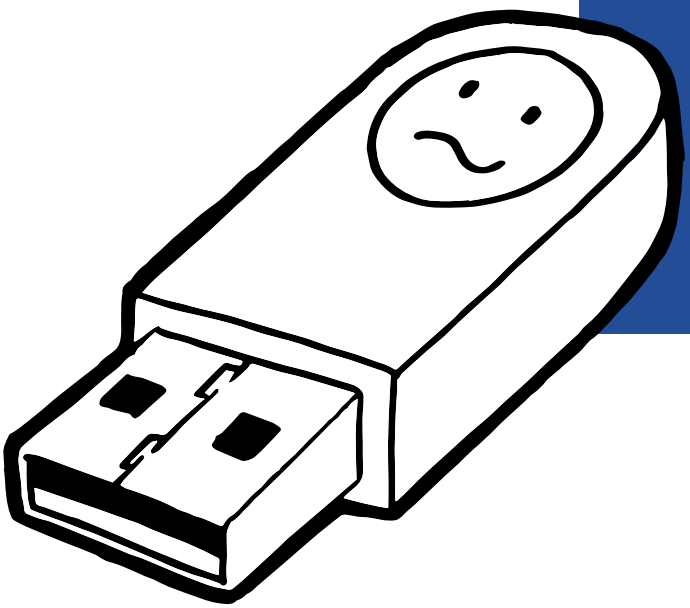
Dilema 91

Autores responsables:

Esteban Puentes
Damián Gildemeister

Caso ficticio, creado con fines académicos



Cuando el origen de los datos es un problema

Sobre las bases de carácter confidencial y su obtención.

En sus investigaciones, los y las economistas utilizan extensivamente bases de datos para realizar sus aproximaciones empíricas y testear sus hipótesis. La información contenida en estas bases es recopilada administrativamente o mediante encuestas, y los entes encargados de estos procesos construyen protocolos para su acceso. Muchas bases de datos contienen datos considerados sensibles, porque su uso indebido puede generar daño a instituciones o personas en particular. En este texto se aborda la historia de Sherlock, un joven economista que realizó una investigación con datos obtenidos de manera irregular. Sherlock accedió a datos con información de crímenes cometidos por niños/as de la ONAME, lo que constituyó un peligro para la privacidad de estos jóvenes y su entorno cercano. El uso de información sensible en las investigaciones económicas genera una serie de cuestionamientos éticos, comúnmente escondidos por la normalización del tráfico irregular de bases de datos. En este caso se plantearán al lector algunos de estos cuestionamientos y se entregan los antecedentes necesarios para un discernimiento informado.

Sherlock es un joven economista egresado el 2016 de la Universidad del Arrollo, quien deseaba comenzar su vida laboral desempeñándose como investigador. Desde que era "mechón", siempre estuvo interesado en estudiar temas sociales, especialmente de marginación y desigualdad. En marzo de 2017, Sherlock entró a trabajar para una Organización No Gubernamental (ONG) enfocada en estudios de derechos humanos dentro de las cárceles chilenas. En el último tiempo ha estado muy presente en el debate público la calidad de los hogares de la ONAME, y en la ONG de Sherlock le pidieron que realizara una investigación empírica sobre el contexto de los niños/as de esta institución. Sherlock sabía que esta investigación podía ser un buen antecedente para postular al Magíster de Políticas Públicas en Espacio Harvard que siempre había querido.

Como buen economista, Sherlock sabía que para llevar a cabo su investigación de manera rigurosa y sofisticada

era imprescindible contar con micro datos de delincuencia juvenil. La principal ventaja de trabajar con micro datos es que al presentar información muy desagregada (generalmente a nivel de individuo u hogar), se cuenta con más observaciones y se pueden utilizar herramientas econométricas más precisas. Lamentablemente para Sherlock, la información de delincuencia a nivel individual, es bastante restringida. La Defensoría Pública Penal (DPP) es el organismo estatal encargado de sistematizar esta información y la da a conocer mediante informes descriptivos, con estadísticas agregadas, pero sin la posibilidad de identificar individualmente a las personas que han sido condenadas por delitos o cumplido penas en recintos penitenciarios. En el caso de la delincuencia juvenil, la DPP a través de la Unidad Juvenil tiene la obligación de hacer pública información estadística de carácter descriptivo sobre el comportamiento del sistema de justicia penal juvenil en el marco de la Ley 20.084 de Responsabilidad Penal del Adolescente.

