

Informe Grupo de Trabajo 2: Plan de Gestión de Datos

Noviembre 2022

Contenido

Introducción.....	3
Herramientas para la realización de Planes de Gestión de Datos	4
Introducción y Metodología.....	5
Herramientas	6
DMP Online	6
DSW: Data Stewardship Wizard	8
EasyDMP	10
EINA.DMP (CORA)	12
PGDOnline.....	14
Argos	16
Conclusiones	19
Recursos necesarios para acometer adecuadamente la Gestión de Datos en los IIS.....	20
Introducción y metodología.....	21
Repositorios y plataformas analizados	22
Global Health Data Exchange (GHDx)	22
OpenNeuro	22
GDC Data Portal GenomicData Commons	23
REPISALUD	24
RECOLECTA.....	26
ELIXIR.....	27
Definición de perfiles profesionales necesarios para la Gestión de Datos	29
Conclusiones	30
Necesidades formativas para la realización de Planes de Gestión de Datos	31
Introducción y Metodología.....	32
Resultados.....	34
Conclusiones	37

Introducción

El presente informe es el resultado del trabajo realizado por 10 Institutos de Investigación Sanitaria [IIB SANT PAU, ISABIAL, INIBIC, ISPA, IDIBAPS, IIS-PRINCESA, IDIPHISA, IDIVAL, IR-HUVH, IIS ARAGÓN (coordinador)] en el marco de la Alianza de los IIS promovida por el Instituto de Salud Carlos III.

Los objetivos fijados para este Grupo de Trabajo han sido los siguientes:

1. Establecer una guía dirigida a los investigadores para la elaboración de los planes de gestión de datos, siguiendo las recomendaciones internacionales de buena práctica y desde el respeto a la normativa vigente.
2. Analizar recursos necesarios para acometer adecuadamente la Gestión de Datos en los IIS:
 - a. Repositorios de datos: situación actual, necesidades y acciones a desarrollar.
 - b. Definición de perfiles profesionales necesarios para la Gestión de datos
3. Identificar las necesidades formativas para el correcto desarrollo de los planes de gestión de datos.

Para dar respuesta a los objetivos fijados, el Grupo de Trabajo ha elaborado este informe con los siguientes apartados:

- Herramientas para la realización de Planes de Gestión de Datos
- Recursos necesarios para acometer adecuadamente la gestión de datos en los IIS
- Necesidades formativas para la realización de Planes de Gestión de Datos

Herramientas para la realización de Planes de Gestión de Datos

Objetivo: Establecer una guía dirigida a los investigadores para facilitar la elaboración de los *Planes de Gestión de Datos*, siguiendo las recomendaciones internacionales de buena práctica y desde el respeto a la normativa vigente.

Introducción y Metodología

Para la generación de planes de gestión de datos (PGD) se apoya en herramientas basadas en plantillas predefinidas que hacen de guía. Hay multitud de plantillas que cubren las necesidades de un PGD, siendo una de las más utilizadas la del Programa Marco. Se han evaluado 6 herramientas que asisten en la generación de Planes de Gestión de Datos. La selección de las herramientas se ha realizado debido a la experiencia que los miembros del grupo de trabajo han tenido sobre ellas, así como la posibilidad de acceder a su uso. Todas las herramientas son utilizadas de forma regular por instituciones de investigación europeas, aunque algunas de ellas tienen un ámbito nacional.

Para que la comparativa sea homogénea se ha acordado realizar una evaluación cualitativa fijando para todas ellas los siguientes criterios:

- Gratuito
- Opensource
- Acepta plantillas adecuadas a los requerimientos de varios financiadores
- Permite cargar nuevas plantillas
- Autocompletar campos
- Formatos a los que permite exportar
- Usabilidad de la herramienta
- Colaborativo: varias personas pueden trabajar a la vez
- Curva de aprendizaje
- Idiomas
- Capacidad de depositar datos
- Capacidad de trabajar con versiones del PGD
- Adaptación al ámbito salud
- Capacidad de incorporar presupuesto de la gestión de datos

Herramientas

Se presenta el análisis, en cada una de las 6 herramientas seleccionadas, de todos los criterios señalados.

