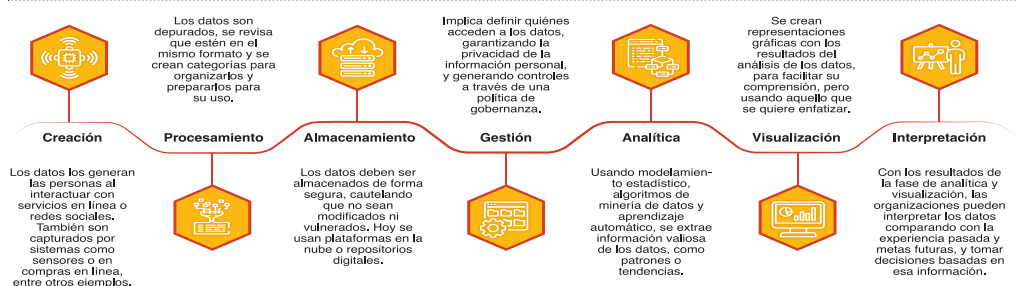


LA CIENCIA DE LA COMPUTACIÓN BUSCA SOLUCIONES QUE RESGUARDEN ASPECTOS TÉCNICOS Y ÉTICOS:

La invisible ruta de los datos digitales y sus crecientes desafíos

Detrás de todas las aplicaciones, servicios y tecnologías digitales, incluyendo los sistemas de IA generativa, existe un flujo constante y silencioso de datos. Su análisis alimenta la toma de decisiones en áreas que van desde las ofertas en los sistemas de compras en línea modeladas por los patrones de consumo de una persona, hasta la creación de perfiles de riesgo en salud que determinan el tratamiento que recibirán los pacientes.

Datos: desde su creación hasta la interpretación



Fuente: DCC

Infografía Estudio Pixel

Una experiencia cotidiana que suele ser utilizada para entender qué son los datos digitales y cómo se usan es la interacción que las personas tenemos a diario con nuestras redes sociales. Un "Me gusta" en cualquiera de estas plataformas activa un flujo de procesos que, partiendo en esa acción, va generando un perfil con las preferencias de las personas y explica la aparición en nuestros perfiles de recomendaciones y publicidad personalizadas.

Sin embargo, para comprender en profundidad el impacto que puede tener la utilización de nuestros datos, es bueno pensar en aquellos que son más privados, personales. Tras la aparición de síntomas preocupantes de salud, viene una cadena lógica de eventos: la solicitud de una hora médica o la posibilidad de una consulta en línea; la compra de un bono; la solicitud de exámenes, e ir por medicamentos a una farmacia donde entregamos nuestro RUT a cambio de descuentos.

Los datos generados por cada una de esas acciones entran a lo que se conoce como el ciclo de vida o la ruta de los datos digitales (ver infografía). Y si analizamos las interrogantes que se plantean en muchos de esos pasos, es comprensible que las ciencias de la computación estén trabajando con sentido de urgencia en el estudio de sus posibilidades, de sus potenciales riesgos, de los desafíos tecnológicos para solucionarlos y en la generación de conciencia sobre el peligro de la utilización de los datos más sensibles de una persona.

LOS DATOS QUE RECOGEMOS SON LOS QUE NOS DEFINEN

A la hora de investigar los síntomas con los que se presenta un infarto al corazón, lo usual es alertar cuando hay dolor agudo al lado izquierdo que se extiende hasta el brazo. Pero dicha descripción no se aplica a las mujeres; en ellas, las señales pueden ser fatiga,

mareos o dolor abdominal, diferencia que inició durante mucho tiempo en la falta de detección y mayor mortalidad femenina por infartos. Esto se debe a que gran parte de los perfiles en el área de salud han sido construidos históricamente con datos provenientes de la población masculina. Lo mismo ha sucedido, por ejemplo, con el instrumental médico o dispositivos como marcapasos.

Considerando que en la actualidad los procesos están altamente digitalizados, una pregunta clave es cómo los datos están siendo recolectados y si —en la captura— hay diversidad de variables como género, edad, etnia y hábitos. Uno de los desafíos técnicos es asegurar la calidad y representatividad; si los nuevos datos siguen siendo incompletos, el sesgo persiste. Y esta definición no la realizan los sistemas; la decisión de qué información se recaba la toman seres humanos, que muchas veces no visualizan que la exclusión de variables y grupos de la sociedad puede acrecentar diferencias tan esenciales como el acceso con equidad a un sistema de salud que no discrimine, comenta Iván Sipiran, del DCC.

