transcurso de un día, no tener claro el proceso de agregación nos puede llevar a conclusiones erradas (Taylor Chin, Harvard University). Por ejemplo, el hecho de que un dispositivo muestre la misma geolocalización en dos días consecutivos no implica que el dispositivo no se haya movido, y mucho menos que la persona asociada con aquel dispositivo no se haya movido. Más allá de la conclusión lógica de que un dispositivo puede estar siendo usado por más de una persona, la forma en la que se genera la geocodificación de este tipo de dato (la torre celular más cercana que ha sido *más usada durante un día*), diluye todos los posibles movimientos que un individuo pueda tener en un día particular. Aún más, si un individuo constantemente hace uso de su celular en su casa, día tras día, el sistema de agregación va a ocultar cualquier otro movimiento. Cada sistema de agregación va a derivar en un nivel particular de resolución y esto a su vez va a sugerir lo que es posible hacer con los datos.

Además de su importancia al momento de sacar conclusiones de los datos, los sistemas de agregación ayudan a responder preguntas sobre la privacidad de los datos. Aunque comúnmente se asocia agregación con anonimidad, la agregación frecuentemente no es suficiente para garantizar la privacidad. Para ello es necesario recurrir a marcos de trabajo como lo es la privacidad diferencial (Nishant Kishore, Harvard University). Sistemas de agregación inteligentes son fundamentales para poder extraer la mayor cantidad de información posible sin comprometer la privacidad. Agregar a lo largo de líneas de marcadores de equidad (e.g. género) permite hacer análisis más granulares que puedan en un futuro informar políticas con enfoques diferenciados (e.g. enfoque de género) (Andrew Schroeder, Crisis Ready).

Panel 2: Migración y desplazamiento

En este panel se tocaron dos temas fundamentales. Por un lado se habló acerca de cómo se pueden usar los datos para lidiar con crisis de migración y desplazamiento, y por otro se habló de las infraestructuras de talento humano necesarias para apoyar este tipo de sistemas de información. Acá se empieza a ver la distribución de los distintos roles a lo largo del espectro de actores (academia, industria y tomadores de decisión) y los enormes retos que existen para conectarlos.

Andrew Schroeder (Crisis Ready) mostró el trabajo realizado en Ucrania para intentar complementar los datos de migración oficiales. Usando datos de Meta fue posible tener una idea general de las ciudades donde hubo mayor incremento poblacional y así intentar acomodar los recursos necesarios para sostener el flujo. Este uso de datos para la planeación también pudo observarse en el ejemplo que nos trajo UNHCR: en Brasil se están usando datos de la Fundación Cáritas para planear en tiempo “casi-real” las capacidades de los refugios disponibles para migrantes venezolanos. Estos modelos pueden ser complementados con señales adicionales para expandir el tipo de conclusiones que se pueden obtener. En el caso de los refugiados provenientes de Ucrania, Crisis Ready está enriqueciendo la información de crecimiento poblacional con información sobre vínculos (conectividad social inferida por amistades en Facebook) para intentar predecir no sólo el destino de los refugiados, sino el lugar más probable de permanencia. Aunque en estos ejemplos los datos están fluyendo desde las poblaciones vulnerables hacia las organizaciones civiles, es posible encontrar flujos en dirección opuesta. La aplicación de GIFMM “GIFMM contigo” ilustra este ejemplo. Esta aplicación busca consolidar la oferta de servicios de apoyo a refugiados y migrantes para proveer información a los usuarios (en este caso, migrantes venezolanos). Es importante recalcar que aunque la privacidad es fundamental en cualquier ejercicio de uso de datos, cuando hay captura de datos directamente en los dispositivos celulares de poblaciones vulnerables, estas preocupaciones son de particular importancia.

La segunda línea de discusión durante este panel fue sobre la necesidad de infraestructuras robustas para apoyar la toma de decisiones basadas en datos. Diana Gualteros (Instituto Nacional de Salud, Colombia) hizo un recorrido sobre el funcionamiento básico del sistema de vigilancia

desarrollado por el Instituto Nacional de Salud, empezando por el monitoreo de señales de 108 eventos de salud pública, a través de un sistema de alerta temprana que se evalúa en la sala de análisis de riesgo para decidir si es necesario activar los equipos de respuesta inmediata. Este proceso recae sobre los hombros de personal capacitado, por lo cual el sistema de formación del INS es igualmente importante. Por otro lado, como lo recordó Humberto Mendoza (Secretaría de Salud de Barranquilla), es fundamental poder garantizar la continuidad del talento humano a lo largo de los cambios de gobierno.

