

Red temática Maredata

Seminario sobre gestión de los datos de investigación

La opinión y los hábitos de los investigadores sobre los datos de investigación.

Antonia Ferrer y Rafael Aleixandre

Con la colaboración de: Fernanda Peset , Antonio Vidal-Infer y Adolfo Alonso-Arroyo

www.datasea.es

¿Por qué compartir los datos de investigación?

- Aumentan las posibilidades de trabajar con grandes volúmenes de datos y de hacer análisis estadísticos ciegos.
- Estimula los descubrimientos adicionales
- Permite establecer nuevas colaboraciones entre los usuarios de los datos y aquellos que los generaron
- Permite la comprobación, previniendo los fraudes: transparencia

¿Por qué compartir los datos de investigación?

- Fomenta la validación de los métodos de investigación
- Evita la repetición de costosos proyectos que utilizan las mismas intervenciones que se utilizaron previamente.
- Aumenta la visibilidad y el impacto de la investigación.
- Aporta recursos para la educación y la formación

Objetivo

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

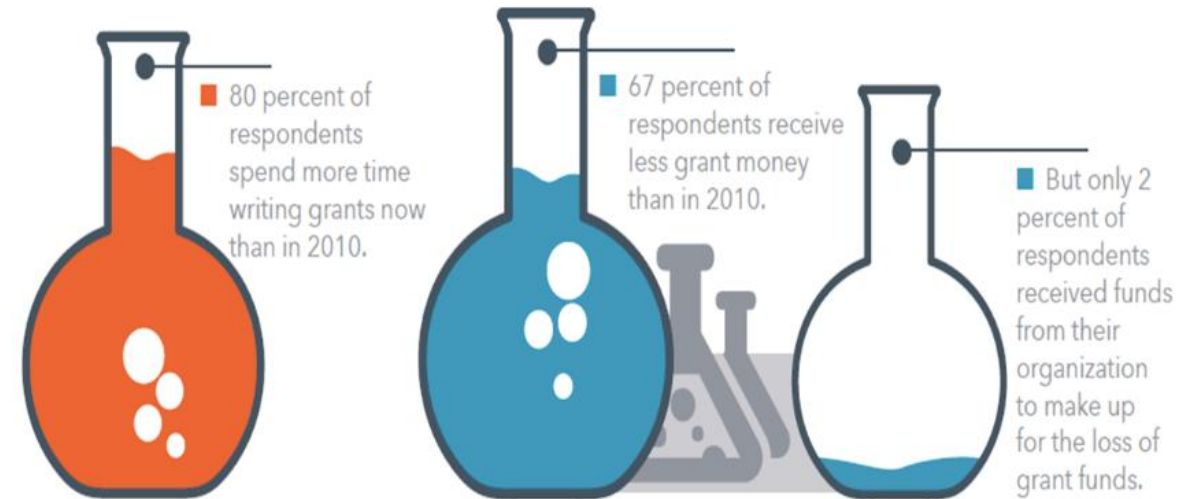


<http://diario16.pe/noticia/53666-video-cientificos-espanoles-bailan-para-financiar-investigaciones-biomedicas>

Método

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.



<http://lacienciaysdemonios.com/2013/09/20/el-principio-del-fin-de-la-hegemonia-cientifica-norteamericana/>

Método

El cuestionario se envió por correo electrónico a investigadores de varias instituciones españolas

Este cuestionario se inscribe en un proyecto de investigación I+D+i cuyo objetivo es fomentar el intercambio de los datos científicos en formato digital.

Entendemos por "datos de investigación todo aquel material que ha sido registrado durante la investigación, reconocido por la comunidad científica y que sirve para certificar los resultados de la investigación que se realiza. Además de estas características también se especifica que debe provenir de una fuente única y deben ser difíciles o imposibles de obtener de nuevo" (Torres-Salinas; Robinson-García; Cabezas-Clavijo, 2012). Ejemplos: vídeos, gráficos, tablas, rawdata, datos numéricos...

Por reutilización de datos entendemos la posibilidad de crear nuevo conocimiento a partir de datos existentes, combinándolos de nuevas maneras, no imaginadas para su forma original. En un sentido estricto son más reutilizables los datos en bruto que un vídeo, por ejemplo.

Solicitamos su colaboración rellenando la encuesta que figura a continuación, que consiste en 40 preguntas divididas en 3 apartados y debería durar entre 10 y 15 minutos.

Siendo conscientes de la multitud de encuestas que llegan por correo, agradecemos que dedique unos minutos a la nuestra

* Los resultados de esta encuesta serán compartidos con proyectos similares

*Obligatorio



A. Datos personales

1. Edad *

2. Sexo *

- Hombre
 Mujer

3. Escoja de la lista la Universidad o Centro de Investigación donde haya realizado la mayor parte de su trabajo científico en los últimos 5 años *

- Universitat Politècnica de València

Método

Variables del apartado A:

- edad,
- sexo,
- nombre del centro de investigación,
- área,
- puesto
- años de investigación.

Método

Variables del apartado B:

- Existencia de planes de gestión de los datos en el centro;
- Motivos por los que se desarrolló o por los que no existe
- Dispositivo de almacenamiento y de copias de seguridad;
- Si les gustaría utilizar los datos de otros investigadores;
- Qué datos externos suelen utilizar;
- Si alguna vez le han solicitado los datos otros investigadores;
- Problemas o prejuicios que tiene para compartir.

Método

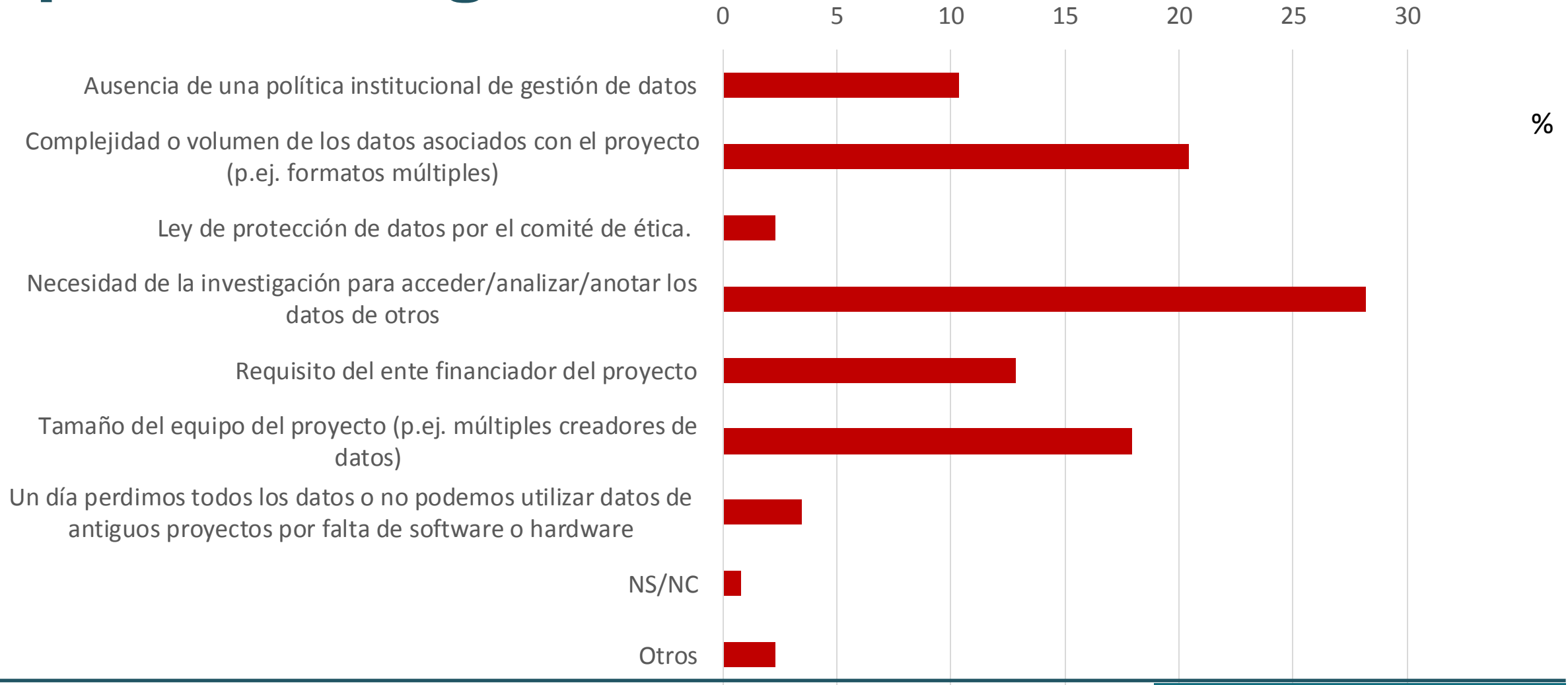
- **Variables del apartado C:**
 - Prácticas de preservación de los datos;
 - Amenazas que pueden derivarse de compartir los datos;
 - Sistema de almacenamiento de los datos para su conservación y utilización en el futuro;
 - Dispositivos empleados para su preservación;
 - Planes para el almacenamiento y preservación;
 - Ayudas para llevar a cabo la gestión de los datos.