DMP Online
Gratuito
Sí
Opensource:
Sí
Acepta plantillas adecuadas a los requerimientos de varios financiadores (p. ej.: Horizonte Europa)
La herramienta permite identificar la entidad financiadora y utilizar la plantilla proporcionada por el financiador (ejemplos European Research Council o European Commission). Si, por el contrario, el financiador no se encuentra en la lista o bien no existe un financiador asociado al PGD, se crea una plantilla basada en la plantilla “DCC Template” proporcionada por el Digital Curation Centre y que responde a su propia guía: Checklist for a Data Management Plan. v.4.0. Edinburgh: Digital Curation Centre. Enlace: http://www.dcc.ac.uk/resources/data-management-plans
Permite cargar nuevas plantillas
No. Las plantillas disponibles se suben por el centro de gestión de la plataforma (DCC) y no están actualizadas a Horizonte Europa.
Autocompletar campos
Permite autocompletar los campos estructurados en distintas secciones. En el caso de la plantilla genérica “DCC template” son: 1) Data Collection; 2) Documentation and Metadata; 3) Ethics and Legal Compliance; 4) Storage and Backup; 5) Selection and Preservation; 6) Data Sharing; 7) Responsibilities and Resources. En el caso de la plantilla “Horizon 2020 DMP” proporcionada por European Commission, además permite elaborar distintas versiones del PGD dependiendo del estado de curso del proyecto: PGD inicial, PGD detallado y PGD revisión final.
Formatos a los que permite exportar
Los formatos en los que permite exportar son: csv, html, pdf, text, docx, json . Además, permite el formato de letra en el pdf y 3 opciones de descarga adicionales:
<ul style="list-style-type: none"> • Preguntas sin contestar
<ul style="list-style-type: none"> • El texto de las preguntas y las secciones
<ul style="list-style-type: none"> • Ficha de detalles del proyecto
Usabilidad de la herramienta
Fácil lectura y descarga, funciones y menús sencillos, uso cómodo, contiene guía en cada uno de los apartados que conforman el PDG.
Colaborativo: varias personas pueden trabajar a la vez

Permite agregar a las personas que van a contribuir a la elaboración del PGD. Por cada colaborador se solicita: nombre, correo electrónico, ORCID, afiliación y rol en el PGD. Entre los roles se incluyen: Administrador de datos, Investigador Principal, Administrador de proyectos y Otro.
Curva de aprendizaje
Aprendizaje rápido e intuitivo. Puedes revisar planes públicos de tu mismo ámbito de actuación, convocatoria o compañeros de la institución/universidad.
Idiomas
Español, inglés (especifica GB o US), francés, alemán.
Capacidad de depositar datos
No se trata de un repositorio de datos.
Capacidad de trabajar con versiones del PGD
DMPonline permite generar copias de un mismo PGD. Para cada una de las versiones o copias se puede invitar a colaboradores y se especifica si se invita a leer o a editar el plan. Las personas invitadas reciben una notificación por correo electrónico de que tienen acceso al PGD. También permite configurar la visibilidad del plan, siendo necesario responder al menos el 50% de las preguntas para habilitar la visibilidad “Pública” y “Organizacional”.
Adaptación al ámbito salud
Falta de especificidad en cuestiones clave como son cuestiones éticas y jurídicas. Sin embargo, fácilmente adaptable a través de otras plantillas que tengan esos apartados (p. ej.: la plantilla por defecto DCC).
Capacidad de incorporar presupuesto de la gestión de datos
En general no, pero existen plantillas que sí la contemplan (Data Management costs como en la plantilla Data Management Plan (September 2020) de Netherlands Organisation for Scientific Research (NWO)).