Panel 3: Uso de datos en situaciones de emergencia y transmisión de enfermedades

Tradicionalmente existen flujos de datos y protocolos determinados que conectan la industria, la academia y el gobierno. Un ejemplo de esto es la sala de análisis de riesgo en Cartagena (Luisana del Carmen Cárcamo Marrugo, DADIS Cartagena) donde cotidianamente se incorporan datos de entradas y salidas aeroportuarias, monitoreo de redes sociales y poblaciones especiales. Además, existen protocolos para involucrar expertos no gubernamentales en caso de que el nivel de riesgo lo amerite. Durante el transcurso de situaciones de emergencia sin embargo, es posible que se establezcan nuevos flujos basados en necesidades particulares. Este fue el caso de Chile, donde el Ministerio de Ciencia y Tecnología puso a la disposición de los académicos los datos de Covid con el propósito expreso de facilitar la investigación al respecto y la integración con el sector salud. Gracias a estos datos fueron posibles investigaciones como la presentada por Pamela Martinez (University of Illinois Urbana Champaign). En ésta investigación se exploraron los factores diferenciales asociados al estrato socio-económico durante la pandemia de COVID-19. Este tipo de investigaciones que esclarecen los vínculos entre factores sociales y diferentes afectaciones en salud son fundamentales para guiar los esfuerzos gubernamentales en la forma de acciones puntuales y enfocadas. Como ejemplo, el equipo de Pamela Martinez logró hacer un ejercicio analítico de los factores socio-económicos responsables de las tasas de contagio y mortalidad en Santiago de Chile (e.g. movilidad, tasas de pruebas, subregistro). Esto permite que los tomadores de decisión concentren sus esfuerzos en mejorar la adquisición de datos y evaluar los factores implicados en el contagio para establecer objetivos adecuados en la asignación de recursos.

Sin embargo, para asegurar el éxito de estrategias de contención y mitigación en momentos de crisis no es suficiente únicamente la incorporación de entidades gubernamentales y la competencia de grupos académicos. El problema de traducción de conocimiento juega un papel central en explicar las fallas a la hora de tomar decisiones basadas en datos. Con relación a este problema Felipe González y Andrea Parra (Universidad Nacional, Universidad del Rosario) proponen una redefinición de la tecnología informática. Tradicionalmente la tecnología informática se usa como sistema de monitoreo. Un ejemplo de esto es el Monitor OCHA que tiene capacidades de recolección, procesamiento y visualización de datos, y es una herramienta de gran utilidad tanto para el público como para los tomadores de decisiones (Silvia Echeverri, OCHA). La

propuesta de Felipe González y Andrea Parra consiste en usar la tecnología informática como motor de diseño. Esto implica crear sistemas portables, escalables y flexibles para facilitar la colaboración entre la academia y el sector público. Este puente da cabida a la co-creación de métricas relevantes para tomadores de decisiones pero basadas en datos y el estado del arte. Un ejemplo de cómo se puede usar la tecnología para crear puentes entre actores proviene de IMMAP. Jeffrey Roberto Villaveces (IMMAP) presentó el caso de estudio del uso de imágenes de detección remota para el monitoreo de nuevos asentamientos. Aunque una parte del proyecto sigue los paradigmas tradicionales (toma de imágenes, entrenamiento y uso de modelos de aprendizaje, verificación y validación en campo y publicación de resultados) una de las ramas que están desarrollando es el mapeo comunitario basado en maquetas de impresión 3D. Este puente entre la comunidad y la investigación facilita los esfuerzos de traducción y asimilación de conocimiento.