Método

El cuestionario se basó en otros ya validados en otras instituciones, convenientemente adaptado a las particularidades del caso español:

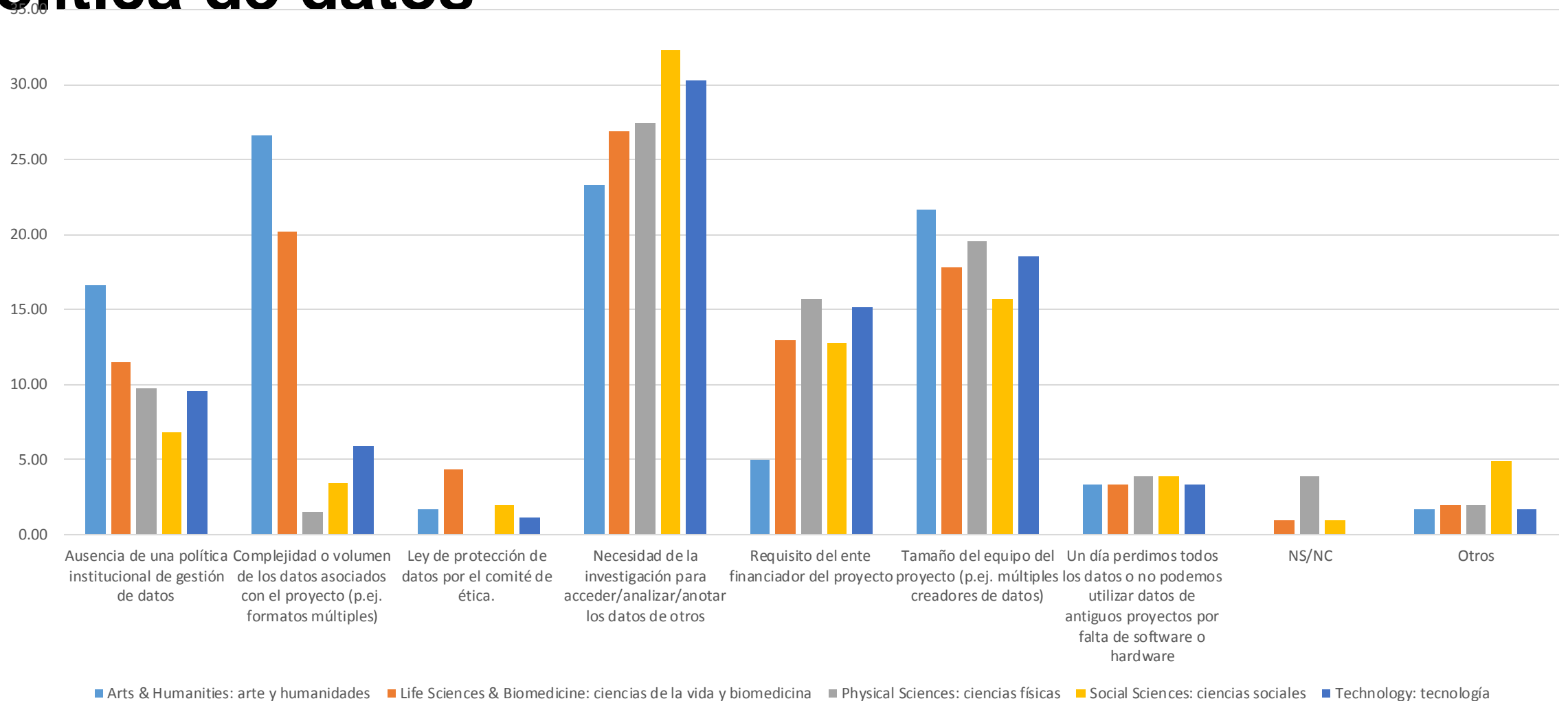
- PARSE insight;
- University of Southampton Questionnaire;
- Imperial College of London DAF Survey Questionnaire;
- Edinburgh Data Audit Implementation Online Questionnaire;
- Data Asset Framework;
- ANU Data Management Manual.
- CSIC

Motivos por los que se desarrolla una política de gestión de los datos



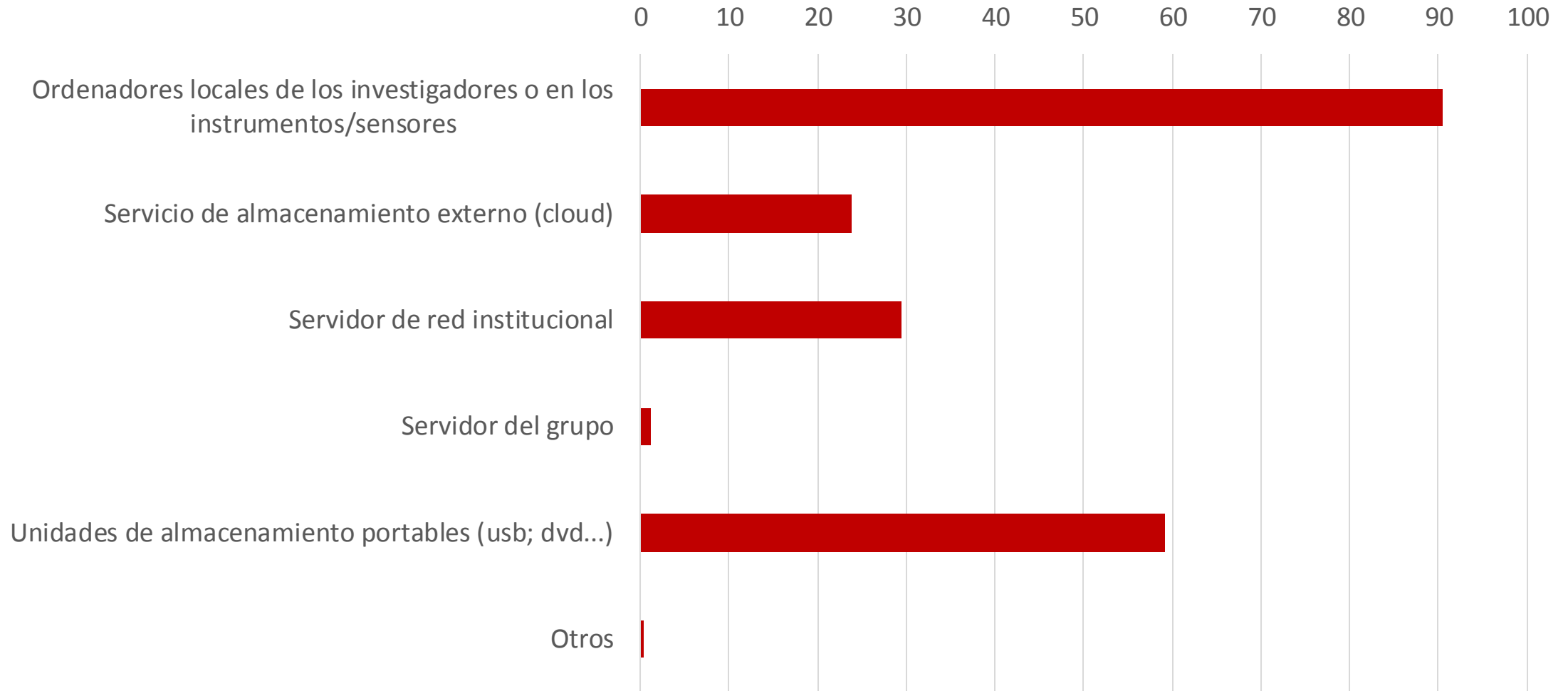
La opinión y los hábitos de los investigadores sobre los datos de investigación.

