

Gestión de los datos brutos de investigación en los investigadores españoles en ciencias de la salud

Research raw data management in Spanish researchers on health sciences

R Aleixandre-Benavent ¹, A Vidal-Infer ², A Alonso-Arroyo ², A Ferrer-Sapena ³, F Peset ³, A García García ⁴

¹ Instituto de Gestión de la Innovación y del Conocimiento (CSIC-Universidad Politécnica de Valencia). UISYS-Universidad de Valencia, España. ² Departamento de Historia de la Ciencia y Documentación. Universitat de Valencia. ³ Departamento de Comunicación Audiovisual, Documentación e Historia del Arte. Universidad Politécnica de Valencia. ⁴ Instituto de Documentación y Tecnologías de la Información-INDOTEL. Universidad Católica San Vicente Mártir. Valencia.

Esta investigación ha sido financiada por FUNDACIÓN MAPFRE

Resumen

Objetivo: Identificar los hábitos y experiencias actuales de los investigadores españoles en ciencias de la salud en relación con la gestión y el intercambio de los datos brutos de investigación.

Material y métodos: Se diseñó un cuestionario compuesto por 40 preguntas distribuidas en tres bloques: A) Datos personales; B) Creación y reutilización de los datos; C) Preservación de los datos. El cuestionario se envió por correo electrónico a investigadores de varias instituciones sanitarias españolas.

Resultados: La mayor parte de los investigadores españoles de las ciencias de la salud carecen de políticas o planes de gestión de los datos. Los principales motivos por los que no existen estas políticas son la ausencia de una obligación institucional, la falta de tiempo, el esfuerzo que se requiere para llevarlas a cabo y la falta de apoyo y de orientación. Desconocen la existencia de repositorios, instituciones y otros sistemas de gestión y preservación de los datos. A una amplia mayoría les gustaría utilizar los datos de investigación recogidos por otros investigadores.

Conclusión: Es necesario adquirir una conciencia colectiva de la importancia de compartir datos como práctica que mejora la transmisión del conocimiento y, en definitiva, la eficiencia de las inversiones en investigación. También es necesario que las agencias e instituciones que financian la investigación proporcionen el apoyo necesario para llevarla a cabo.

Palabras clave:

Datos brutos de investigación, datos abiertos, compartir datos, ciencias de la salud.

Abstract

Objective: To identify the habits and current experiences of Spanish researchers in health sciences in relation to the management and sharing of raw research data.

Material and method: A questionnaire composed of 40 questions divided into three blocks was designed: A) Personal information; B) Creation and reuse of data; C) Preservation of Data. The questionnaire was sent by email to researchers from several Spanish healthcare institutions.

Results: Most of the Spanish health science researchers lack policies or plans of management data. The main reasons why there are no such policies is the absence of an institutional duty, lack of time, the effort required to carry it out and lack of support and guidance. They are unaware of repositories, institutions and other systems of management and data preservation. A vast majority would like to use the research data collected by other researchers.

Conclusion: It is necessary to acquire a collective awareness of the importance of data sharing as a practice that improves the transmission of knowledge and, ultimately, the efficiency of research investments. It is also required that the agencies and institutions that fund research provide the necessary support to carry it out.

Key words:

Raw research data, open data, data sharing, health sciences.

Correspondencia

R Aleixandre-Benavent
Plaza Cisneros, 4. 46003 Valencia, España.
E-mail: Rafael.aleixandre@uv.es

Introducción

La investigación científica puede mejorarse mediante la gestión y el intercambio de los datos de la investigación. El uso compartido de los datos permite generar nuevos conocimientos, formular nuevas hipótesis y verificar con certeza los resultados [1][2]. Los principios en los que se basa están teniendo cada vez un mayor respaldo por parte de las agencias financiadoras de la investigación, dada la creciente preocupación por obtener el mayor retorno posible de la inversión, tanto en términos de la calidad de los productos derivados de la investigación como en la reutilización de datos [3-5].

Los datos de investigación son un recurso con un gran valor que, por lo general, requieren mucho tiempo y dinero para ser producidos. Muchos datos tienen un valor significativo fuera del uso de la investigación inicial. Las virtudes de los datos abiertos y las ventajas de compartirlos están hoy en día firmemente establecidas en la literatura, a pesar de que su implementación sea todavía limitada [4][6-9].