Panel 4: Privacidad y proveedores de datos

Durante este panel se discutió el uso de datos en el contexto de, por un lado, el riesgo a vulneraciones de derechos de privacidad y por otro la expectativa de hacer el bien (Carolina Botero, Fundación Karisma). Se compartió evidencia de momentos donde, por muy bienintencionada que sea la recolección de datos, si no se es sumamente cuidadoso con mantener protocolos de privacidad, estos datos pueden dar lugar a abusos de poder a tensiones entre distintos grupos de interés. Como ejemplo concreto, la alcaldía de Bogotá, con el apoyo de grupos académicos, usó datos celulares para hacer modelos de sobredispersión. Estos modelos buscaban mitigar los contagios sin tener que recurrir a clausuras del espacio público o cuarentenas durante la epidemia de COVID. En otras ciudades del país sin embargo, estos datos fueron usados por las autoridades para ir a las casas de personas que se consideraban focos de transmisión en lo que se considera una grave violación a los derechos de privacidad. La peligrosa narrativa de la que nos advierte la organización Karisma es la de “normalizar la idea que los datos deberían estar disponibles siempre y se deberían usar para todo.” Esta es una narrativa que suele primar durante momentos de crisis, ya que como bien lo expresó Edna Margarita Valle (coordinadora GIT Estadísticas Vitales, DANE, Colombia), en momentos de crisis los gobiernos están bajo grandes presiones para servir a la ciudadanía y con muy poca información disponible. Es acá donde es importante recordar que hay una diferencia entre el uso de datos por parte del gobierno y el uso de información por parte de entidades públicas u organizaciones académicas. Uno de los papeles fundamentales de la academia en momentos de crisis es actuar como guardián de la privacidad, y traducir datos sensibles a información accionable y anonimizada que los gobiernos puedan implementar. Para cumplir esta labor, Carlos Ahumada (Meta), propone un paradigma de desarrollo llamado “privacidad como diseño.” La agregación de datos rara vez es suficiente para garantizar la anonimidad, por lo tanto es necesario recurrir a herramientas como la privacidad diferencial o el aprendizaje federado. En la primera, los datos se traducen en una distribución a la que se le agrega ruido y luego se muestrea para preservar la anonimidad sin comprometer la

calidad analítica. En la segunda, se hace entrenamiento de máquina distribuido donde el algoritmo nunca ve la totalidad de los datos, sino solo muestreos locales. Herramientas como estas son importantes para cerrar el ciclo de traducción de datos a política pública, pero son sólo el primer paso. Zulma Cucunubá (Universidad Javeriana, Imperial College) describe este ciclo como un proceso iterativo donde primero se identifica el problema, luego se entra en una etapa de co-desarrollo de métricas y modelos. Se sigue con el uso de datos anonimizados para las validaciones de los modelos y la comunicación de los hallazgos. Esta comunicación debe estar ya traducida a un lenguaje relevante a la política pública y debe incluir las implicaciones concretas de los mismos. Finalmente, debe darse un periodo de transferencia del modelo y de entrenamiento humano. Así, el modelo queda en manos de las entidades públicas que van a usarlo para la toma de decisiones, concluyendo así el proceso de traducción de conocimiento.

Taller

Dado que uno de los objetivos del taller era generar una red de actores, se llevó a cabo un ejercicio de mapeo de actores y procesos con métodos de “Design Thinking” dividido en 4 actividades:

- Mapa de Actores: se llevó a cabo un ejercicio de arquetipos asociados a un listado de actores, con el fin de identificar los actores, papeles e interacciones que se dan en la respuesta a una crisis.
- Mapa de procesos: se llevó a cabo un ejercicio de construir la línea de tiempo de la respuesta a una crisis humanitaria.
- Diseño de dashboard: se llevó a cabo un ejercicio de mapeo de las necesidades de información para toma de decisiones.
- Itinerario de información / *data journey*: con base en los insumos levantados el día anterior, se consolidó la información de mapas de actores, procesos y necesidades de información en forma de un itinerario de información, con el fin de identificar diferentes tipos de respuesta y el papel de la analítica de datos en cada uno de ellos.

Basado en el taller se construyó el siguiente flujo de respuesta como ejercicio de síntesis.