Cual fue el motivo principal para desarrollar la política de datos

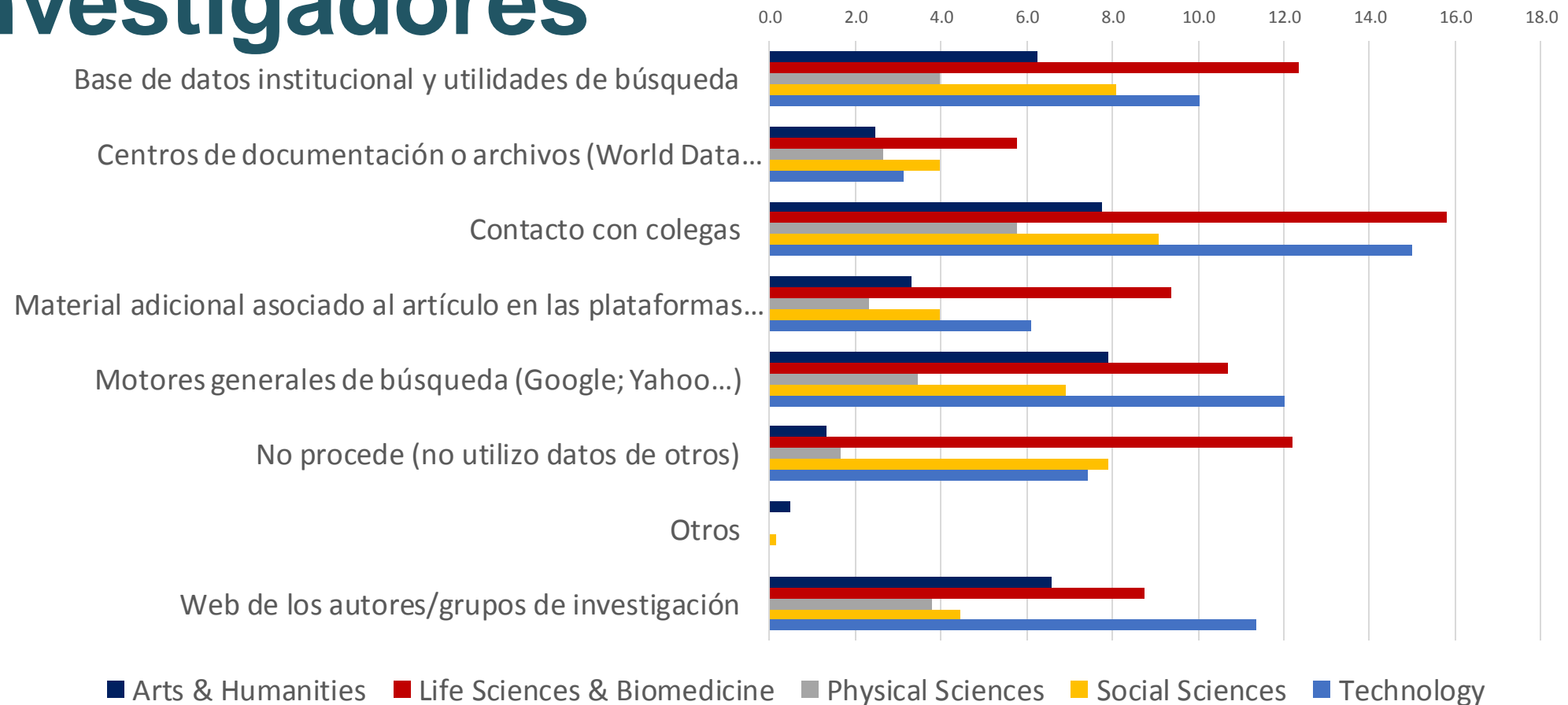


La opinión y los hábitos de los investigadores sobre los datos de investigación.

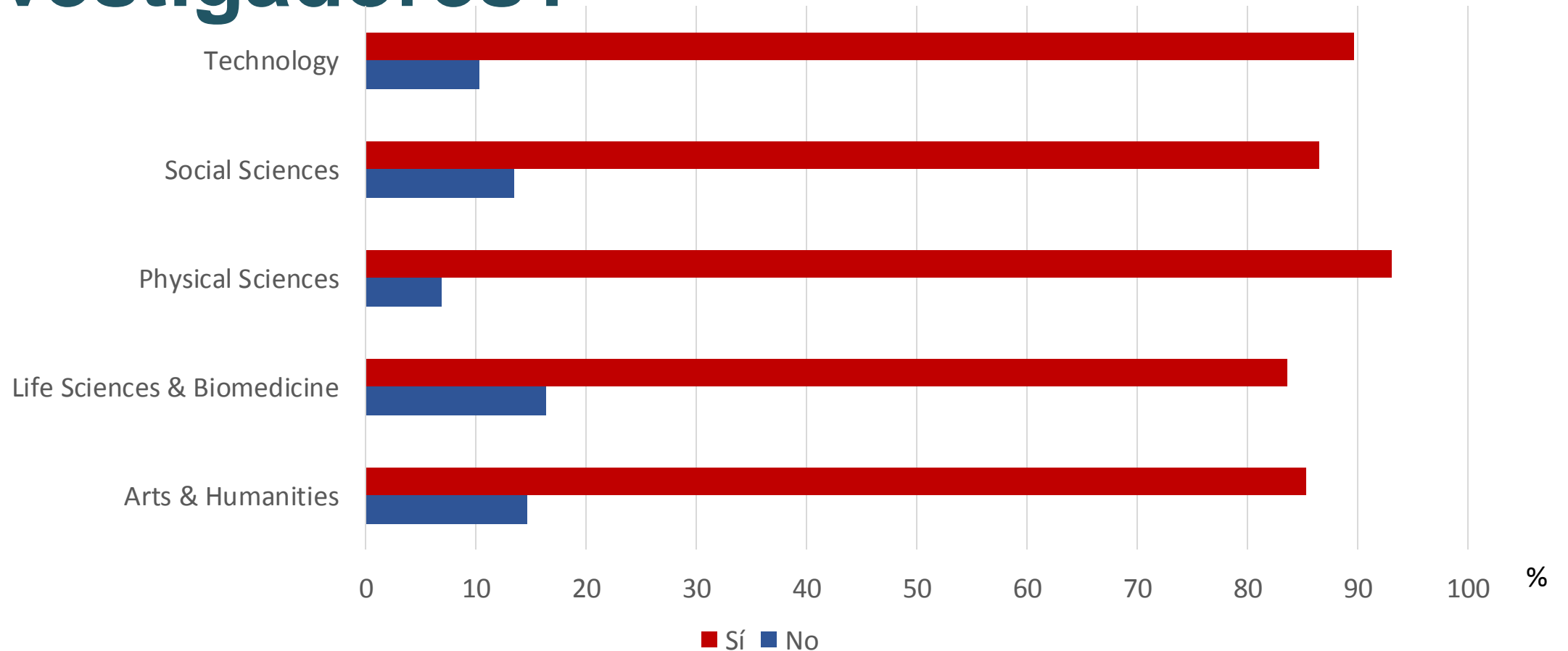
Donde suele almacenar los datos de investigación