Uno de los aspectos que debe investigarse en el estadio inicial en el que nos encontramos es conocer cuáles son las prácticas actuales de los investigadores españoles en este proceso. El objetivo de este trabajo es identificar los hábitos y experiencias actuales de los investigadores españoles en ciencias de la salud en relación con la gestión y el intercambio de los datos brutos de investigación, así como sus motivaciones para compartir, las circunstancias que lo impiden o dificultan y las necesidades de apoyo técnico y de infraestructuras para su puesta en práctica.

Material y métodos

Para conocer las prácticas de los investigadores que trabajan en centros de investigación españoles en biomedicina se diseñó un cuestionario compuesto por 40 preguntas. El cuestionario se basó en otros ya validados en otras instituciones, convenientemente adaptado a las particularidades del caso español. Los cuestionarios se enviaron en noviembre de 2013 por correo electrónico a investigadores de varias instituciones sanitarias españolas, entre ellas, el Hospital Severo Ochoa (Leganés, Madrid), la Facultad de Medicina de Valencia, el Hospital Clínico de Valencia y el Hospital General de Castellón. Las preguntas del cuestionario se dividieron en tres bloques: A) Datos personales; B) Creación y reutilización de los datos; C) Preservación de los datos. En el apartado A, sobre datos personales, se identificaban las siguientes variables: edad, sexo, nombre del centro de investigación, área, puesto y años de investigación. En el apartado B sobre Creación y reutilización de los datos se investigaba la existencia de políticas o planes de gestión

de los datos en la institución; los motivos por los que se desarrolló esta política o por los que no existe; quién es la persona en el grupo que tiene la responsabilidad de gestionar los datos; dónde se almacenan y de qué tipo de datos se hacen copias de seguridad. También se preguntaba a los investigadores si les gustaría utilizar los datos de investigación no publicados de otros investigadores con los que no colabora y qué medios utiliza para localizarlos; qué datos externos suelen utilizar; si alguna vez le han solicitado los datos otros investigadores o si le han solicitado que estén en abierto; cuáles son los problemas o prejuicios que tiene para compartir. El apartado C estaba destinado a indagar sobre las prácticas de preservación de los datos una vez terminado el proyecto. Se preguntaba sobre las amenazas que consideran que pueden derivarse de compartir los datos; cómo los almacenan para preservar su acceso y utilización futura; dónde es más adecuado que se preserven; la existencia de planes en la institución para la creación de un repositorio de datos para su almacenamiento y preservación; la existencia en su disciplina de guías o recomendaciones para preservar datos; la importancia que tiene la preservación y, por último, si la institución presta alguna ayuda para llevar a cabo la gestión de los datos en todo su ciclo.

Resultados

De los 155 profesionales que respondieron el cuestionario, 97 eran hombres (62,6%) y 58 mujeres (37,4%). La distribución de los 155 respondientes según la función habitual desempeñada en los equipos de trabajo en los últimos años ha sido la siguiente: el 44,5% eran investigadores principales, el 32,2% investigadores colaboradores, el 16,1% investigadores *senior*, el 4,5% asistentes o técnicos de investigación y el 1,3% estaban realizando estudios de doctorado e eran investigadores en formación. Según el número de años que llevan investigando, la distribución ha sido la siguiente: el 13% llevaba menos de cinco años; el 19,3% de cinco a 10 años; el 30,3% de 11 a 20 años; el 37,4% más de 20 años.

La pregunta relativa a si en los proyectos de investigación en los que ha participado el investigador existe alguna política o plan de gestión de los datos (implícita o por escrito) ha obtenido los siguientes resultados (Figura 1): la repuesta más frecuente ha sido que no se dispone de ellas (22,8%); el 15,4% de los encuestados respondió sobre la existencia de políticas de «Normas de anonimización y confidencialidad»; el 12,2% indicaba «Unidades donde se almacenan para el uso del grupo de investigación»; el 11,6% marcó «Niveles de autorización para acceder y modificar datos».



Fig. 1. Política o plan de gestión de los datos.

Respecto a los motivos por los que no existe esa política (Figura 2), el 25,1% respondió que era debido a «la ausencia de una obligación institucional de gestión de los datos» y el 17,8% «por el tiempo y el esfuerzo requerido». Otras respuestas cercanas al 13% y al 14% fueron: «la falta de apoyo u orientación», «no requerido por el ente financiador del proyecto» y «no requerido para el campo de investigación». El 13,6% indicaba que no lo sabía o no podía contestar a la pregunta.

Respecto al lugar de almacenamiento de los datos (Figura 3), el 43,2% lo hace en ordenadores locales de los investigadores, el 30,8% en unidades de almacenamiento portables, el 15,6% en servidores de red institucionales y el 8,4% en un servicio de almacenamiento externo.

