



MARTA BELTRÁN

IMPLICACIONES DE LA RECOGIDA Y EL TRATAMIENTO MASIVO DE DATOS

La ética de la investigación

El *big data* está demostrando ser una herramienta muy potente para la ciencia en casi todas las áreas del conocimiento. Es necesario que los investigadores reflexionen acerca de los aspectos éticos de su utilización y que se hagan responsables de la protección de los datos que recogen y utilizan a lo largo de todo su ciclo de vida.

ILUSTRACIÓN: DANIEL TORNERO

Una de las ramas más importantes de la ética es la ética normativa, encargada de proponer criterios y normas que regulen la conducta humana dentro de un determinado ámbito. En el científico, la ética normativa suele plantearse desde el enfoque de la deontología, es decir, se sostiene que los investigadores deben cumplir ciertos deberes para actuar moralmente, independientemente de las consecuencias que no cumplirlos pudieran tener.

Cuando pensamos en la deontología científica, a todos nos vienen a la mente algunas de sus aplicaciones más mediáticas. La bioéti-

ca, en todo lo que tiene que ver con la regulación de la investigación genética, la eutanasia o la donación de órganos; la ética relacionada con la experimentación con animales y humanos durante las pruebas de nuevos fármacos; la ética que tiene que ver con la investigación en áreas que tienen aplicación militar directa como la producción de nuevas armas; o en los últimos tiempos, la ética de la inteligencia artificial o la roboética, por ejemplo, que plantean preguntas como si es ético que un coche autónomo atropelle a una persona para salvar a otras cinco, si es correcto que un algo-

ritmo que toma decisiones importantes —como si se nos concede un crédito o no, o si se nos somete a una operación quirúrgica o no— no sea transparente en sus criterios, o qué consideración debemos darle a una máquina con inteligencia dentro de nuestra sociedad.

Como afirmaba Carl Sagan: “Cada gran avance tecnológico en la historia de la especie humana, desde el invento de las herramientas de piedra y la domesticación del fuego, ha sido éticamente ambiguos”. Es lógico que entre tantos temas apasionantes y, sobre todo, controvertidos, pase desapercibida