Medios utilizados para localizar y acceder a los datos de otros investigadores

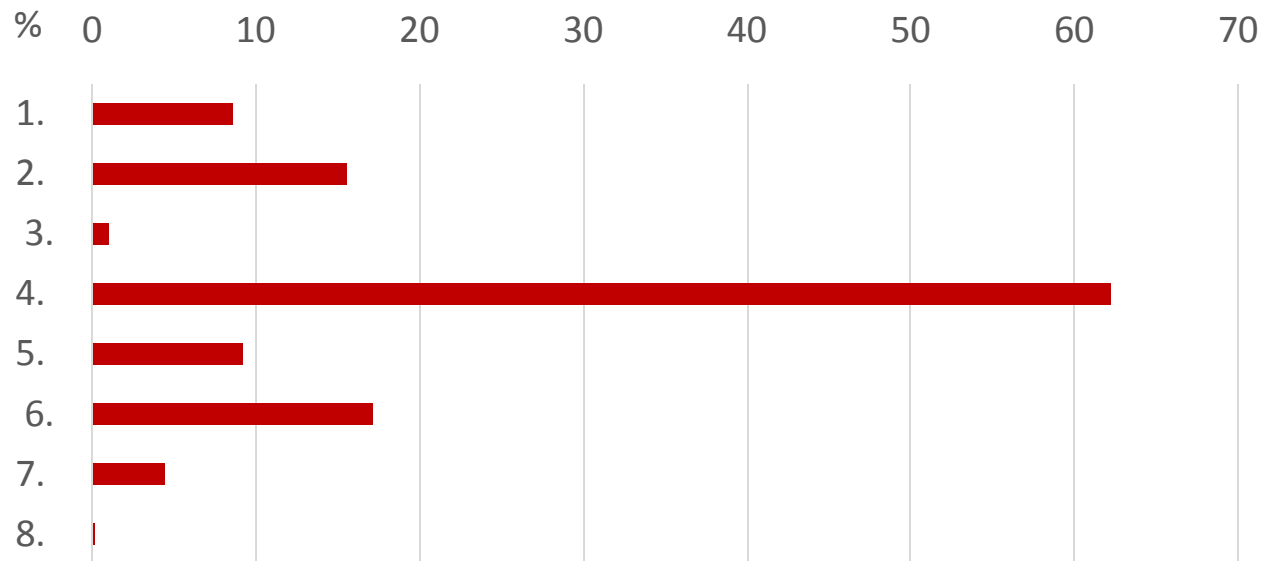


¿Le gustaría utilizar datos de investigación de otros investigadores?



La opinión y los hábitos de los investigadores sobre los datos de investigación.

En relación a la disponibilidad de los datos



1. Mis datos están disponibles en abierto para mi disciplina científica

2. Mis datos están disponibles en abierto para todo el mundo

3. Mis datos están disponibles mediante el pago de una tasa

4. Mis datos están disponibles para mi grupo de investigación y/o colegas colaboradores en la investigación

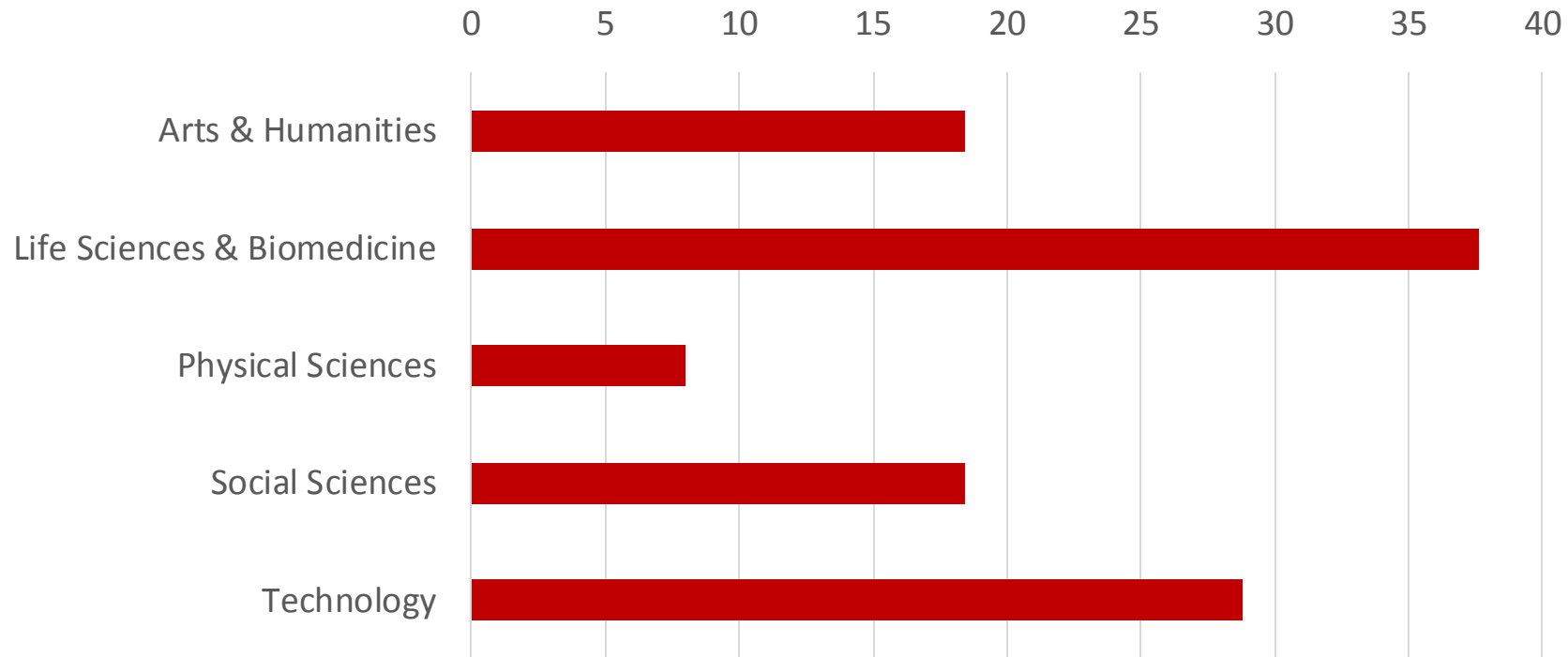
5. Mis datos podrían estar disponibles con los cambios apropiados (por ejemplo: datos clínicos anónimos o con la obligación de citarlos)