Ante la pregunta de si les gustaría utilizar datos de investigación recogidos por otros investigadores, las respuestas han sido afirmativas mayoritariamente, con los porcentajes siguientes: 50% en los estudiantes de doctorado, 75,3% en los investigadores principales, 85,7% en los asistentes y técnicos de investigación, 88% en los investigadores *senior*, 92% en los investigadores colaboradores y 100% en los investigadores en formación.

La pregunta «Utiliza actualmente datos de investigación no publicados de otros investigadores con los que no colabora» tenía varias posibilidades de respuesta (Tabla 1). La mayor parte de los respondientes (64,7%) manifestaron que «mis datos están disponibles para mi grupo de investigación y/o colegas colaboradores en la investigación». Las otras respuestas obtuvieron porcentajes mucho menores que fueron del 9,8% en la respuesta «No comparto mis datos, pero me gustaría hacerlo en un futuro»; 9,2% en la res-

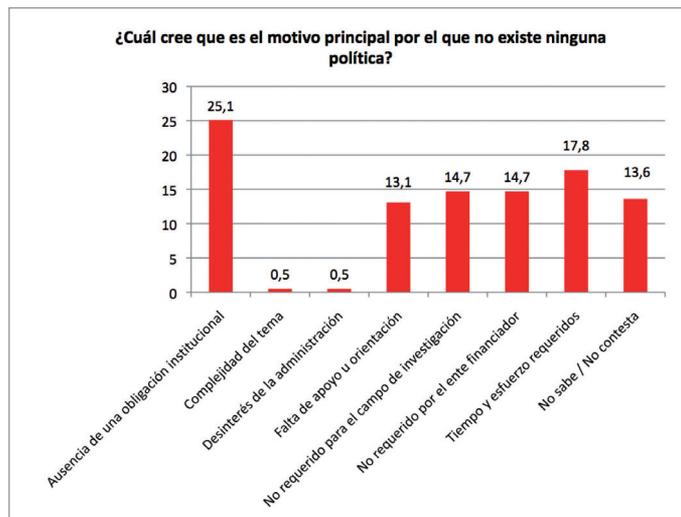


Fig. 2. Motivo principal por el que no existe ninguna política de gestión de los datos de investigación.

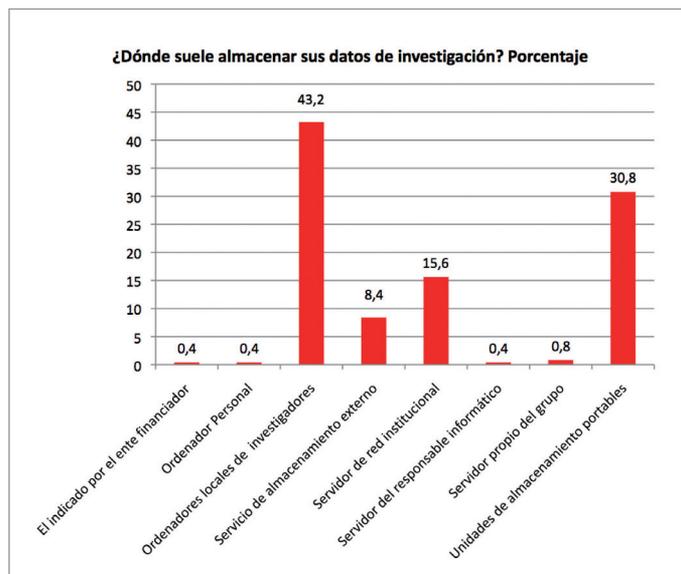


Fig. 3. Dispositivo de almacenamiento de los datos de investigación.

puesta «Mis datos podrían estar disponibles con los cambios apropiados (por ejemplo, datos clínicos anónimos o con la obligación de citarlos)»; 6,9% en la respuesta «Mis datos están disponibles en abierto para todo el mundo»; y 6,4% en la respuesta «Mis datos están disponibles en abierto para mi disciplina científica». Solo un 2,3% respondió que no compartía sus datos y que no quería hacerlo en el futuro y apenas un 0,6% respondió que sus datos estaban disponibles mediante el pago de una tasa.

La pregunta «Qué le preocupa a la hora de compartir los datos en un servicio externo» (Tabla 2) ha obtenido las siguientes respuestas, en orden decreciente: el 25,3% adu-

Tabla 1. Utilización de los datos de otros investigadores

¿Utiliza actualmente datos de investigación no publicados de otros investigadores con los que no colabora?

