



Entrevista a Marcelo Arenas

Director Instituto Milenio
Fundamentos de los Datos



Foto: IMF D



Por Federico Olmedo
y Ana Martínez.



Marcelo Arenas es director del Instituto Milenio Fundamentos de los Datos (IMFD), y desde esa posición formó parte de la Mesa de Datos que el Gobierno conformó en abril de 2020 para el asesoramiento técnico en el manejo de la pandemia por COVID-19. En la siguiente entrevista recogemos su lectura sobre lo ocurrido en los últimos meses, poniendo especial énfasis en el manejo de datos relativo a la pandemia.

Marcelo, ¿nos podrías explicar qué es el Instituto Milenio Fundamentos de los Datos?

El IMFD es un instituto interdisciplinario que realiza investigación en diversos temas relacionados con datos. Para poder abordar problemas complejos, en parti-

cular, el impacto que los datos tienen hoy día en la sociedad, contamos con investigadores del área de computación y estadística —que son fundamentales—, pero también necesitamos una mirada desde las ciencias sociales; sin ellos nos quedaríamos cortos. Por eso en el IMFD participan también investigadores de ciencias

políticas y de comunicaciones. Hay muchas otras áreas que también deberían estar presentes, pero lamentablemente los recursos son limitados.

¿Y con qué objetivo se fundó la Mesa de Datos?

Lo que se veía más urgente cuando empezó la pandemia era tener modelos predictivos... tener alguna idea de lo que iba a pasar. Y para generar esos modelos había un problema de datos: se necesitaban datos que estuvieran limpios, integrados, bien modelados, etc. La Mesa de Datos nació entonces con ese objetivo, técnico, bien definido. Y como mesa técnica tenía tres partes. Había una primera parte que generaba los datos; ahí estaba el Departamento de Epidemiología y el Departamento de Estadísticas e Información de Salud (DEIS). Luego había una parte intermedia, el IMFD, que se iba a encargar de desarrollar una plataforma para tomar esos datos, modelarlos, anonimizarlos y publicarlos. Y la tercera parte, los consumidores de los datos, eran varios grupos de modeladores: el Centro de Modelamiento Matemático, la Fundación Ciencia y Vida, el Instituto de Sistemas Complejos de Ingeniería y en menor medida, el IMFD. Pero el objetivo principal del IMFD era disponibilizar los datos del Ministerio de Salud para los grupos que iban a hacer el modelamiento matemático de la epidemia.

¿Sólo para esos grupos o al público en general?

Al público en general. Si bien lo que en ese minuto urgía era tener buenos datos que nos dijeran hacia dónde iba esto, el objetivo era que esos datos, anonimizados y desagregados, fueran puestos a disposición de todo el mundo.



Foto: IMFD



¿Y qué nos perdemos si esos datos no están, o si los tenemos pero están agregados?

Lo explicaría de la siguiente forma. Cuando una persona llega al consultorio se le pregunta "¿desde cuándo tiene usted síntomas?". Por ley, los médicos deben declarar esa información en Epivigila, que es un sistema centralizado para llevar cuenta de los pacientes con COVID y otras enfermedades transmisibles. Ahí se registra también todo lo que viene después: cuándo se realiza el examen, el resultado del examen, cuándo llega el resultado, etc.

Por ejemplo, algo que es fundamental para el modelamiento matemático es el acceso a la fecha de los primeros síntomas. Sin eso no podemos saber desde cuándo el infectado está contagiando.

Si uno podía tomar ese sistema y ver las fechas, podía calcular, por ejemplo, el tiempo entre los primeros síntomas y el resultado del examen PCR, y ese tiempo no es el mismo en todas las comunas, por supuesto. Ahí uno se da cuenta que hay comunas con sistemas sanitarios colapsados y otras que no. Cuando uno tiene la información desagregada, esas diferencias muestran las realidades comunales.