6. No comparto mis datos pero me gustaría hacerlo en un futuro

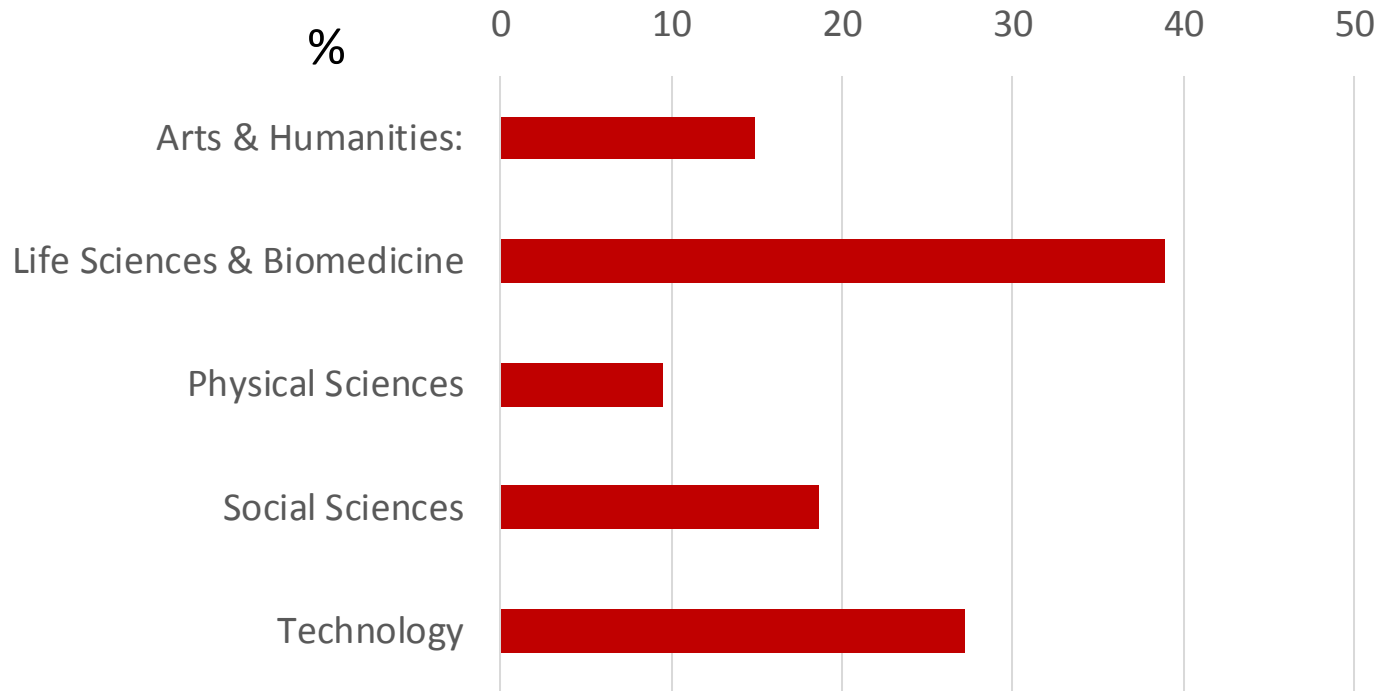
7. No comparto mis datos y no quiero hacerlo en el futuro

8. Otros

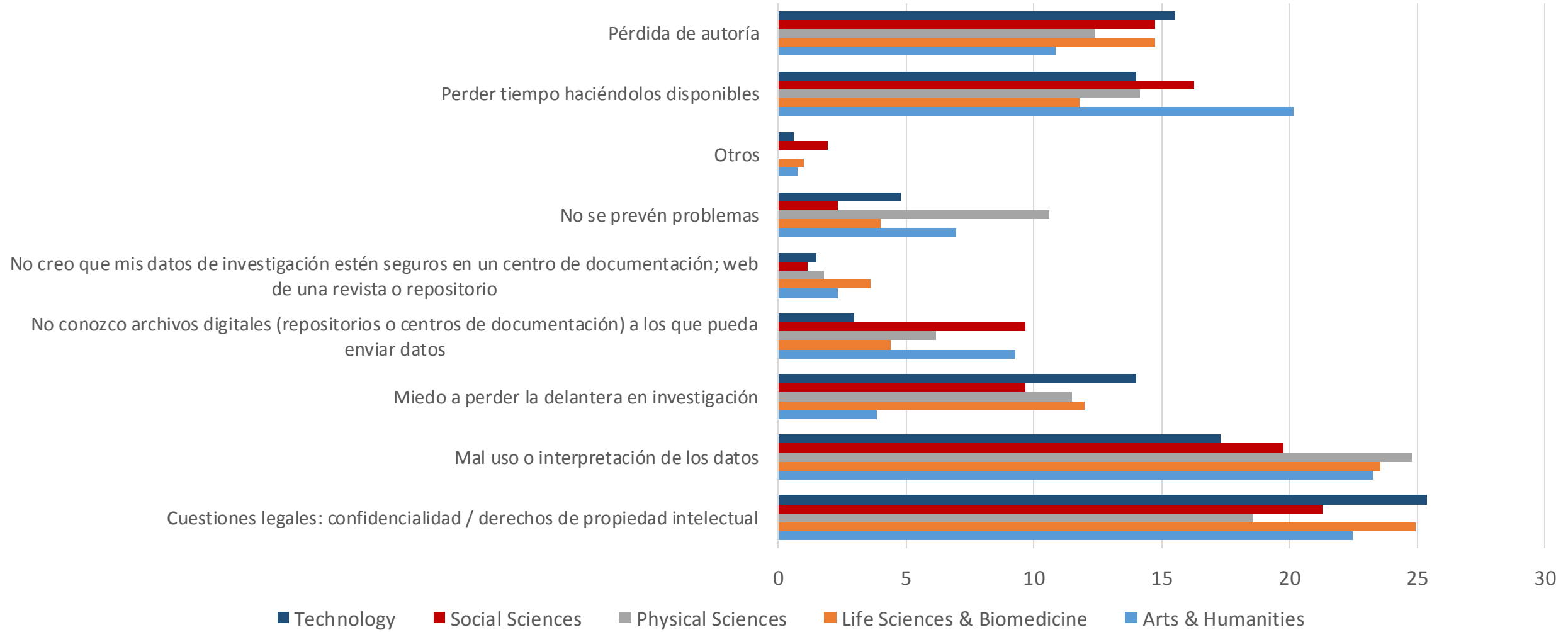
Alguna vez le han solicitado que sus propios datos estén disponibles en abierto



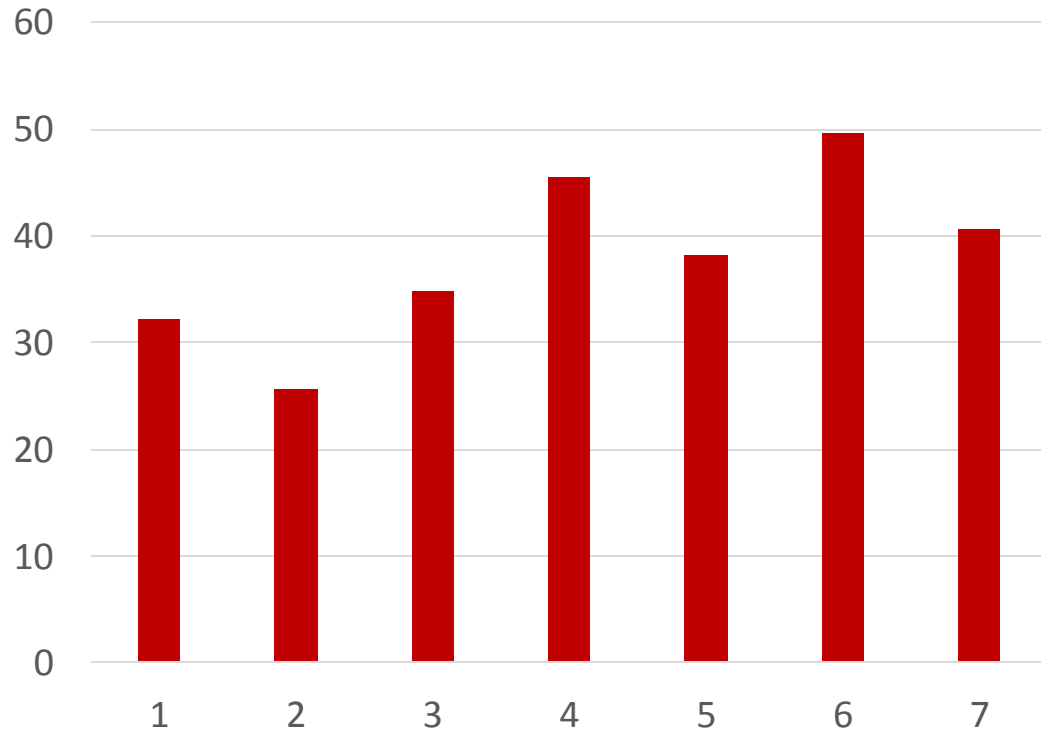
¿Alguna vez le han solicitado sus datos otros investigadores?