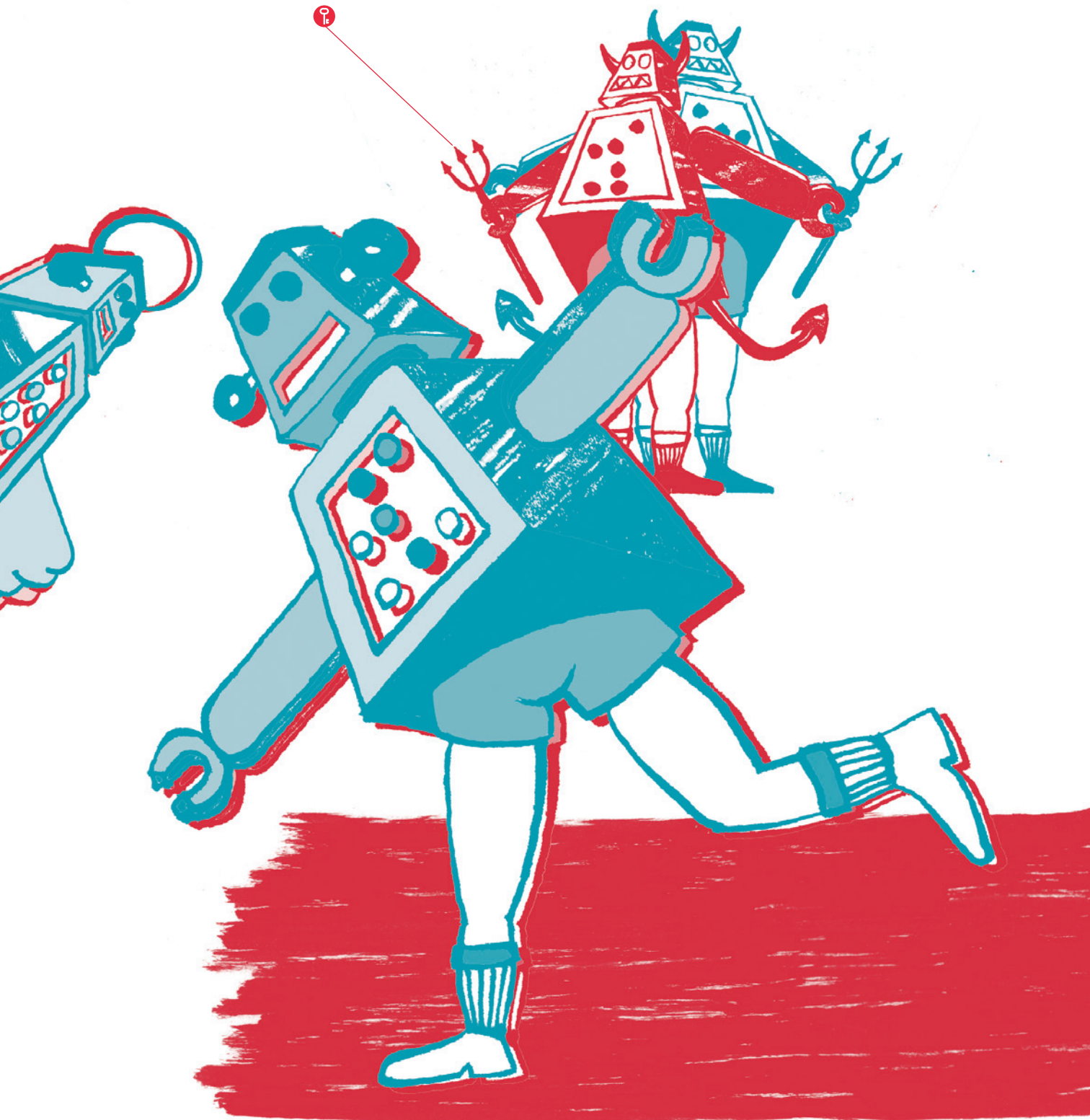


Implications of massive data collection and processing **RESEARCH ETHICS**

Big data is proving to be a very powerful tool for science in almost every area of knowledge. Researchers need to think about the ethical aspects of its use and take responsibility for the protection of the data they collect and use throughout their life cycle.

Keywords: data, science, ethics, privacy, data protection, regulation, big data

Palabras clave:
big data,
ciencia, ética,
privacidad,
protección
de datos,
regulación