Sólo como dato, la comuna de San Ramón "no tuvo" pacientes COVID en todo marzo. Era impresionante ver cómo Santiago iba cambiando de colores, iban aumentando los casos en las distintas comunas, pero San Ramón que está al medio, seguía en blanco. Entonces uno dice: "¡Esto no puede ser! San Ramón tiene casos y no han sido detectados".

¿Y si efectivamente hubiese habido voluntad de hacer públicos esos datos, se sabe cómo hacerlo respetando la privacidad de los pacientes?

Sí, se sabe cómo publicar datos. Epivigila tiene una cantidad enorme de información; si mal no recuerdo son más de 170 variables que se guardaban por pacien-

Hay que entender que estos procesos de manejo de datos tienen una parte social muy importante [...] Es un proceso mucho más complejo que sólo el proceso técnico.

te. Claro, si uno publica todo el historial de viajes de los últimos meses, y alguien realizó varios viajes, eventualmente se va a poder identificar a la persona... Pero publicar, por ejemplo, las fechas importantes por paciente sabemos que es seguro; uno puede mostrar que esa información no se puede desanonimizar. De hecho el IMFD desarrolló un sistema para llevar a cabo la anonimización de los datos, de manera automática.

¿Cómo te sentiste viendo que éste era el momento en el que podías aportar a la sociedad como científico y el Ministerio de Salud no lo permitía?

Por un lado hay mucha frustración, pero es una combinación de cosas. Finalmente nosotros —el IMFD— nos salimos de la Mesa de Datos porque no teníamos acceso a datos. No teníamos el insumo básico con el cual necesitábamos trabajar, y sin ese insumo no podíamos participar ni en las discusiones, ni en las decisiones.

También hay que entender que estos procesos de manejo de datos tienen una parte social muy importante. Dicho de otra manera, hay una parte que es culpa de nosotros. El pensar que simplemente vamos a llegar a un ministerio a solucionar los problemas es dar un salto muy anticipado. Es un proceso mucho más complejo que sólo el proceso técnico.

¿No estás hablando de nada político, sino de algo social?

Claro. Desde el Ministerio de Ciencia y desde la parte técnica del Ministerio de Salud vi la mejor disposición para tratar de ayudarnos. Pero finalmente esto se fue enredando en todo tipo de cosas extrañas; fue un proceso muy frustrante donde perdimos mucho tiempo en tér-

minos de lucha contra la pandemia. En abril seguíamos teniendo números relativamente buenos —una oportunidad gigante para cerrar comunas y hacer cuarentenas estrictas—, pero en mayo la situación se nos escapó de las manos, se acabó la trazabilidad y los números se fueron al cielo. Puede ser que hubiéramos terminado igual, no sé. Pero sólo estadísticamente, llegar al séptimo lugar en el número de casos con una población que está como en el lugar 60, no tiene sentido. No deberíamos haber terminado en esa situación.

Éstos son datos del Ministerio de Salud, pero todos los ministerios tienen sus datos y cada uno los maneja a su manera. Pareciera que falta una gobernanza de datos un poco transversal, ¿no?

De hecho a ese término quería llegar en algún momento. La gobernanza de datos es un problema en este minuto. Y eso se ve muy claro en los ministerios, donde la gobernanza de datos es muy distinta en uno y en otro. Por ejemplo, el Ministerio de Transportes tiene una parte de telecomunicaciones, que es una industria completamente distinta, mucho más moderna y acostumbrada a manejar datos. Muy distinto a la cultura en un hospital donde, a pesar de que los datos son importantes, es mucho más importante estar atendiendo un paciente que, finalmente, manejar la información.

¿Cómo evalúas la participación del ministro de Ciencia en todo esto? Al fin y al cabo él presidía la Mesa de Datos y él —más que nadie— debería haber velado por la toma de decisiones basada en evidencia.

Varias cosas al respecto... Estoy completamente de acuerdo contigo en que más



La figura del ministro de Ciencia era muy nueva para este Gobierno, y para Chile en general.

allá de su rol político, tiene un deber con las ciencias; tiene que ser independiente, llevar una opinión lo más técnica posible, decir con la evidencia en la mano "ésta es la realidad", y después la autoridad correspondiente tendrá que decidir en base a eso y muchas otras cuestiones.