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

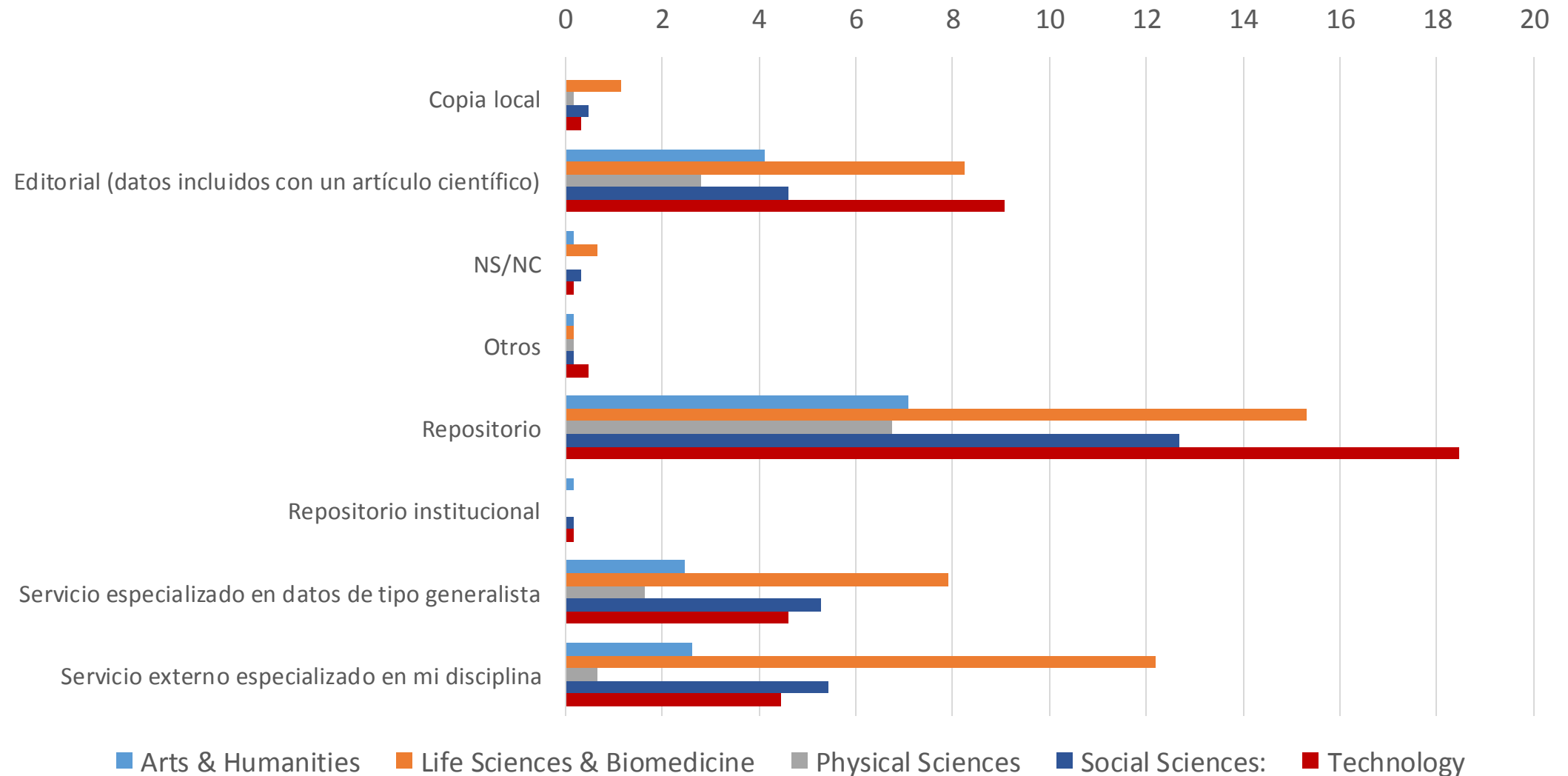


Principales amenazas a compartir

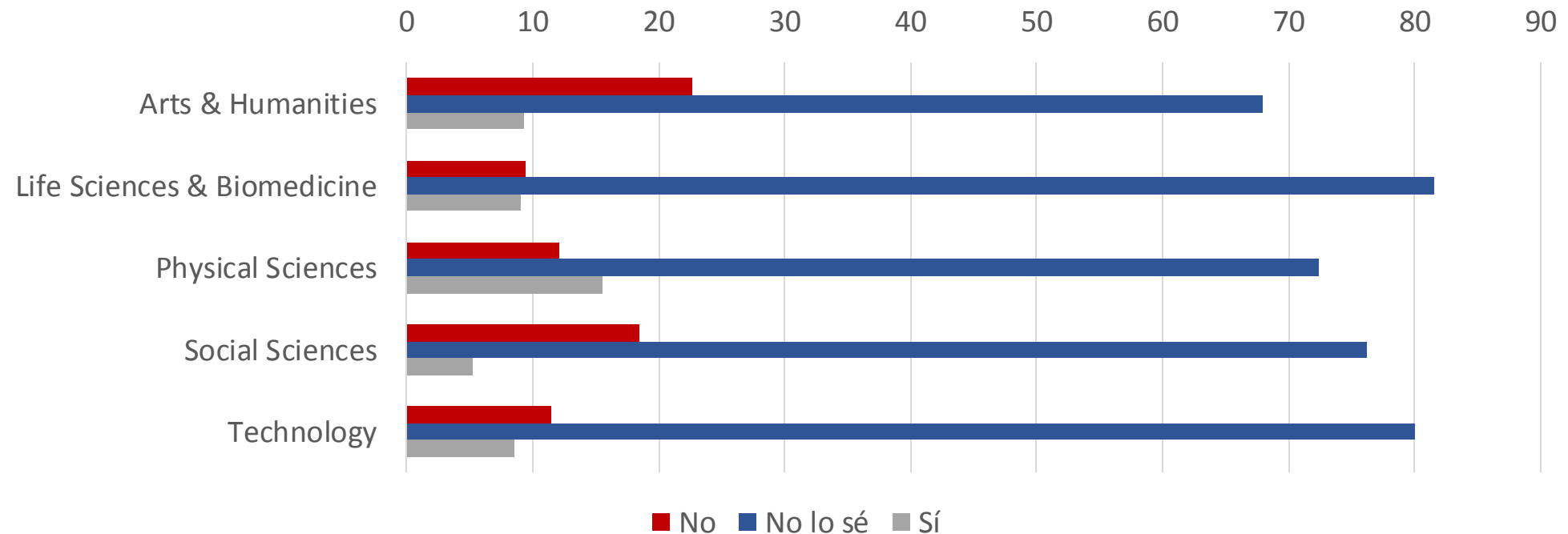


- 1** Los usuarios pueden ser incapaces de comprender o utilizar los datos (semántica, formatos o algoritmos implicados)
- 2** La falta de software, hardware o soporte sostenibles de los ordenadores
- 3** El origen y autenticidad de los datos puede ser incierto
- 4** Las restricciones de acceso y uso (propiedad intelectual)
- 5** Pérdida de capacidad para localizar los datos
- 6** La custodia actual de los datos, en una organización o proyecto, puede dejar de existir en algún punto en el futuro
- 7** Aquellos en quienes confiamos para cuidar los datos, pueden fallarnos