DSW: Data Stewardship Wizard	
Sitio web	
	https://ds-wizard.org/
Descripción	Herramienta open source desarrollada por ELIXIR para facilitar la creación de un plan de gestión de datos basado en lo que (actualmente) se consideran «buenas prácticas»
Coste	La herramienta es gratuita, aunque existe una opción comercial («FAIRWizard» para «Grupos e Instituciones» que, probablemente es lo que una institución con el ISCIII necesitaría)
Código abierto	Es de código abierto y configurable.
Plantillas	<ul style="list-style-type: none"> • Permite utilizar plantillas de distintos financiadores (https://ds-wizard.org/document-templates/) • Permite el desarrollo de nuevas plantillas
Autocompletar campos	Este es uno de los puntos fuertes de la herramienta. No sólo permite auto completar campos, sino que tiene varios sistemas de preguntas para ayuda.
Formatos a los que se permite exportar	No dispone, o no he sabido localizar de una opción de exportación del DMP. Sí que permite exportar su «Knowledge Model» en formato JSON y los administradores pueden realizar una «migración» a otro modelo de conocimiento. Si por «exportación» se refiere a como generar el documento del DMP esto puede hacerse en Word, Latex o PDF.
Usabilidad de la herramienta	La herramienta dispone de múltiples ayudas en todo momento del proceso, para facilitar su uso. Parece potente en este aspecto.
Colaborativo	Es posible compartir un proyecto, una vez creado, para que las personas con acceso puedan trabajar simultáneamente sobre él.
Idiomas	La herramienta está en inglés y no se ve a simple vista si es posible usar otros idiomas. Diría que no.
Curva de aprendizaje	El hecho, indicado anteriormente de que dispone de múltiples ayudas en todo momento del proceso, simplifica su aprendizaje. Hay mucho material de cursos y talleres en su web.
Capacidad de depositar datos	No parece disponer de esta capacidad.

Capacidad de trabajar con versiones del PGD
Incorpora control de versiones y se puede acceder a un histórico de versiones y cambios realizados.
Adaptación al ámbito salud
Es una herramienta de carácter general sin plantillas o indicaciones específicas para proyectos de salud.
Capacidad de incorporar presupuesto de la gestión de datos
Dispone de un «Storage Costs Evaluator» que genera un presupuesto orientativo de los costes de almacenaje a partir de la cantidad de información y el tiempo que se desea almacenar la misma.
Referencias
https://datascience.codata.org/articles/10.5334/dsj-2019-059/

EasyDMP
Gratuito
Gratuito para todos investigadores que accedan desde Europa y Noruega. Se puede acceder libremente desde https://easydmp.sigma2.no
Opensource
Software libre con licencia MIT que permite el uso del software, así como su modificación y redistribución. El código fuente está accesible para su descarga.
Acepta plantillas adecuadas a los requerimientos de varios financiadores
Al crear un nuevo PGD se selecciona la plantilla con la que se va a acceder. Actualmente tiene definidas: Science Direct y EU Horizon 2020 template.
Permite cargar nuevas plantillas
En la guía de ayuda se da la posibilidad de definir nuevas plantillas contactando con los desarrolladores. Sin embargo, el que sea software libre, permite realizar una instalación propia que incluya plantillas adaptadas a las necesidades de la institución.
Autocompletar campos
No disponible.
Formatos a los que permite exportar
Formato Acrobat Reader pdf y html.
Usabilidad de la herramienta
Muy sencilla. Las preguntas definidas en la plantilla aparecen en orden secuencial, siendo sencilla la cumplimentación.
Colaborativo: varias personas pueden trabajar a la vez
Permite definir coeditores que reciben un enlace a través del cual pueden editar el PGD.
Curva de aprendizaje
Corta. Es muy sencilla de utilizar, ya que no tiene muchas opciones y está enfocada a la cumplimentación secuencial de la información predefinida.
Idiomas
Únicamente en inglés.
Capacidad de depositar datos
No disponible directamente desde la herramienta.
Capacidad de trabajar con versiones del PGD
No disponible.
Adaptación al ámbito salud
No hay adaptación específica.
Capacidad de incorporar presupuesto de la gestión de datos
No disponible.

Horizon 2020: TEST DMP

[← Go to plan summary](#)

Your answers are saved every time you press "prev" or "next"

1 Data Summary

2 FAIR data

3 Allocation o...

4 Data Security

5 Ethical aspe...

6 Other

1 Data Summary

Covers a brief summary of the data you intend to use and how it will be used in your project.

Q1.5 Please specify where the data originate from.*

Derived covers data from simulations, models, etc. Experimental covers data arising from experiments made on objects (such as particle physics, chemical reactions etc). Observational covers data from observation of objects (such as earthquakes, sunspot activity, etc).

- Observational**
- Experimental**
- Derived**
- Other**

More information

If you need to go more in depth, do so here. This will be shown in the generated text

Datasets 1 & 2 origin from observational experiments on animals perform during the project.

Dataset 3 origins from data provided from the health authority. These personal data are anonymised and comply with current regulation.