Función desempeñada	Mis datos están disponibles en abierto para mi disciplina científica	Mis datos están disponibles en abierto para todo el mundo	Mis datos están disponibles mediante el pago de una tasa	Mis datos están disponibles para mi grupo de investigación y/o colegas colaboradores en la investigación	Mis datos podrían estar disponibles con los cambios apropiados (por ejemplo, datos clínicos anónimos o con la obligación de citarlos)	No comparto mis datos y no quiero hacerlo en el futuro	No comparto mis datos, pero me gustaría hacerlo en un futuro	Total
Asistente/Técnico de investigación	1 (12,5%)	1 (12,5%)	0	4 (50%)	1 (12,5%)	0	1 (12,5%)	8 (4,6%)
Estudiante de doctorado	0	0	0	2 (100%)	0	0	0	2 (1,2%)
Investigador colaborador	3 (5,4%)	3 (5,4%)	1 (1,8%)	38 (67,7%)	3 (5,4%)	2 (3,6%)	6 (10,7%)	56 (32,4%)
Investigador en formación	1 (33,3%)	0	0	1 (33,3%)	0	0	1 (33,3%)	3 (1,7%)
Investigador principal	4 (5,3%)	6 (7,9%)	0	47 (61,8%)	9 (11,8%)	1 (1,3%)	9 (11,8%)	76 (43,9%)
Investigador <i>senior</i>	2 (7,1%)	2 (7,1%)	0	20 (71,4%)	3 (10,7%)	1 (3,6%)	0	28 (16,2%)
Total	11 (6,4%)	12 (6,9%)	1 (0,6%)	112 (64,7%)	16 (9,2%)	4 (2,3%)	17 (9,8%)	173 (100%)

ce «cuestiones legales: confidencialidad y derechos de propiedad intelectual»; con un porcentaje cercano al anterior, 24,4%, «mal uso o interpretación de los datos»; la «pérdida de autoría» preocupa al 14,2%; el «miedo a perder la delantera en investigación» al 11,6%. Otras cuestiones como la «confidencialidad de datos del paciente», «gastar tiempo haciéndolos disponibles» y «pérdida de autoría» son asuntos que no preocupan, a juzgar por el bajo número de encuestados que respondieron afirmativamente a estas cuestiones.

Las siguientes preguntas indagaban sobre las amenazas que perciben los investigadores relacionadas con los datos. En la Tabla 3 se presentan las respuestas a una de las preguntas de este apartado sobre si «los usuarios pueden ser incapaces de comprender o utilizar los datos (semántica, formatos o algoritmos implicados)». Como se puede apre-

ciar, no existe unanimidad en estas percepciones y los resultados son dispares, aunque por lo general no se percibe este riesgo, ya que predominan los porcentajes en las valoraciones que perciben la amenaza como menor.

Dentro del capítulo sobre la preservación de los datos, es interesante conocer cómo se almacenan los datos de investigación para preservar su acceso y utilización futura (Figura 4). Más de la mitad de los encuestados (51,6%) respondió que los almacena en su ordenador personal, el 21,6% contestó que lo hace en el servidor de la organización y el 9,3% en un servicio web externo como Google Drive o Dropbox. Como se aprecia en la figura, el uso de servicios externos especializados es utilizado por un porcentaje muy bajo de los encuestados (1,6%), y lo mismo ocurre con otras posibilidades.

Ante la pregunta de «Dónde considera más adecuado

Tabla 2. Preocupaciones para compartir datos en servicios externos

¿Qué le preocupa a la hora de compartir los datos en un servicio externo?