Pero no todos los ministerios tienen el mismo peso. Hay ministerios que son políticos —muy cercanos a la presidencia— y hay otros que no. Y el primer problema que yo veía es que la figura del ministro de Ciencia era muy nueva para este Gobierno, y para Chile en general. Lo que me tocó ver a mí es que el Ministerio de Salud y el Gobierno tenían problemas para escuchar al Ministerio de Ciencia; no les gustaba la historia tan catastrófica que les contaba. De alguna manera lo dejó ver Andrés Couve cuando Jaime Mañalich se fue, al reconocer que habían cosas en la relación de ellos que podían haber sido mejores.

A mí me parece que el Ministerio de Ciencia ha tratado de ayudar, y hay que reconocer varias de sus iniciativas. Por ejemplo el desarrollo de laboratorios para PCR —transformar a las universidades en laboratorios, por así decirlo— fue un muy buen proyecto, con mucho impacto. Pero también hay una serie de iniciativas que no dieron los frutos que deberían haber dado, la Mesa de Datos es una de ellas en términos de su objetivo inicial.

¿Por qué crees que el Minsal nunca escuchó lo que tenía para decir la comunidad científica? Me cuesta entender esa reticencia cuando hay tanto en juego...

A mí también. Lamentablemente no es una realidad sólo de Chile y hemos visto varios gobiernos que son reticentes a escuchar. En Estados Unidos, por ejemplo, tienen un desastre. Para nadie es fácil escuchar críticas, pero si en algún momento uno se da cuenta que las

cosas no están saliendo bien, uno tiene que agachar la cabeza y decir "bueno, no está funcionando, es el momento de pedir ayuda o de escuchar a otros".

Y estoy de acuerdo en que esto se transformó en una especie de guerra de no escuchar. Y peor todavía, porque puedes terminar escuchando sólo a aquellos que van en tu dirección. En una situación como ésta definitivamente hay que escucharlos a todos, incluso a los que te cuentan las peores historias.

¿Cuál consideras que fue el peor error a nivel de manejo de la pandemia?

El gran problema fue tener un discurso cambiante. Si uno mira Alemania, el discurso fue siempre el mismo: "Ésta es una situación difícil". En cambio, si tú en algún momento dices que es difícil, luego empiezas a hablar de nueva normalidad, después de que vas a abrir el país y empiezas a preparar el regreso de los colegios... Ese discurso ambiguo es un problema porque te encuentras con una realidad que no es así y hay que cambiarla en la otra dirección.

Por otro lado, el manejo de los datos no ha sido bueno y ha generado desconfianza. Jaime Mañalich decía que estaba cansado de dar explicaciones, ¡por supuesto! Si en todas las conferencias de prensa le volvían a preguntar cómo se calculaban los fallecidos porque na-





die lo entendía. El haber tenido datos más transparentes habría ayudado a comunicar mejor el mensaje: “La situación es difícil, lamentablemente se van a morir personas, pero ésta es la realidad que nos va a tocar vivir, y éstos son los números”. En lugar de haber sido un insumo para tener mejores políticas, los datos se transformaron en un problema.

¿Y qué países sí adoptaron un buen manejo de datos?

Muchas veces llegaba el mensaje de que en Latinoamérica nos estábamos enfrentado a una infraestructura que no teníamos... que era propia sólo de países europeos o países como Estados Unidos. Sin embargo en lugares como Colombia o México tienen datos de lujo y han hecho muy bien ese trabajo. Por ejemplo en México tú puedes ver los pacientes desagregados, con las fechas, etc. Nosotros podríamos haber estado en una realidad mucho mejor.

Cuando hayamos superado esto, ¿crees que vamos a mirar con otros ojos la toma de decisiones basada en datos?