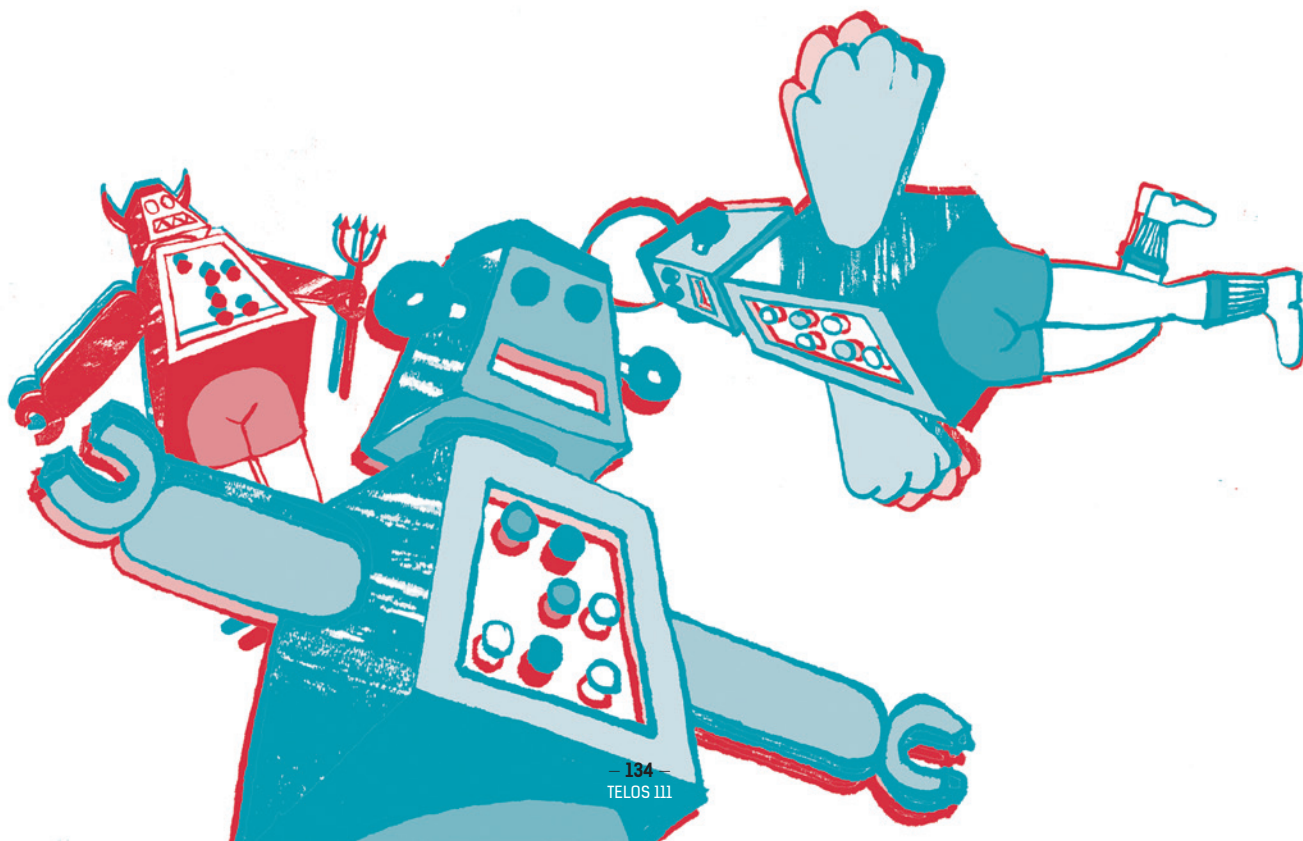
en muchos casos una ética del día a día que es esencial para el funcionamiento de los centros de investigación, para el desarrollo digno y honesto de la carrera de los investigadores y para la producción de resultados que tengan un verdadero impacto en la sociedad.

Esta ética o deontología del investigador plantea cuestiones como el ejercicio de la duda metódica y un correcto planteamiento del método científico, el adecuado diseño y desarrollo de los experimentos que justifican las hipótesis propuestas en los proyectos de investigación,

el reconocimiento necesario al trabajo de los demás —en el pasado mediante las bibliografías y sistemas de citas pero, también en el presente, respetando la autoría— o el uso eficaz de los recursos, ya que siempre hay una o varias entidades que realizan un esfuerzo para financiar un proyecto de investigación. Pero obviamente, también se proponen normas que evitan el plagio, la manipulación o la fabricación de datos, poner en peligro a terceros, la apropiación de resultados, etcétera. Todos los investigadores deberían conocer y aplicar los principios básicos

que se recogen en la Declaración sobre la ciencia y el uso del saber científico¹ adoptada por la Conferencia Mundial sobre la Ciencia de la UNESCO del 10 de Julio 1999.

Por desgracia, igual que a un soldado el valor se le supone, a un investigador la ética se le supone. En pocas instituciones y proyectos de investigación se realizan formaciones explícitas —y periódicas, no sirven si solo se realizan para los investigadores en formación, que suelen tener escasa responsabilidad en la gestión de los proyectos— acerca de temas éticos y relacionados con la



deontología, se proponen y comunican códigos éticos, se evalúa su grado de cumplimiento, se debaten de manera abierta y transparente los temas más controvertidos. La excepción son aquellos centros y proyectos que, por sus líneas de investigación y áreas de trabajo, están obligados a que la ética sea uno de los pilares básicos de su labor —en temas antes mencionados como la bioética, la investigación con aplicaciones militares o la inteligencia artificial— y a demostrarlo.