Un día durante la hora de almuerzo Sherlock le comentó a Harriet, una colega del trabajo, sobre sus aspiraciones como investigador y los problemas que había tenido para contar con los datos necesarios. Ante esto, Harriet le comentó que tenía acceso a una base de datos de la ONAME (generada por la DPP), de tipo panel, que hace un seguimiento a través de 12 años a todos los estudiantes que, en los años 2000, 2001, 2002 y 2003 estaban en primero básico. Específicamente, contiene información de todos los crímenes por los que han sido condenados los/as jóvenes incluidos, identificándolos mediante el *mrun*, código único para cada estudiante que permite cruzar esta base con las de pruebas estandarizadas como el Sistema de Medición de Aprendizaje Estandarizado (SIMAE) y otras fuentes de datos del Ministerio de Educación y Escuelas (MEE), como las notas.

- ¿Cómo tuviste acceso a esa base?, preguntó Sherlock.

- Soy ayudante de investigación de una profesora que solicitó las bases con Gendarmería hace unos años. Te mueres lo complicado que fue que le dieran los permisos.

Aquí Sherlock dudó. Él sabía que si las bases son de difícil acceso es por algo. Después de un rato de reflexionar se acordó que su amigo de la Facultad, Franco Garay, hizo su tesis con una base que sólo pudo acceder por el trabajo de su papá y con la que obtuvo la calificación máxima. Pensó en lo típico que resultaba conseguirse bases con conocidos cuando era estudiante, así que finalmente dijo:

- ¿Podrías pasármela?

- … Sí, obvio, ¿tienes un pendrive?

Así, Harriet compartió esta base con Sherlock, quien estaba muy contento pues sabía que estos datos le permitirían hacer una investigación sobre delincuencia juvenil y temáticas educacionales, lo que le motivaba mucho porque sabía que podían constituir un aporte para hacer frente a la crisis de la ONAME y plantear políticas públicas de inserción escolar para jóvenes condenados por delitos.

El objetivo de la investigación de Sherlock correspondía a cómo la presencia de estudiantes con prontuario delictual afectaba al aprendizaje de sus pares. Después de varios meses de trabajo encontró que por cada compañero/a con prontuario delictual, el puntaje promedio del curso en evaluaciones estandarizadas disminuía en 5 desviaciones estándar. La enorme magnitud de este efecto lo motivó a profundizar en su investigación, construyendo un mapa de calor para el Gran Santiago que permitía identificar las comunas que presentaban mayor número de estudiantes con antecedentes judiciales en sus establecimientos, encontrándose marcadamente concentrados en el sector norponiente de la ciudad. Para hacer más amigable el acceso a la información le pidió a Eugenio, un informático que también trabajaba en la ONG, que programara un buscador con el cual mediante el nombre de una comuna o establecimiento se pudiese acceder a todas las estadísticas de crimen juvenil asociadas.

Completamente convencido de que su investigación debía ser considerada para políticas públicas, Sherlock logró concretar una reunión con el centro de estudios del MEE para presentarle sus resultados. En dicha reunión, Sherlock les mostró el mapa de calor y mencionó, con nombre y ubicación, cuáles eran los establecimientos con más estudiantes procesados por delitos. Además, les mostró detalles de los delitos cometidos en las 5 comunas con más antecedentes del país y los resultados SIMAE que estos establecimientos tenían. También les mostró el buscador que había programado Eugenio y les comentó sobre el gran potencial que tenía como herramienta para que los padres pudiesen escoger informadamente el colegio de sus hijos/as.