Prev

Question 6/8

Next

EINA.DMP (CORA)
Gratuito
Sí
Opensource
Sí (DMPRoadmap).
Acepta plantillas adecuadas a los requerimientos de varios financiadores
Sí, hay plantillas para Horizon Europe, H2020, ERC, Plan Estatal y otras que no responden a una agencia de financiación concreta. También hay disponible una plantilla de Software Management Plan, así como un modelo simplificado dirigido a doctorandos.
Permite cargar nuevas plantillas
Sí, las plantillas se pueden ir generando. Si una institución/s tiene la suya propia también se pueden cargar dichas plantillas.
Autocompletar campos
Sí, aunque en número limitado.
Formatos a los que permite exportar
Pdf, docx, html, csv.
Usabilidad de la herramienta
Fácil.
Colaborativo: varias personas pueden trabajar a la vez
Sí, cualquier persona puede registrarse y utilizar la herramienta. Si una institución tiene un espacio creado dentro de Eina.dmp existe una limitación de visualización en algunos campos para usuarios de fuera de esa institución. Adicionalmente, se pueden dar permisos de co-owner (todos los permisos), colaborador (añadir, comentar, pero no compartir) o lectura (solo visualización).
Curva de aprendizaje
Aprendizaje rápido e intuitivo.
Idioma
Inglés.
Capacidad de depositar datos
Sí, dispone de repositorio propio CORA al cual hay que registrarse por separado. Se trata de otra herramienta independiente y que en los siguientes meses presentará funcionalidades adicionales que permitirá la “comunicación” entre esta y Eina.dmp.
Capacidad de trabajar con versiones del PGD
Sí, se pueden crear y recuperar distintas versiones del historial.
Adaptación al ámbito salud
No, por ejemplo, no hay funcionalidades/campos ampliados en temas relevantes como seguridad o ética en el uso de datos sensibles.
Capacidad de incorporar presupuesto de la gestión de datos

Sí, aunque la herramienta no permite subir ficheros (por ejemplo, Excel), hay un apartado específico donde incluir información sobre el presupuesto en el que se puede incluir, de manera narrativa, toda la información económica del PGD.

PGDOnline
Descripción
Herramienta web desarrollada por el Consorcio Madroño para elaborar planes de gestión de datos adaptada partir de DMPonline, creada por DCC (Digital Curation Centre de UK)
Sitio web
https://pgd.consorciomadrono.es/
Gratuito
Sí, pero actualmente sólo contempla como organismos las Universidades de la Comunidad de Madrid.
Opensource
Sí
Acepta plantillas adecuadas a los requerimientos de varios financiadores
No, actualmente sólo funciona con la plantilla para Horizonte 2020
Permite cargar nuevas plantillas
No
Autocompletar campos
No. La cumplimentación hay que realizarla de forma manual. La herramienta tan sólo te guía en qué campos hay que cumplimentar y te ofrece una descripción de qué es lo que se pide en cada apartado.
Formatos a los que permite exportar
csv, html, pdf, texto, doc.
Usabilidad de la herramienta
La herramienta es bastante intuitiva. No requiere de aprendizaje previo
Colaborativo: varias personas pueden trabajar a la vez
Sí. Se pueden asignar colaboradores y diferentes niveles de visibilidad del plan.
Curva de aprendizaje
Sencillo.
Idiomas
Alemán, Inglés, Español, Francés, Portugués. La traducción es a nivel de plantilla.
Capacidad de depositar datos
No.
Capacidad de trabajar con versiones del PGD
No. Tan sólo permite definir un Plan Inicial, un Plan detallado y un Plan Final. En cada fase se incrementa el número de preguntas a contestar.
Adaptación al ámbito salud
No, es una herramienta de generación de PGD genérica.
Capacidad de incorporar presupuesto de la gestión de datos

Plan de Gestión de Datos

No.