Donde se considera más adecuado la preservación y compartir los datos



Existe en su disciplina alguna instalación de datos digitales que pueda utilizar



Cómo cree que debería ser esta infraestructura

Por tipología de datos

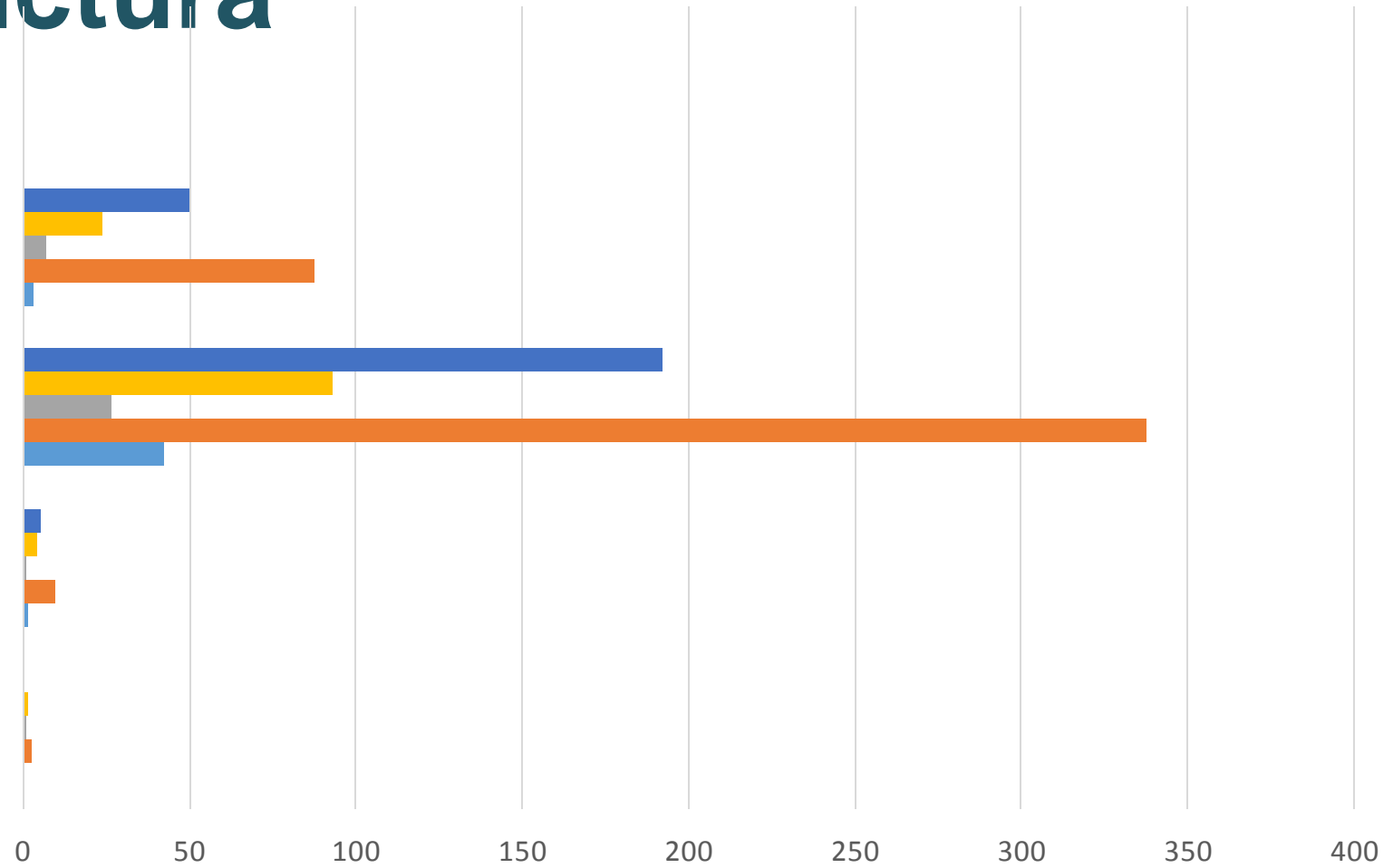
Por organismos

Por disciplina

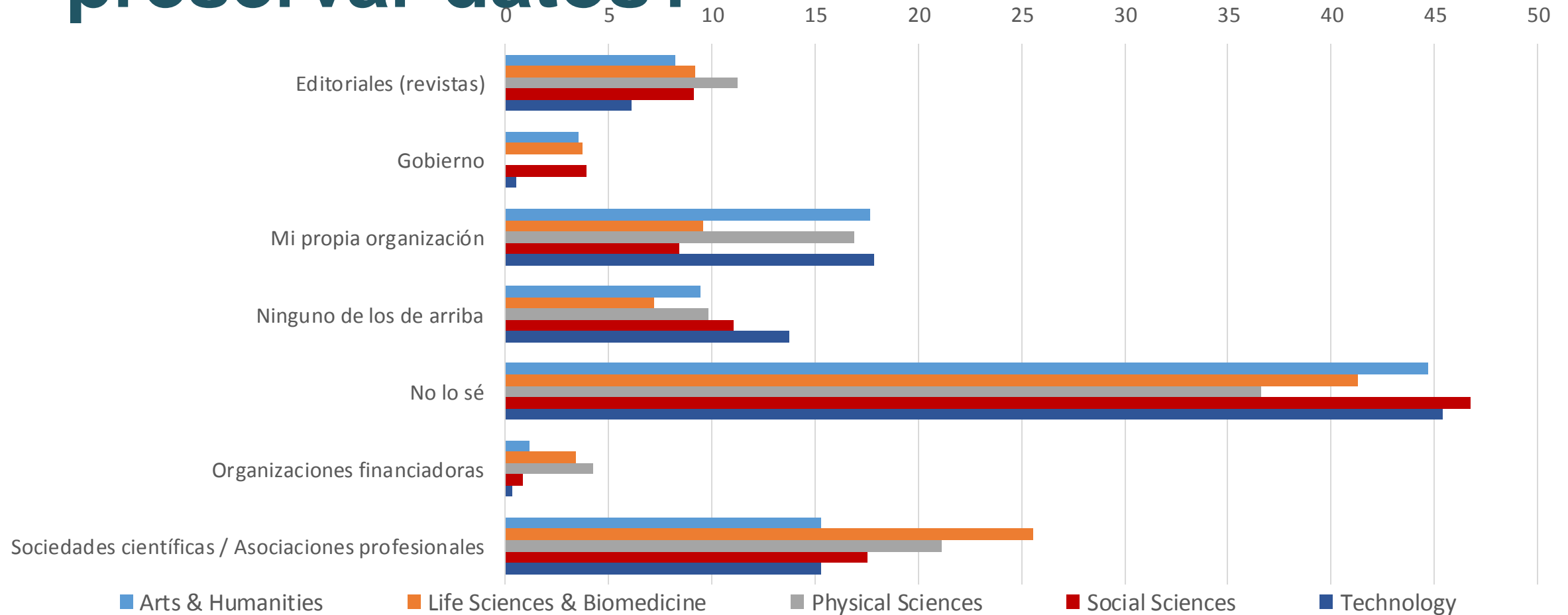
Otros

NS/NC

■ Technology ■ Social Sciences ■ Physical Sciences ■ Life Sciences & Biomedicine: ■ Arts & Humanities



En su disciplina, ¿quién proporciona guías o recomendaciones para preservar datos?