Función desempeñada	Confidencialidad de datos del paciente	Cuestiones legales: confidencialidad / derechos de propiedad intelectual	Falta de interés para otros	Mal uso o interpretación de los datos	Miedo a perder la delantera en investigación	No conozco archivos digitales (repositorios o centros de documentación) a los que pueda enviar datos	No creo que mis datos de investigación estén seguros en un centro de documentación, web de una revista o repositorio	No se prevén problemas	No se prevén problemas, (no perder) tiempo haciéndolos disponibles	No tengo seguridad de cómo fueron recogidos	Perder tiempo haciéndolos disponibles	Pérdida de autoría	Pérdida de autoría, hay quien aspira a tener datos sin trabajarlos...	Total
Asistente/Técnico de investigación	0	3 (27,3%)	0	3 (27,3%)	0	0	0	2 (18,2%)	0	0	0	3 (27,3%)	0	11 (3,2%)
Estudiante de doctorado	0	1 (25%)	1 (25%)	1 (25%)	0	0	0	0	0	0	1 (25%)	0	0	4 (1,2%)
Investigador colaborador	1 (0,9%)	31 (29,2%)	0	27 (25,5%)	14 (13,2%)	4 (3,8%)	5 (4,7%)	0	0	0	11 (10,4%)	13 (12,3%)	0	106 (30,8%)
Investigador en formación	0	1 (20%)	0	1 (20%)	0	1 (20%)	0	0	0	0	0	2 (40%)	0	5 (1,4%)
Investigador principal	0	36 (23,1%)	0	38 (24,4%)	16 (10,3%)	10 (6,4%)	9 (5,8%)	5 (3,2%)	1 (0,6%)	1 (0,6%)	17 (10,9%)	22 (14,1%)	1 (0,6%)	156 (45,3%)
Investigador senior	0	15 (24,2%)	0	14 (22,6%)	10 (16,1%)	1 (1,6%)	2 (3,2%)	0	0	0	11 (17,7%)	9 (14,5%)	0	62 (18%)
Total	1 (0,3%)	87 (25,3%)	1 (0,3%)	84 (24,4%)	40 (11,6%)	16 (4,6%)	16 (4,6%)	7 (2%)	1 (0,3%)	1 (0,3%)	40 (11,6%)	49 (14,2%)	1 (0,3%)	344 (100%)

que se preserven y se compartan los datos de las investigaciones» (Figura 5), las respuestas cambian claramente respecto a la pregunta anterior, de manera que, en este caso, la consideración de preservación en el ordenador personal solo es del 1,7%. La mayoría de los respondientes indicaron que el repositorio de la organización es el lugar apropiado (36,4%), el 23,9% opinó que debería ser el servicio externo especializado en su disciplina, el 17,9% en un servicio especializado en datos de tipo generalista y el 16,8% en una editorial.

Un hecho que llama mucho la atención es la falta de información de los usuarios sobre la existencia de instalaciones que permitan preservar los datos digitales. Un 82,6% de los respondientes indicaron que ignoraban su existencia

frente a un 10,3% que señalaron que conocían su existencia (Figura 6).

Al solicitarles que priorizaran en una escala de 1 a 7 la afirmación «Si la investigación está financiada públicamente, los resultados deberían convertirse en propiedad pública y, por tanto, estar debidamente preservados» (Tabla 4), la mayor parte de los investigadores principales de los proyectos le dieron la máxima prioridad (45%), mientras que en los investigadores *senior* y en los colaboradores, esta prioridad fue algo menor. Ningún investigador principal dio la mínima prioridad a este aspecto, pero sí un investigador *senior* y otro colaborador.

En otra serie de preguntas se indagaba sobre las razones para preservar los datos y se plantearon las siguientes cuestio-

Tabla 3. Las restricciones de acceso y propiedad intelectual como amenaza a los datos

Califique del 1 al 7 las siguientes amenazas A SUS DATOS para los próximos 10 años

Función desempeñada	[Las restricciones de acceso y uso (propiedad intelectual)]							Total
	1	2	3	4	5	6	7	
Asistente/Técnico de investigación	2 (28,6%)	0	0	3 (42,9%)	0	2 (28,6%)	0	7 (4,7%)
Estudiante de doctorado	1 (50%)	1 (50%)	0	0	0	0	0	2 (1,3%)
Investigador colaborador	3 (6,5%)	6 (13%)	6 (13%)	5 (10,9%)	9 (19,6%)	12 (26,1%)	5 (10,9%)	46 (30,7%)
Investigador en formación	0	0	0	0	0	0	1 (100%)	1 (0,7%)
Investigador principal	5 (7,2%)	7 (10,1%)	5 (7,2%)	14 (20,3%)	14 (20,3%)	10 (14,5%)	14 (20,3%)	69 (46%)
Investigador senior	0	2 (8%)	3 (12%)	2 (8%)	3 (12%)	9 (36%)	6 (24%)	25 (16,7%)
Total	11 (7,3%)	16 (10,7%)	14 (9,3%)	24 (16%)	26 (17,3%)	33 (22%)	26 (17,3%)	150 (100%)



Fig. 4. Almacenamiento de los datos de investigación.

nes: a) Estimula el avance de la ciencia basada en el conocimiento existente; b) Sirve para validar las investigaciones; c) Permite el re-análisis de datos existentes; d) Puede estimular colaboraciones interdisciplinarias; e) Tiene valor económico potencial; e) Son únicos e irremplazables. La mayor parte de los investigadores respondieron positivamente ante estas afirmaciones. La valoración fue más positiva en las cuatro primeras aseveraciones, mientras que obtuvo una menor valoración la referida al valor económico potencial de los datos.