PGDOnline

Mi panel

Crear planes

Referencia

Ayuda

Idioma

Detalles del proyecto

Resumen del plan

PGD inicial

PGD detallado

PGD final

Compartir

Solicitar información

Descargar

expandir todo | contraer todo

0/35

1. Resumen de los datos (0 / 7)



2.1 Crear datos localizables suministrando metadatos [datos FAIR] (0 / 10)



2.2 Crear datos accesibles en abierto [datos FAIR] (0 / 5)



2.3 Crear datos interoperables [datos FAIR] (0 / 2)



2.4 Incremente las posibilidades de reutilizar los datos (usando licencias claras) [datos FAIR] (0 / 5)



3. Necesidades de recursos (0 / 3)



4. Seguridad de los datos (0 / 1)



5. Cuestiones éticas (0 / 1)



6. Otros (0 / 1)



Argos
Descripción
Es una herramienta online para la creación, gestión, difusión y enlace de un PGD, que hace énfasis en la aplicación de los principios FAIR y en las mejores prácticas para fomentar la accesibilidad a los datos de investigación. Tiene un funcionamiento abierto y colaborativo y ha sido desarrollada de forma conjunta entre OpenAIRE
Sitio web
(https://www.openaire.eu/) y EUDAT (https://www.eudat.eu/).
Gratuito
Sí, además Argos no requiere un registro específico en la herramienta, ya que permite iniciar sesión seleccionando uno de los proveedores de la página de inicio de sesión (OpenAIRE, ORCID, EUDAT, GMAIL o a través de perfiles de redes sociales como Facebook o Twitter).
Opensource
Sí, https://argos.openaire.eu/opensource-licences
Acepta plantillas adecuadas a los requerimientos de varios financiadores
Sí. Ofrece la mayoría de las plantillas necesarias para acceder a los fondos de investigación, así como la necesaria para solicitar los fondos ofrecidos en Horizonte Europa.
Permite cargar nuevas plantillas
Sí, pero con cuenta de administrador. Con cuenta de administrador existe la posibilidad de proponer plantillas nuevas a medida de las necesidades de los autores y agencias de financiación. La cuenta de administrador debe solicitarse a argos@openaire.eu .
Autocompletar campos
Sí. La plantilla permite la interoperabilidad con otras herramientas openaire. Argos permite autocompletar los datos de manera automática, si previamente existe un recurso cargado en alguna de esas herramientas o repositorios. Durante el proceso aparecen cuadros dialogo que facilita la tarea de autocompletado.
Formatos a los que permite exportar
Exportación del PGD en formatos comprensibles por máquinas (xml, json) y por personas. En concreto: PDF, documento, XML, RDA JSON (se puede importar a otras herramientas DMP compatibles con RDA).
Usabilidad de la herramienta

<ul style="list-style-type: none"> • Asistente para la creación del PGD en función del financiador. • Asistente para la descripción de los datos de investigación, de manera que permite reutilizarlos en diferentes PGDs. • Asignación de un DOI para el PGD si se publica, por lo que se facilita la visibilidad y citación del propio PGD. • Exportación del PGD en formatos comprensibles por máquinas (xml, json) y por personas. • Servicio incluido en el European Open Science Cloud (EOSC), iniciativa promovida por la Comisión Europea. • Enlaza con la cuenta que el usuario tiene en ZENODO. De esta manera, relaciona los planes de gestión de datos con los datasets depositados en ZENODO, el repositorio multidisciplinar creado por la Unión Europea para cumplir con el proyecto general de datos abiertos. • No existe limitación de tamaño ni para los planes ni para los conjuntos de datos.
<p>Colaborativo</p>
<p>Sí. Los PGD generados en ARGOS se gestionan a su vez como productos de investigación a los que se pueden asignar DOI, licencias, etc. Además, permite que los PGDs se hagan de forma colaborativa entre varias personas.</p>
<p>Curva de aprendizaje</p>
<p>A lo largo de todo el proceso de autocompletado, aparecen cuadros dialogo que facilita ayuda al investigador para completar el plan de gestión de datos conforme a las indicaciones prácticas que se incluye en cada uno de los campos. Esto permite al investigador obtener una visión general de la gestión de los datos en el marco del proyecto, a la vez que le permite adquirir destrezas de cara a la elaboración de futuros planes.</p>
<p>Idiomas</p>
<p>9 idiomas, entre ellos el español.</p>
<p>Capacidad de depositar datos</p>
<p>Sí. Enlaza con la cuenta que el usuario tiene en ZENODO. De esta manera, relaciona los planes de gestión de datos con los datasets depositados en ZENODO, el repositorio multidisciplinar creado por la Unión Europea para cumplir con el proyecto general de datos abiertos.</p>
<p>Capacidad de trabajar con versiones del PGD</p>
<p>Sí, además de descargar en diferentes formatos, invitar investigadores para trabajar en grupo, permite subir nuevas versiones.</p>
<p>Adaptación al ámbito salud</p>
<p>No. No obstante, el cuestionario o plantilla contiene campos que permite completar y dar respuesta a algunas de las claves específicas derivadas de tratamiento de los datos en materia de salud, como son las relativas a: consentimiento informado, anonimización, tratamiento de los datos, asignación de roles en el tratamiento, así como a cuestiones éticas.</p>
<p>Capacidad de incorporar presupuesto de la gestión de datos.</p>