Habiendo visto atentamente los resultados de la investigación, Minerva, la directora del Centro de Estudios, le preguntó por el origen de los datos. Nervioso, Sherlock respondió vagamente que "se los pasaron". Minerva, indignada, le dijo que si no eliminaba esos datos inmediatamente de su computador lo denunciaría con la

brigada del cibercrimen de la Policía del Peritaje (PDP) ya que estaba violando no sólo la Ley 19.223, sino que las condiciones de uso de las bases de datos empleadas, atentando contra la propiedad intelectual . A continuación, Minerva tildó a Sherlock de inescrupuloso y le dijo que era una vergüenza y un peligro para la investigación de ciencias sociales por no manejar ningún protocolo de uso de datos nominados. Sherlock, medio aturdido por la reacción de Minerva y sin saber bien de qué se le acusaba, acató la orden, pidió disculpas y decidió nunca más trabajar en esa o cualquier otra investigación.

¿Por qué es algo incorrecto, desde el punto de vista ético, compartir bases de datos sin los debidos permisos?

¿Cuál fue la acción errónea desde un comienzo?

¿Qué podría haber hecho Sherlock para no incurrir en una falta ética?

Si ves alguna vez a compañeros/as pasándose bases de datos por pendrive, ¿qué deberías aconsejarles?

¿Qué estaba en juego con el uso de esos datos?

Anexos

Anexo 1: Información contenida en base de datos ONAME

Si bien el caso descrito en este dilema es de carácter ficticio, en este anexo se busca problematizar los posibles riesgos de mal manejo de esta información. En el caso se dijo que la base de datos del Oname contaba con el identificador único *mrun*, además de contar con información sobre los delitos cometidos por los estudiantes de tres generaciones distintas. Al tener identificador de *mrun*, se puede cruzar con datos administrativos de notas del MEE, así como con información del SIMAE. De este modo se puede identificar el establecimiento al que asisten los/as jóvenes procesados, así como también se pueden identificar a sus compañeros/as de curso. Se puede obtener información de la carrera a la que entraron, de su rendimiento en todas las evaluaciones estandarizadas y escolares, así como información socioeconómica y psicoemocional declarada por los apoderados y los propios estudiantes.

Fuente: Elaboración propia.

Anexo 2: Fragmento de la Ley 19223: Tipifica figuras penales relativas a la informática

TIPIFICA FIGURAS PENALES RELATIVAS A LA INFORMÁTICA

Teniendo presente que el H. Congreso Nacional ha dado su aprobación al siguiente Proyecto de Ley:

"Artículo 1°. - El que maliciosamente destruya o inutilice un sistema de tratamiento de información o sus partes o componentes, o impida, obstaculice o modifique su funcionamiento, sufrirá la pena de presidio menor en su grado medio a máximo. Si como consecuencia de estas conductas se afectaren los datos contenidos en el sistema, se aplicará la pena señalada en el inciso anterior, en su grado máximo.

Artículo 2°. - El que, con el ánimo de apoderarse, usar o conocer indebidamente de la información contenida en un sistema de tratamiento de la misma, lo intercepte, interfiera o acceda a él, será castigado con presidio menor en su grado mínimo a medio.

Artículo 3°. - El que maliciosamente altere, dañe o destruya los datos contenidos en un sistema de tratamiento de información, será castigado con presidio menor en su grado medio.

Artículo 4°. - El que maliciosamente revele o difunda los datos contenidos en un sistema de información, sufrirá la pena de presidio menor en su grado medio. Si quien incurre en estas conductas es el responsable del sistema de

información, la pena se aumentará en un grado.". Y por cuanto he tenido a bien aprobarlo y sancionarlo; por tanto, promúlguese y llévese a efecto como Ley de la República.

Santiago, 28 de mayo de 1993.-

ENRIQUE KRAUSS RUSQUE, Vicepresidente de la República. -

FRANCISCO CUMPLIDO CERECEDA, Ministro de Justicia.

Fuente: Ministerio de Justicia, 1993.

Referencias

Ministerio de Justicia. (7 de junio de 1993). Tipifica Figuras Penales Relativas a la Informática. Chile: Biblioteca del Congreso Nacional.
Recuperado de: *https://www.leychile.cl/Navegar?idNorma=30590*