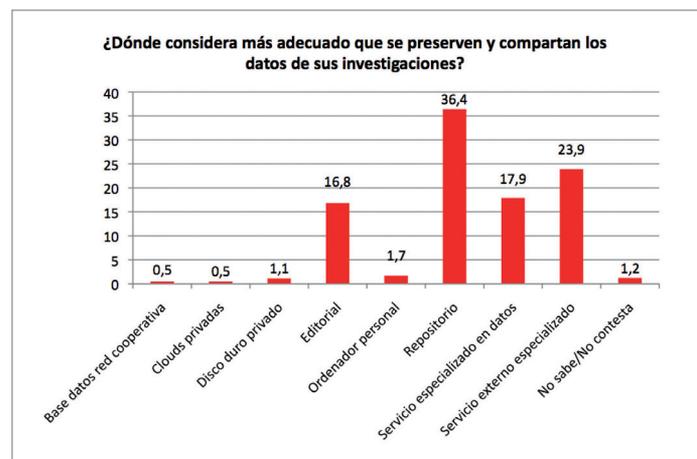


Fig. 5. Lugar adecuado de preservación de los datos de investigación.

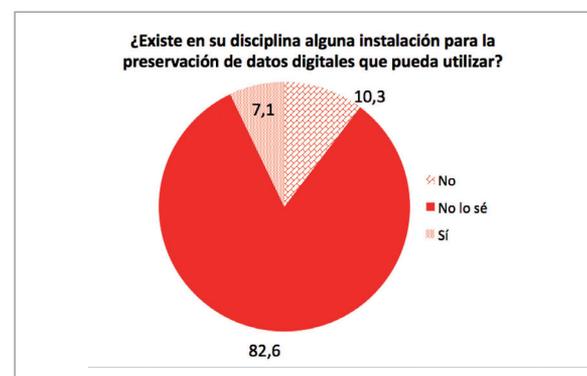


Fig. 6. Existencia de instalaciones para la preservación.

Tabla 4. La financiación de la investigación pública como razón para preservar datos

Priorice del 1 al 7 la importancia de las siguientes razones para preservar datos

Función desempeñada	[Si la investigación está financiada públicamente, los resultados deberían convertirse en propiedad pública y, por tanto, estar debidamente preservados]							Total
	1	2	3	4	5	6	7	
Asistente/Técnico de investigación	2 (28,6%)	0	0	3 (42,9%)	0	2 (28,6%)	0	7 (4,7%)
Estudiante de doctorado	1 (50%)	1 (50%)	0	0	0	0	0	2 (1,3%)
Investigador colaborador	3 (6,5%)	6 (13%)	6 (13%)	5 (10,9%)	9 (19,6%)	12 (26,1%)	5 (10,9%)	46 (30,7%)
Investigador en formación	0	0	0	0	0	0	1 (100%)	1 (0,7%)
Investigador principal	5 (7,2%)	7 (10,1%)	5 (7,2%)	14 (20,3%)	14 (20,3%)	10 (14,5%)	14 (20,3%)	69 (46%)
Investigador <i>senior</i>	0	2 (8%)	3 (12%)	2 (8%)	3 (12%)	9 (36%)	6 (24%)	25 (16,7%)
Total	11 (7,3%)	16 (10,7%)	14 (9,3%)	24 (16%)	26 (17,3%)	33 (22%)	26 (17,3%)	150 (100%)

Discusión

Este trabajo ha permitido conocer los hábitos, actitudes, motivaciones y amenazas que perciben los investigadores españoles en ciencias de la salud sobre la gestión y preservación de los datos brutos de sus investigaciones.

Los organismos financiadores de la investigación siguen cada vez más la orientación de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) que propone que los datos en los que se basan las investigaciones financiadas con fondos públicos deben, en lo posible, estar en abierto y ser puestos a disposición de la comunidad [10]. En junio de 2013, la carta pública de los líderes del G8 indicaba: «Los datos abiertos son un recurso sin explotar con un enorme potencial para fomentar la construcción de sociedades más fuertes, más interconectadas, que satisfagan mejor las necesidades de nuestros ciudadanos y permitir que florezca la innovación y la prosperidad». Por su parte, Horizon 2020, el VII Programa Marco de la Unión Europea, prevé que en el acuerdo de subvención de los proyectos financiados por el Consejo Europeo de Investigación se podrán establecer las condiciones en las que se deberá ofrecer el acceso abierto a los resultados de los proyectos [11].