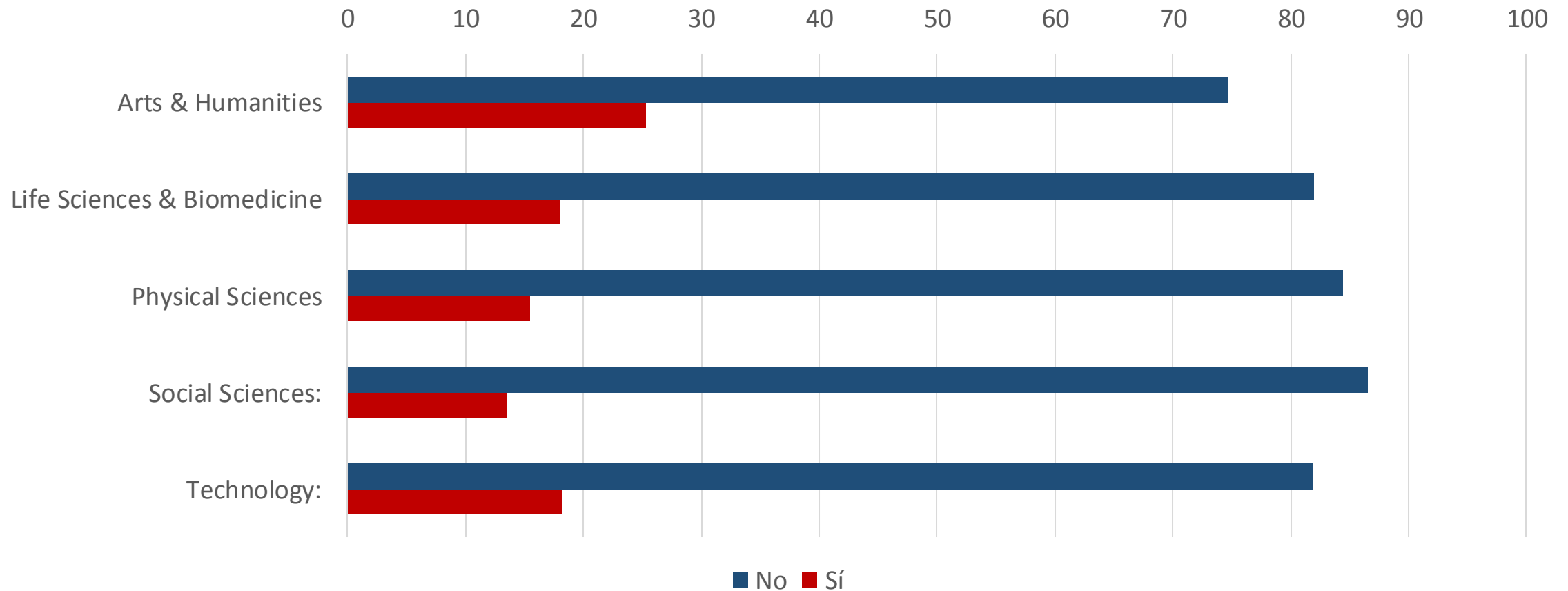


La opinión y los hábitos de los investigadores sobre los datos de investigación.

Razones para la preservación

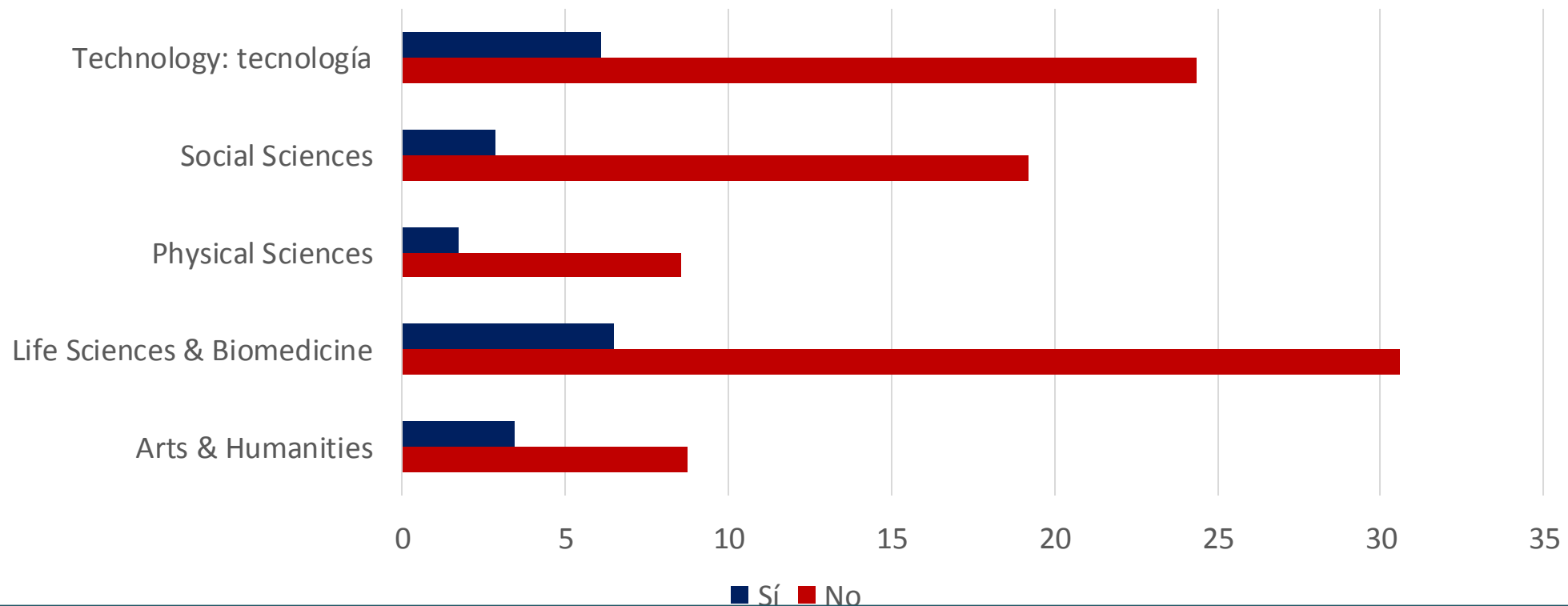
	1	2	3	4	5	6	7
Si la investigación está financiada públicamente, los resultados deberían convertirse en propiedad pública y, por tanto, estar debidamente preservados	2,0	2,0	3,8	8,4	12,2	25,2	41,4
Estimula el avance de la ciencia basada en el conocimiento existente	1,2	0,8	4,3	7,1	15,7	27,7	37,7
Sirve para validar las investigaciones	0,5	0,8	2,8	9,7	16,8	29,2	34,9
Permite el re-análisis de datos existentes	0,5	1,5	2,6	6,3	17,1	28,8	37,1
Puede estimular colaboraciones interdisciplinarias	0,2	1,5	3,5	10,0	19,4	29,0	30,5
Tiene valor económico potencial	8,4	13,3	14,0	18,5	15,0	12,7	12,4
Son únicos e irremplazables	4,8	8,9	11,5	15,3	16,1	16,0	21,1

Su organización le ayuda con la preservación de datos



¿Su organización le ayuda con la preservación los datos?

Modos de ayuda: con procedimientos para nombrar ficheros, guías de servidores externos donde almacenar, personal para subir los ficheros...



Conclusiones

- Suelen utilizar ordenadores locales o memorias digitales portátiles como sistemas de almacenamiento y depósito.
- A una amplia mayoría les gustaría utilizar los datos de investigación recogidos por otros investigadores.
- La principal amenaza a sus datos son es la restricción de acceso y uso relacionada con la propiedad intelectual.
- El principal motivo para desarrollar una política de datos fue por la necesidad de la propia investigación en todas las áreas excepto en Arte y Humanidades, área en la que la complejidad o el volumen de datos asociados con el proyecto fue el motivo principal.

Conclusiones

- Es necesario realizar:
 - Campañas de sensibilización para liberar los datos
 - Crear infraestructuras que faciliten su depósito y preservación
 - Definir las reglas de juego que aclaren muchas de las implicaciones que supone el uso compartido de datos
 - Deben ser los organismos oficiales los impulsores de este movimiento.

Red temática Maredata

Seminario sobre gestión de los datos de investigación

La opinión y los hábitos de los investigadores sobre los datos de investigación.

Antonia Ferrer y Rafael Aleixandre

Con la colaboración de: Fernanda Peset , Antonio Vidal-Infer i Adolfo Alonso-Arroyo

www.datasea.es