Ojalá que sí, que nos demos cuenta. Uno ve que hay organizaciones que se están tomando muy en serio el manejo de datos y su rol en la toma de decisiones, por ejemplo, con la construcción de infraestructuras de datos.

Creo que la pandemia nos va a ayudar en ese sentido; es una forma bien dramática de aprender... Pero también necesitamos una infraestructura y una gobernanza de datos mejor. Y eso porque ahora nos pasó en salud, pero más adelante nos puede pasar en otra parte.

En todo este proceso que nos ha tocado vivir aprendimos que no sólo tener datos de calidad es importante, sino que también necesitamos datos más transparentes, cuyos procesos de creación, limpiado y comunicación sean auditables. Si yo te digo “estos son mis datos desagregados y esto es lo que

Llegar al séptimo lugar en el número de casos con una población que está como en el lugar 60, no tiene sentido. No deberíamos haber terminado en esa situación.

hago para calcular los fallecidos por COVID”, tú puedes repetir exactamente el mismo cálculo. Podríamos no estar de acuerdo si hay que hacer el cálculo de ésta u otra manera, pero tú sabes que lo estoy calculando de esta manera y que lo calculé bien.

Una cosa positiva que veo de la pandemia es que impulsó (o nos obligó a) una transformación digital que hubiera llevado años. ¿Crees que, por ejemplo, el teletrabajo o la educación online vinieron para quedarse?

Creo que sí. También nos empezamos a dar cuenta que teletrabajar no significa sólo estar sentados frente a Zoom; hay tareas que funcionan mucho mejor cuando uno las hace en base a objetivos en lugar de en base a tiempo destinado. Va más allá de obligarnos a estar juntos a través de la tecnología, sino el obligarnos a ser más eficientes en términos del manejo de la tecnología.

En educación todavía lo veo menos. Es cierto, la democratización del conocimiento —que uno pueda tomar un curso, por ejemplo, con el mejor profesor de algoritmos del MIT— me parece increíble. Pero las clases a través de Zoom, o la plataforma que sea, me parecen complejas: veo a los alumnos cansados, a los profesores cansados...

¿Qué tan bueno ha sido el rol de los científicos como comunicadores? ¿Han estado a la altura de las circunstancias?

Es un rol fundamental que en general debiéramos tener en términos de difusión. Creo que hay gente que lo ha hecho muy bien. Se han dado el tiempo y el trabajo para ello. Un ejemplo que me gusta mucho es Ricardo Baeza-Yates.

Ha dedicado mucho tiempo y se ha vuelto una voz que uno puede escuchar... el mostrar porqué el cálculo de fallecidos no está bien hecho, o cómo interpretar ese resultado de una manera tan sencilla. Tú lo escuchas y dices: “Claro, tiene razón, esto hay que mirarlo de esta forma”. Otro ejemplo que me gusta mucho es Jorge Pérez, ha hecho un aporte enorme a través de Twitter. Es algo que tiene que quedarse, no sólo con el COVID que es una cosa gigante y súper importante, sino también con nuestros temas de investigación, tratar de explicar qué es lo que estamos haciendo, por qué lo estamos haciendo, etc.

Y por último Marcelo, ¿qué has aprendido como científico de esta pandemia y qué crees que la sociedad en todo su conjunto ha aprendido?

A nivel profesional lo que me ha llegado más es que en el IMFD ya teníamos esta idea muy fuerte de que los problemas de datos hay que atacarlos de manera multidisciplinaria, y aquí quedó muy claro que eso es así. Nos quedamos completamente cojos si no los abordamos desde múltiples ángulos; se necesita gente del lado de computación, de estadística, de ciencias sociales, etc.

Como sociedad la lección más importante que creo que hemos aprendido es que somos bastante más frágiles en algunos sentidos de lo que pensábamos. Dado los avances médicos, tecnológicos, etc., nos costaba imaginar que una pandemia nos iba a poner en una situación como ésta, en la que llevamos meses y aún no tenemos idea hacia dónde vamos. ■