Existen numerosas iniciativas mundiales en este campo y también en España se están llevando a cabo proyectos destinados a fomentar el intercambio y reutilización de los

datos primarios de investigación. En España, el grupo de trabajo de «Depósito y Gestión de Datos en Acceso Abierto» del proyecto Recolecta [12] ha elaborado un informe que refleja los aspectos relevantes que intervienen en la gestión de los datos de investigación en los repositorios científicos. Digital CSIC, el repositorio de publicaciones y de datos del Consejo Superior de Investigaciones Científicas, también pone a disposición de los investigadores una guía de buenas prácticas y políticas de datos de investigación [13]. Por otra parte, el Ministerio de Economía y Competitividad financia el proyecto «Opendatascience, centro de recursos para la preservación y gestión de datos abiertos de investigación» (<http://www.datasea.es/dt/>), mientras que FUNDACIÓN MAPFRE sufragó el proyecto «Estrategias para la gestión y la promoción del uso compartido de datos científicos sobre la salud» (<http://www.healthdatasea.es/>). Entre las acciones de estos proyectos destacan las campañas de sensibilización mediante la edición de guías de buenas prácticas destinadas a convencer a autores y editores de la conveniencia de liberar y compartir los datos brutos de investigación para un aprovechamiento más óptimo. Se hace hincapié en la necesidad de poner al servicio de los investigadores el asesoramiento técnico y las infraestructuras necesarias para que puedan llevar a cabo estas prácticas, así como la elabora-

ción de una normativa que regule todo este proceso para evitar el uso indebido o fraudulento.

En un estudio elaborado por investigadores que en 2008 habían publicado artículos en la revista *Annals of Internal Medicine*, se les preguntó bajo qué condiciones pondrían sus protocolos, datos estadísticos y datos a disposición de otros, y solamente el 13% manifestó que sus protocolos estaban disponibles sin ninguna condición, el 3% dijo que estaban disponibles los datos estadísticos y el 4% los datos. Este porcentaje es superior al encontrado en nuestro estudio, que se sitúa cerca del 7%. Por otra parte, la mayoría de los autores que indicaron que dichos materiales estaban disponibles requerían que los interesados contactaran con ellos primero, y muchos establecieron condiciones específicas para proceder al intercambio de estos materiales [14]. En definitiva, los investigadores estaban más dispuestos a compartir plenamente sus protocolos y menos dispuestos a compartir datos. Los investigadores españoles coinciden con los de otros estudios que han mostrado que la falta de financiación y la falta de tiempo para destinarlo a la gestión y preservación de los datos son los principales obstáculos para poner los datos a disposición de otros [15]. Por otra parte, se pone de manifiesto que, para proceder a la apertura de datos y para compartirlos, es necesario establecer tempranamente un plan de gestión de los datos en el que se describa todo el proceso [16].

Aunque la responsabilidad de la gestión de datos recae principalmente en los investigadores, las instituciones deben proporcionar la orientación, las herramientas, la infraestructura y el personal de apoyo necesario para ayudar en muchas tareas relacionadas con la gestión de los datos.

La mayor parte de los investigadores españoles de las ciencias de la salud carecen de políticas o planes de gestión de los datos. Los principales motivos por los que no existen estas políticas son la ausencia de una obligación institucional, la falta de tiempo, el esfuerzo que se requiere para llevarlas a cabo y la falta de apoyo y de orientación. Desconocen la existencia de repositorios, instituciones y otros sistemas de gestión y preservación de los datos y, por lo tanto, no utilizan servicios externos especializados como Dryad, ArrayExpress, Figshare o DataOne. Por otra parte, suelen utilizar ordenadores locales o memorias digitales portátiles como sistemas de almacenamiento y depósito. Sin embargo, a una amplia mayoría les gustaría utilizar los datos de investigación recogidos por otros investigadores. La principal amenaza a sus datos es la restricción de acceso y uso relacionada con la propiedad intelectual.

Es necesario adquirir una conciencia colectiva de la importancia de compartir datos como práctica que mejore la

transmisión del conocimiento y, en definitiva, la eficiencia de las inversiones en investigación. También es necesario un cambio en la cultura de la ciencia que comprometa no solo a los investigadores, sino que también requiera compromisos importantes de las agencias e instituciones que financian la investigación. Se podría ampliar la actual ley de la ciencia de manera que no solamente promueva el acceso abierto a las investigaciones financiadas con fondos públicos, sino también el depósito de los datos originarios en repositorios especiales que permitan que estos sean compartidos y, potencialmente, reutilizables [17][18].