Por eso no ha de extrañarnos que una revolución como la que ha supuesto la utilización de la ciencia de datos en casi todos los campos de la investigación científica, no haya supuesto una revolución también en las prácticas deontológicas asociadas a esos campos. Las implicaciones de la recogida y tratamiento masivo de datos, o de la producción de modelos descriptivos y predictivos en multitud de áreas de conocimiento, todavía no se gestionan de manera adecuada en la mayor parte de los casos. Y no solo en la dimensión ética que tiene que ver con la privacidad y la protección de los datos, sino también en la organizativa —quién se hace responsable de la protección de los datos, cómo se hace un diseño adecuado de la investigación para proteger la privacidad—, administra-

tiva —cómo debe cambiar la burocracia asociada a la gestión de los proyectos y a la publicación de resultados—, técnica —en lo que se refiere a las contramedidas y mitigaciones que protegen los datos cuando se capturan, analizan, transmiten o almacenan, económica —ya que es más que probable que surja un nuevo centro de costes en los proyectos— y regulatoria —atendiendo a la obligación de cumplir la regulación nacional e internacional—. Haciendo mención solo a este último aspecto ¿cuántos investigadores saben cómo les afecta el GDPR o la nueva LOPD y cómo deben tener estas normativas en cuenta en su labor? ¿Qué apoyo reciben en este sentido de los centros en los que investigan o de las instituciones que financian sus proyectos?

Es verdad que el marco regulatorio europeo ha previsto relajar algunas de sus premisas en el caso de los proyectos de “interés público, científico o de investigación histórica o estadística”. En este sentido, se permite a los miembros de un equipo de investigación, por ejemplo, que el consentimiento que utilizan cuando recogen datos sea amplio y no especifique de manera concreta y explícita los tratamientos que se les van a dar, ya que en las primeras fases de una investigación en muchos casos esto todavía

El uso masivo de datos puede ser una herramienta para el avance del conocimiento y una fuente de amenazas para los sujetos de las investigaciones y su privacidad al mismo tiempo

no se ha decidido. O que se almacenen los datos tras la finalización de un proyecto por si pueden ser de utilidad para otro, siempre y cuando se protejan adecuadamente. Pero esta relajación de la regulación no hace sino poner más énfasis en los límites que la propia ética de los investigadores debe poner a algunas prácticas.

Veamos algunos ejemplos de proyectos de investigación en los que la protección de datos se convierte claramente en un aspecto ético, a pesar de que pueda no parecerlo en un principio.

Hace unos años colaboramos en un proyecto de investigación con financiación pública y privada sobre eficiencia energética en el que se utilizaban técnicas de *big data* (minería de datos) que permitían analizar las lecturas de contadores inteligentes instalados en los domicilios particulares de millones de usuarios de una compañía energética. Desde el punto de vista utilitarista que se considera muchas veces en este tipo de proyectos, se valoraba el beneficio que los resultados de la investigación podían tener para el medioambiente y, por lo tanto, para la mayoría. Pero analicemos las amenazas a las que podrían haber estado sometidos los usuarios cuyos datos se tuvieron en cuenta para realizar la investigación si no se hubiera planteado el proyecto correctamente. ■■■



Pérdida de privacidad

Impactos de la pérdida de privacidad de los sujetos de una investigación según el National Institute of Standards and Technology (NIST)

1

Pérdida de libertad

2

Pérdida de autonomía o independencia

3

Exclusión

4

Daño físico

5

Estigmatización

6

Desequilibrio de poder

7

Pérdida de confianza

8

Pérdida económica

En primer lugar, pudo haber revelación bajo presiones, ya que la compañía que financiaba el proyecto ofrecía un bono descuento a los usuarios que participaran en la investigación cediendo sus datos.

En segundo lugar, revelación imprevista, ya que durante el proyecto se constató cómo a partir de los datos cedidos por los usuarios se podía predecir cuántas personas vivían en la casa, de qué género, edad y nivel socio-económico, sus hábitos de alimentación y sueño, sus horarios y hobbies, etcétera. Y los usuarios podían no saber o no comprender esto al proporcionar su consentimiento para participar en la investigación.

En tercer lugar, podría haberse observado inseguridad, ya que si los datos recogidos no se hubieran protegido adecuadamente podrían haber terminado en manos de un tercero.