QUIÉN PROTEGE Y ACCEDE A NUESTROS DATOS

Una vez capturada la información personal, debe ser procesada: obtener duplicados, corregir formatos y, un paso clave, definir las categorías para organizar los datos. Aunque hoy existen herramientas para detectar su potencial pérdida y corregir sesgos durante la limpieza, la categorización podría eliminar patrones que a primera vista parecen atípicos o errados, pero que en realidad reflejan diferencias, por ejemplo, biológicas o sociales. Por ende, el procesamiento tampoco es neutral.

En cuanto al almacenamiento, en una época en que se produce una cantidad inabarcable de datos digitales, los sistemas en la nube son una solución que permite acceder a ellos desde cualquier lugar. El reto técnico está en

asegurar su integridad y la diversidad de información que contienen, además de su seguridad y la soberanía sobre ellos, para que no sean vulnerados, advierte el académico Matías Toro.

CÓMO SE MANEJA LA CARRETERA DE LOS DATOS

Un punto crítico en este ciclo es su gestión. Uno de los primeros pasos es decidir quiénes los utilizan y con qué nivel de detalle. Considerando que las normativas actuales apuntan a proteger la información sensible, se espera que estos criterios formen parte de una política y gobernanza de datos, auditable y transparente.

La gestión también implica asegurar que la calidad de los datos se preserve en el tiempo, es decir, que la intervención humana —el más común de los errores— no genere alteraciones. Otro factor esencial es la representatividad y fidelidad del registro, es decir, garantizar que desde su base se integre la mayor cantidad y profundidad de variables.

Ahora bien, los desafíos actuales apuntan a procesos aún más complejos en las ciencias de la computación. Entre ellos, cómo integrar distintos tipos de datos sin que estos pierdan su calidad, y se pueda —de manera efectiva— enriquecer la información con la que se cuenta. Además, se investigan las capacidades técnicas para su procesamiento y la infraestructura misma que permita acceder a ellos con velocidad y agilidad, explica Aidan Hogan, director del DCC.

DEL ANÁLISIS Y LA VISUALIZACIÓN A LA TOMA DE DECISIONES

Los pasos anteriores representan el ideal para llegar a la etapa de análisis, pero, pese a que son críticos, hay ocasiones en que alguno de ellos no se realiza para acelerar los procesos.

En la etapa de análisis, se descubren patrones —como el aumento sostenido en la compra de bonos para un área médica específica por parte de un grupo

de la población— o hacer modelos predictivos, como la mejora progresiva en la pesquisa de infarto en mujeres al integrar los detalles de sus síntomas particulares.

Las posibilidades para analizar todo tipo de fenómenos en la actualidad han aumentado de manera vertiginosa, de la mano de algoritmos capaces de procesar millones de casos y encontrar patrones invisibles, técnicas de procesamiento de lenguaje natural para extraer información de datos escritos, y la captura de material visual y su integración con distintos tipos de información, entre otras técnicas. Sin embargo, con la escala también aumentan los desafíos: los modelos aprenden de los datos con los que los entrenamos y de su calidad dependerá la fidelidad con que representan aquello que se quiere estudiar, agrega Sipiran.

Para ayudar a comprender mejor la información que emerge del análisis de datos, la visualización de información permite representar de manera gráfica, más comprensible, patrones que —de otra forma— sería difícil entender, ejemplifica Benjamin Bustos. Los desafíos aquí apuntan a la información que se decide visualizar: en esa decisión, es clave evitar la simplificación o el refuerzo de estereotipos. Pero otro riesgo emergente es la frecuencia con la que estas visualizaciones están siendo delegadas a sistemas de inteligencia artificial, lo que si bien permite tomar decisiones en tiempo real y a una escala más grande, conlleva muchos riesgos.

La interpretación es el punto en la ruta de los datos donde el proceso queda en manos de las personas. Con la información obtenida, y a la luz de la experiencia, los patrones históricos y las metas que se desea conseguir, es posible tomar decisiones basadas en datos. La gran interrogante es el contenido de dicha decisión y si esta permitirá que la información que cada uno de nosotros genera contribuya a un futuro con mayor equidad y bienestar para toda la sociedad.

