

EL OBSERVADOR

Matías Castro

TECNOLOGÍA > PANDEMIA

Ofrecen un software uruguayo para trazar la ruta del coronavirus y otras enfermedades

Ponen a disposición en Uruguay, Argentina y Brasil un sistema informático que podrá analizar las relaciones sociales entre infectados, población de riesgo y aquellos que son casos sospechosos



Fue cuestión de horas. El sábado 14 de marzo de mañana, unas 12 horas después de que Presidencia anunciara el primer caso y las primeras medidas por el coronavirus, dentro de la empresa Quanam se planteó un desafío. Especializados en análisis de datos, entre otras tareas, preguntaron si alguno de sus empleados podía crear una solución para contribuir desinteresadamente al combate de la pandemia. En la madrugada del lunes 16, el ingeniero Héctor Coteló ya había programado una.

Se trata de un programa que no tiene nombre, pero sí una finalidad específica. Es una gran base de datos que contiene información sobre contagiados, posibles contagiados, sus relaciones y sus recorridos. Muestra todo eso en forma de gráficos (que se denominan grafos) y permite no solo buscar los grupos, zonas y personas en riesgo, sino tomar decisiones y también hacer análisis para el futuro. La están ofreciendo honorariamente en Uruguay y también fuera de fronteras, a través de IBM y las filiales brasileña y argentina de Quanam.

El programa de Coteló carga y vincula datos de personas, hogares, trabajos, actividades como asistencia a clubes deportivos y similares, eventos a los que asistieron, relaciones entre los individuos (familia, amistad, trabajo), encuentros y otras variables. Cada uno de esos datos es lo que Coteló denomina "entidad" y se va enlazando y conectando con los demás. En el caso del coronavirus, rastrea su propagación y los potenciales focos de contagio. "Si tenés geolocalizadas las casas de la gente en cuarentena, podés crear un mapa de calor

que indique donde hay más focos. Lo mismo sucede con las fechas: podés analizar con una línea del tiempo, que caso de este virus es un análisis crucial”, apunta Cotelo.

“Esto sirve para que las autoridades puedan tomar decisiones basadas en datos”, explica. “El programa es un modelo de grafo con nodos y vértices que enlazan entidades”, agrega. Para traducir su terminología, usa a Facebook como ejemplo. En esa red social cada persona sería una “entidad”, los nodos, vértices y lazos entre ellas son las formas de representar si son amigos, la cantidad de intermediarios y otros datos como si tienen parentesco o no. “Si unís los datos de toda la gente de esa red, te da un grafo”, dice.

Este tipo de software se usa habitualmente para estudios en entidades bancarias, en redes de telefonía y en redes de Internet. En este último caso sirve para analizar dónde se generan cuellos de botella y se enlentece el tránsito de los datos, o para entender cuáles son los puntos críticos que podrían hacer caer una red entera.

“Volviendo a la analogía de Facebook, el punto crítico sería la persona con más amigos del mundo. O quien esté en medio de la red, es decir, la persona a la que se llega con menos intermediación de amigos. Si hubiera virus en esas entidades, o personas, ellos serían clave. Por eso es interesante analizar este tipo de grafos y redes”, explica Cotelo.

El programa fue creado sobre un sistema de IBM llamado i2 Analyze, que sirve para analizar grandes cantidades de datos en empresas. Es un conjunto de herramientas que se han utilizado, entre otras cosas, para la investigación de fraudes y delitos.

Una simulación

Sobre esta base Cotelo ya había desarrollado un software para el Sistema Nacional de Información Ganadera, del Ministerio de Ganadería. Con él se analizaban predios desde el lado epidemiológico y se veía cómo una epidemia podía expandirse en el ganado. “Ahora podía darle una vuelta de tuerca para aplicarlo al combate al Coronavirus”, comenta.

Durante las casi 40 horas que trabajó de continuo, cargó su software con 20 mil datos ficticios, obtenidos de servidores que ofrecen esos contenidos para experiencias de este tipo. Empleó una técnica llamada ETL, extracción, transformación y carga, por el que se toman esos datos y se cargan al sistema. Reutilizó datos que ya tenía de trabajos anteriores, como ciudades, barrios y códigos postales y usó la información sobre ubicación y tipo de centros educativos que está en el catálogo de datos abiertos del gobierno.

También puso su propia información y la de su entorno, además de los que son públicos en torno al supuesto paciente cero, Carmela Hontou, y con ellos armó varios ejemplos sobre su funcionamiento.

Para que el programa arroje conclusiones válidas se puede comenzar con algo tan simple como los datos de las personas infectadas, su entorno cercano, dónde trabajan, con quiénes viven y han interactuado. “Así podemos comunicarnos con esas personas y decirles que son plausibles de hacer el análisis. Eso, a su vez, va escalando por saltos cuando el grafo nos muestra que esas personas tuvieron contacto e interactuaron con otras”.

Cotelo propone como fuentes de datos la información del BPS, que mostraría dónde trabaja cada persona. El MSP tiene los datos sobre las mutualistas y afiliados. TickAntel y otras ticketeras, por ejemplo, podrían aportar quiénes fueron a qué eventos masivos (en el ejemplo que armó, él se coloca a sí mismo en un concierto del Antel Arena y agrega a las personas con las que fue y en qué fecha sucedió). Antel y el Codicen podrían dar los datos sobre las escuelas así como la Dirección Nacional de Migraciones los ingresos y egresos al país.

Para procesar tal cantidad de datos estima que no se precisaría un centro de cómputos, sino solamente un servidor preparado, o “polenteado”, como se dice en su jerga. “Somos 3.5 millones de personas, no es tanta data”, afirma. De hecho, el modelo sobre los 20 mil datos lo armó y procesó en su laptop. |

[Tecnología](#) [software](#) [coronavirus](#) [Member](#)

[REPORTAR ERROR](#)