1. Contexto previo y análisis de vulnerabilidades

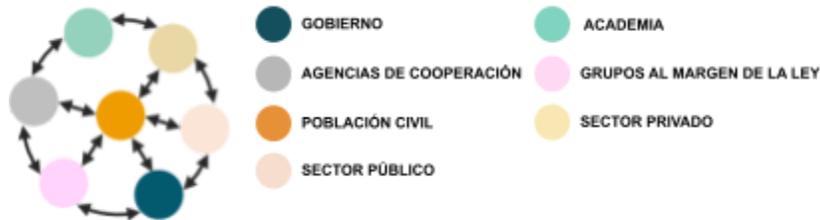
Previo a una emergencia existen una serie de relaciones sistemáticas entre diferentes actores. Esto se refiere a las dinámicas subyacentes entre actores, producto de los sistemas existentes. En el momento de desplegar respuestas a emergencias estas dinámicas pueden influenciar de forma abierta o tácita las interacciones entre actores. Por ejemplo, dinámicas de exclusión a poblaciones marginalizadas pueden traducirse en exclusión a la hora de brindar atención primaria. La responsabilidad de los actores involucrados en esta etapa es estudiar este contexto, identificar las vulnerabilidades y crear protocolos de respuesta integrales. En esta etapa se determinan qué datos están

disponibles así como los protocolos de intercambio de información entre diferentes actores.



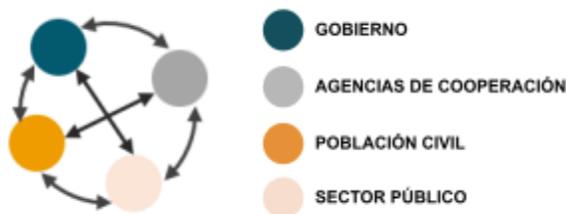
2. Evento desencadenante

El evento desencadenante precipita una cadena de afectaciones. Todos los actores interactúan entre sí. Dado al caos y emergencia desatada por la emergencia, esta interacción inicial está guiada por protocolos preestablecidos y factores sistemáticos. Esto recalca la importancia de la etapa previa, donde el desarrollo consciente de protocolos integrales e inclusivos ayudarán a tener un primer desenlace fluido y equitativo. La academia es el único actor que no suele interactuar directamente con la población civil.



3. Atención Primaria

Se despliega la atención primaria y se entra en periodo de mitigación. Se empiezan a desarrollar respuestas diferenciadas dependiendo de las particularidades de la emergencia. La disponibilidad de datos en tiempo real para guiar la respuesta inmediata (e.g. datos de movilidad, vulnerabilidad, conectividad...) son fundamentales para optimizar el uso de recursos. Para esto es fundamental que en la etapa (1) se haya desarrollado tanto la infraestructura de recolección, procesamiento y análisis de datos como los protocolos para compartirlos. De igual forma es fundamental que en la etapa (1) se hayan definido protocolos claros de protección de privacidad pues es posible que durante momentos de urgencia, bajo la presión de respuesta inmediata, se establezcan precedentes de violación a la privacidad que pueden resultar peligrosos, sobretodo para la poblaciones vulnerables.



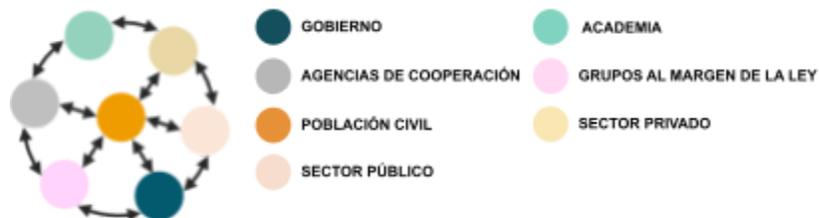
4. Diseño y Estrategia

Los actores encargados del diseño estratégico interactúan entre ellos, colaborando para identificar una estrategia a mediano y largo plazo que dé una solución integral a la Emergencia. Es acá donde se identifica una oportunidad de co-creación de conocimiento entre la academia, el gobierno, las agencias de cooperación y el sector privado. La práctica de traducción de conocimiento es fundamental pues determinará hasta qué punto las metodologías del estado del arte científico pueden ser integradas en el proceso de toma de decisiones frente a una crisis. En cuanto al uso de datos, es importante que la academia actúe como guardián de la privacidad y del uso ético de los datos. Los participantes mencionaron la dificultad de convencer al sector privado de compartir datos en primer lugar, por lo cual es importante pensar en incentivos para que en momentos de emergencia haya mejor colaboración entre las partes.