En cuarto lugar, se podrían haber producido vigilancia particular o masiva a partir de los datos, ya que, aunque se anonimizaron, se ha demostrado en los últimos años que este tipo de estrategias pueden deshacerse con cierta facilidad —sobre todo, si se cuenta con otros

conjuntos de datos masivos relacionados con los mismos usuarios o clientes—.

En quinto lugar, se pudo producir apropiación y/o restricción de acceso, actuando cualquiera de las entidades o grupos de investigación participantes en el proyecto como si fueran los dueños de los datos e impidiendo a los usuarios, o poniendo trabas con falta de transparencia, ejercer sus derechos de acceso. Y en sexto lugar, se pudo producir distorsión, llegando a conclusiones equivocadas con los modelos producidos por utilizar datos con origen equivocado, incompletos u obsoletos, por mencionar solo algunos ejemplos.

El mismo tipo de amenazas las hemos observado después en proyectos de investigación en el área de la demografía basados en datos extraídos de censos electorales, en el área de la teoría de la comunicación y del análisis de las *fake news* basados en datos extraídos directamente de redes sociales, o en el área de la geografía urbana basados en datos extraídos de sensores empotrados en la red de transporte público y en aparcamientos municipales o de bicicletas.

Los principales retos a los que se enfrentan los investigadores en lo que se refiere a la privacidad y a la protección de datos en este tipo de proyectos suelen ser la naturaleza exploratoria de muchas investigaciones —cuando se comienza el proyecto todavía no se sabe muy bien qué uso se le va a dar los datos, por lo que no se suele ser muy explícito en los consentimientos ni se suele minimizar la recogida de información—, la falta de recursos y de formación y, en general, la falta de apoyo por parte de las instituciones y centros.

Pero es una obligación ética de los investigadores conocer las amenazas para los implicados, comunicarlas adecuadamente y poner medidas para evitarlas o, por lo menos, para mitigar sus impactos. Esto debe hacerse desde la fase de propuesta y planificación del proyecto, con un enfoque proactivo —no reactivo cuando ya se ha presentado los problemas, esto es muy importante—, preventivo —de nuevo, en contraposición al enfoque correctivo que suele adoptarse—, por defecto, desde el principio y centrado en el usuario/sujeto de la investigación.

Dos herramientas esenciales para poner en marcha estas buenas prácticas son el plan de gestión de datos y cumplimiento y el comité ético o asesor en cada proyecto de investigación. Ambos deben servir de guía en el cumplimiento de la regulación y de brújula en todos los aspectos éticos que puedan afectar al proyecto. Y por supuesto, la auto-regulación, basada en una educación y concienciación en profundidad gracias a asignaturas de ética en las carreras universitarias, cursos específicos en los másteres de investigación y programas de doctorado y formaciones para los investigadores principales de los proyectos y otros gestores involucrados.

Investigadores, gestores, universidades, centros de investigación y empresas tienen que ser conscientes del reto que afrontamos: los datos, por definición, son la base de cualquier investigación científica de calidad. El *big data* puede ser, por tanto, una herramienta muy potente para el avance del conocimiento en muchos campos, pero sin un adecuado cumplimiento de la regulación vigente y, sobre todo, sin unos buenos planteamientos

éticos de los proyectos de investigación, también puede ser una fuente de amenazas para los sujetos de las investigaciones y su privacidad. Seamos responsables, de nosotros depende que retornen a la sociedad los impactos positivos del tratamiento masivo de datos en ciencia sin que se sufran sus potenciales efectos negativos.

Bibliografía

- Collmann, J. y Adam Matei, S. (2016): *Ethical Reasoning in Big Data: An Exploratory Analysis*. Computational Social Sciences, Springer.
- Loukides, M., Mason, H. and Pati, DJ. (2018): *Ethics and Data Science*. O'Reilly Media.
- Richterich, A. (2018): *The Big Data Agenda: Data Ethics and Critical Data Studies* (Critical Digital and Social Media Studies Series). Londres, University of Westminster Press.
- UNESCO. Declaración sobre la ciencia y el uso del saber científico adoptada por la Conferencia Mundial sobre la Ciencia del 10 de Julio 1999. Disponible en: http://www.unesco.org/science/wcs/esp/declaracion_s.htm
- Unión Europea. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