COLUMNA DE OPINIÓN

Ciencias de la computación: una de las disciplinas más transversales y transformadoras de nuestro tiempo



JOSÉ CORREA,
vicerrector de
Tecnologías de la
Información,
Universidad de Chile

En los últimos años, el desarrollo tecnológico ha sido gigantesco, hasta abrumador. La irrupción de la inteligencia artificial generativa con capacidades insospechadas ha sorprendido al mundo. Estos desarrollos son una

demonstración del poder de las ciencias de la computación para abordar desafíos que, desde sus cimientos teóricos, se materializan en soluciones que atienden las necesidades más apremiantes de la sociedad. Esta es una ciencia cuya definición más elemental es hacer que nuestros computadores funcionen mejor y más rápido, basándose en el modelo fundamental de Turing que permanece impávido frente a los cambios tecnológicos. En la práctica, las ciencias de la computación se han transformado en una de las disciplinas más transversales y transformadoras de nuestro tiempo.

Esta transformación no se explica sin la historia. En nuestro país, los primeros pasos de la computación universitaria los dio la Universidad de Chile, primero con la llegada de máquinas de alta capacidad para la época y luego con la fundación del Departamento de Ciencias de la Computación (DCC). La computación no es solo una herramienta para científicos e ingenieros; son las personas, los especialistas, quienes hacen la diferencia. Así, pues, la formación de los primeros expertos en el área incluía el proceso completo, "desde los fierros y la matemática hasta las aplicaciones". Esto reflejaba la dualidad inherente de las ciencias de la computación: la capacidad para resolver problemas prácticos así como aquella para explorar los límites del pensamiento lógico y abstracto. Esta dualidad permite que las ciencias de la computación aborden desafíos a escalas muy distintas. Por un lado, se ocupa de problemas profundamente estructurales, casi filosóficos, mientras que, por otro, se aboca a proponer mejoras muy concretas para la vida de las personas. A nivel local, el DCC ha sido pionero en ambos ámbitos, desarrollando, por ejemplo, la internet en Chile, pero también haciendo investigación de frontera y formando a quienes irán construyendo el futuro digital del país.

La computación de hoy también nos hace cuestionarnos como sociedad. El avance de la inteligencia artificial es el ejemplo más claro. Nos enfrentamos a preguntas que pertenecían a la ciencia ficción: ¿Podrá la IA reemplazar la actividad científica humana? O, parafraseando a Harvey Lederman, ¿qué perderíamos si la cura del cáncer, por ejemplo, fuera descubierta por una IA? ¿Es más importante el resultado de una investigación o el proceso humano y su consecuente gozo en la creación de conocimiento? En este escenario de crecimiento exponencial y preguntas abiertas queda claro que las ciencias de la computación son mucho más que programar o construir máquinas; son un puente entre la lógica abstracta y las necesidades concretas del mundo. El DCC de la Universidad de Chile fue, ha sido y será pieza clave en este camino.

"Queda claro que las ciencias de la computación son mucho más que programar o construir máquinas; son un puente entre la lógica abstracta y las necesidades concretas del mundo".

DATOS SOBERANOS

Si los datos digitales de un país, como la salud de su población, la ubicación de recursos naturales estratégicos o la infraestructura militar, son generados dentro de su territorio, pero almacenados en servidores que están en otros países, ¿quién tiene derecho a ver y usar esa información? ¿Aplican las normativas de la nación que es fuente de esos datos, o la legislación del lugar en donde están guardados físicamente? ¿Y qué pasa si, para reducir el riesgo de pérdida de información, los datos están

repartidos en servidores de distintos países?

Este es uno de los grandes debates en la actualidad, y de especial urgencia si se considera que las estrategias de digitalización del Estado apuntan a que el volumen de datos soberanos crezca radicalmente.

Para Alejandro Hevia, el desafío que está estudiando el DCC es cómo mantener el control de los datos y, al mismo tiempo, utilizarlos de manera colaborativa para resolver problemas

comunes, interactuar y trabajar en conjunto con otros países, preservando la privacidad.

Es decir, se pueden conciliar los conceptos de soberanía de la información digital con un uso útil y colaborativo de los datos. Este es uno de los grandes debates en la actualidad en materia de soberanía de datos, y de especial urgencia si se considera que las estrategias de digitalización del Estado apuntan a que el volumen de datos soberanos crezca radicalmente.

Primer Servidor Web de Latinoamérica
Este primer servidor comienza a operar en el DCC.

Primer sitio web con información sobre Chile
El DCC desarrolla el primer sitio web, SunSITE, y Chile adquiere así presencia en la red.

NIC Chile

Se consolida como unidad alojada en el DCC, para administrar los nombres de dominio .CL con una reglamentación que, entre otros, resuelve conflictos de dominios.

Nace el programa de Doctorado en Ciencias e Ingeniería en Computación.

Factura electrónica
En conjunto con el SIJ, el DCC diseña el sistema de gestión de facturación electrónica.