Como conclusión, de los resultados de este estudio se deduce que es necesario realizar campañas de sensibilización dirigidas a los profesionales para que liberen sus datos de investigación, pero también crear infraestructuras que faciliten su depósito y preservación, y definir las reglas de juego que aclaren muchas de las implicaciones que una estrategia de estas características supondría. En opinión de los respondientes, deben ser los organismos oficiales los impulsores de este movimiento. ■

REFERENCIAS BIBLIOGRÁFICAS

1. Alsheikh-Ali AA, Qureshi W, Al-Mallah MH, Ioannidis JPA. Public availability of published research data in high-impact journals. *PLoS ONE*. 2011; 6:24357.
2. Piwowar HA, Chapman WW. Public sharing of research datasets: a pilot study of associations. *J Informetrics* 2010; 4:148-56.
3. Gotzsche PC. Strengthening and opening up health research by sharing our raw data. *Circulation: Cardiovascular Quality and Outcomes*. 2012; 5:236-7.
4. González LM, Saorín T, Ferrer A, Aleixandre-Benavent R, Paset F. Gestión de datos de investigación: infraestructuras para su difusión. *Profesional de la Información*. 2013; 22:414-23.
5. Guttmacher AE, Nabel EG, Collins FS. Why data-sharing policies matter. *Proc Nat Acad Sci USA* 2009; 106:16894.
6. Murray-Rust, Peter. Open data in science. [citado 14 Feb. 2014]. Disponible en: <http://hdl.handle.net/10101/npre.2008.1526.1>.
7. Molloy JC. The open knowledge foundation: open data means better science *PLoS Biology* 2011; 9:1001195.
8. Pisani E, AbouZahr C. Sharing health data: good intentions are not enough. *Bull WHO* 2010; 88:462-6.
9. Krumholz HM. Open science and data sharing in clinical research: basing informed decisions on the totality of the evidence. *Circulation: Cardiovascular Quality and Outcomes* 2012; 5:141-2.

10. Organisation for Economic Co-operation and Development (2007). OECD principles and guidelines for access to research data from public funding. [página web]. [citado 8 enero 2014]. Disponible en: www.oecd.org/dataoecd/9/61/38500813.pdf.
11. Horizon 2020. [página web]. [citado 28 Feb. 2014]. Disponible en: <http://ec.europa.eu/programmes/horizon2020/>.
12. Grupo de Trabajo de «Depósito y Gestión de datos en Acceso Abierto» del proyecto Recolecta. La conservación y reutilización de los datos científicos en España. Informe del grupo de trabajo de buenas prácticas. Madrid: Fundación Española para la Ciencia y la Tecnología, FECYT (2012) [citado 14 Ene. 2014]. Disponible en: http://www.recolecta.net/buscador/documentos/informe_datos_cientificos_en_esp.pdf.
13. Digital CSIC. Ciencia en abierto [página web]. [citado 27 Feb. 2014]. Disponible en: <http://digital.csic.es/politicas/politicaDatos.jsp>.
14. Lane A, McAndrew P. Are open educational resources systematic or systemic change agents for teaching practice? *Brit J Educ Technol* 2010; 41:952-62.
15. Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu E, Read E, *et al.* Data sharing by scientists: practices and perceptions. *PLoS ONE* 2011; 6:21101.
16. Aleixandre-Benavent R, Ferrer Sapena A, Alonso-Arroyo A, Peset F. Compartir datos de investigación en cardiología. *Rev Esp Cardiol* 2013; 66:1007-8.
17. Aleixandre-Benavent R, Vidal-Infer A, Alonso-Arroyo A, González de Dios J, Ferrer Sapena A, Peset F. Disponibilidad en abierto de los artículos y de los datos brutos de investigación en las revistas pediátricas españolas. *Anales de Pediatría* 2013; [en prensa]. Disponible en: <http://dx.doi.org/10.1016/j.anpedi.2013.11.014>.
18. Nina-Alcocer V, Blasco Gil Y, Peset F. Datasharing: guía práctica para compartir datos de investigación. *El profesional de la información* 2013; 22:562-8.

Conflicto de intereses

Los autores hemos recibido ayuda económica de FUNDACIÓN MAPFRE para la realización de este trabajo. No hemos firmado ningún acuerdo por el que vayamos a recibir beneficios u honorarios por parte de alguna entidad comercial o de FUNDACIÓN MAPFRE.