5. Respuesta integral

Se despliega la solución integral. Una solución integral requiere que todos los actores estén involucrados y haya interacción dinámica entre ellos. Es fundamental que los ejercicios de traducción del conocimiento mencionados en la etapa anterior continúen. El tipo de datos relevantes durante esta etapa suelen ser diferentes a los necesarios durante la etapa de respuesta inmediata. Acá empieza a ser más útil incorporar metodologías de pronóstico y modelamiento, así como datos con historicidad que den un panorama más completo de la situación.



6. Evaluación y rediseño

Una vez controlada la emergencia los actores encargados del diseño de soluciones hacen evaluación y seguimiento de la solución. Se extraen lecciones aprendidas y se actualizan y evalúan los planes de contingencia. Este es un momento excelente para re-evaluar acuerdos de intercambio de datos. Se propuso en el taller que este es un momento para que la academia y el gobierno se sienten a crear metodologías conjuntas informadas con las necesidades identificadas y dificultades de colaboración que afloraron en el transcurso de la crisis. Surgió la noción de que lo último que se quiere en esta etapa es que cada actor tome por su lado pues esto perpetúa el status quo de falta de colaboración intersectorial.



Conclusiones y pasos a seguir

El simposio fue un primer paso para generar una red de actores, identificar los procesos y discutir de manera abierta el papel de los datos y la analítica en el futuro de la respuesta ante crisis humanitarias. Si bien es cierto que se identificaron dificultades y fricciones, también es cierto que se evidenció un potencial muy grande en términos de procesos de analítica que ya se están desarrollando y que todavía no se incorporan en los flujos de procesos de respuesta a crisis humanitarias.

Por otro lado, se resaltó la importancia de que la analítica no se limite a que existan dashboards con información, sino que estos procesos sean acompañados de profesionales expertos en traducir el conocimiento derivado de la ciencia de datos a implementación y toma de decisiones en política pública. Para ello, es fundamental dedicar recursos y personal a la articulación de equipos de ciencia de datos con la academia y las demás instancias intersectoriales e institucionales. Así mismo se enfatizó la necesidad de negociar acuerdos de intercambio de datos así sea únicamente durante situaciones excepcionales. En el caso del sector privado se identificó la necesidad de crear incentivos para promover dichos acuerdos.

A lo largo del simposio, se resaltó la importancia de llevar a cabo estos encuentros de manera recurrente, para fortalecer los lazos, establecer un diálogo fluido entre actores, y lograr así aumentar la resiliencia de las actividades de atención en crisis humanitarias.

Instituciones participantes

- Agencia Analítica de Datos-AGATA-
- Agencia de la ONU para los Refugiados (ACNUR)
- Crisis Ready
- Cruz Roja Colombiana
- DataLama
- Departamento administrativo distrital de salud (DADIS) Cartagena Secretaría de Salud del Distrito de Bogotá
- Fundación Karisma
- Grupo Interagencial de Flujos Migratorios Mixtos (GIFMM)
- Harvard T.H.Chan School of Public Health, Harvard University
- iMMAP
- Instituto Nacional de Salud (INS)
- Ministerio de Salud y Protección Social de Colombia
- Oficina de las Naciones Unidas para la Reducción del Riesgo de Desastres (UNDRR)
- Oficina de Naciones Unidas para la Coordinación de Asuntos Humanitarios (OCHA)
- Organización Internacional para las Migraciones (OIM)
- Organización Panamericana de Salud (OPS)
- Procalculo
- Profamilia
- Secretaría de Salud de Guainía
- Secretaría de Salud Distrital de Barranquilla
- United Nations High Commissioner for Refugees (UNHCR)
- Universidad de Los Andes
- Universidad El Rosario
- Universidad Javeriana
- Universidad Nacional de Colombia

La transmisión del evento se puede encontrar en los siguientes links:

1. <https://www.youtube.com/watch?v=qvJm3gqPT4g&feature=youtu.be>
2. https://www.youtube.com/watch?v=iPM_70oKacU