La plantilla incluye preguntas relativas a la asignación y gestión de costes relativos a la gestión de datos, pero no incluye una herramienta específica que permita calcular los costes o presupuestos asignados al plan de gestión de datos.

The screenshot displays the ARGOS web application interface. At the top left is the ARGOS logo. The top right contains navigation links: 'Iniciar un nuevo PGD', 'FAQ', 'ES', and 'INICIAR SESIÓN'. A left sidebar menu includes 'Inicio', 'PGDs publicado', 'Descripciones del dataset publicadas', 'Marca compartida', 'Soporte', and 'Enviar feedback'. The main content area features a modal window titled '¿Qué es ARGOS?' with the following text: 'ARGOS es un servicio extensible y abierto que simplifica la gestión, validación, monitorización y mantenimiento de los Plan de Gestión de Datos. Permite a los participantes (investigadores, gestores, supervisores, etc) crear un PGDs visible que puede ser compartido libremente entre distintas infraestructuras para llevar a cabo aspectos específicos del proceso de Gestión de Datos de acuerdo con los propósitos y el compromiso de los propietarios de los datos.' Below the text is a cartoon character holding a laptop. To the right of the modal, 'Public Usage' statistics are shown: 19 PGDs públicos, 26 Descripciones de Dataset Públicos, 18 Subvenciones, and 170 Organizaciones relacionadas. At the bottom of the main area, it says 'Última actividad'.

Conclusiones

Todas las herramientas están basadas en software libre para uso individual, salvo DSW que no lo es para instituciones. El uso de una herramienta de software libre permite adaptaciones y personalizaciones de la herramienta con el fin de que se ajuste a las necesidades de las instituciones: desarrollo de nuevas funcionalidades, nuevas plantillas, nuevos idiomas, etc.

Ninguna herramienta incluye apartados específicos de proyectos de ciencias de la salud, aunque alguna (Argos) han incluido en sus plantillas apartados específicos en relación sobre todo en la gestión de datos personales de conformidad con la legislación vigente. Sería recomendable el desarrollo de plantillas que incluyan textos que apoyen en su cumplimiento, explorando la posibilidad de que esta actividad se gestionase centralizadamente desde el ISCIII.

Únicamente una herramienta contempla la gestión económica que conlleva el desarrollo de un plan de gestión de datos: EINA.CORA

La mayoría de las herramientas asisten en la generación de una primera versión del PGD, aunque hay algunas que también incluyen la funcionalidad para depositar los datos generados en repositorios públicos: EINA.CORA y ARGOS.

Recursos necesarios para acometer adecuadamente la Gestión de Datos en los IIS

Objetivo: Analizar recursos necesarios para acometer adecuadamente la Gestión de Datos en los IIS:

- a. Repositorios de datos: situación actual, necesidades y acciones a desarrollar.
- b. Definición de perfiles profesionales necesarios para la Gestión de datos

Introducción y metodología

El trabajo realizado se ha centrado en la selección de los repositorios más afines a las áreas de investigación en salud. Tras el mapeo de estos por parte del grupo de trabajo se ha realizado un análisis individualizado de 3 repositorios internacionales. Durante la búsqueda de repositorios, se constató que la mayoría estaban incluidos en plataformas de datos, las cuales también hemos procedido a analizar. Las plataformas analizadas han sido RECOLECTA y ELIXIR. Ambas son institucionales, siendo la primera promovida por la FECYT y la segunda por la Red Europea ELIXIR. Tras la finalización del análisis se han redactado unas conclusiones en relación con el uso de los repositorios y la utilidad de las plataformas.

También se ha hecho un análisis sobre los diferentes perfiles profesionales que pueden intervenir en el proceso de gestión de datos. Se hizo una puesta en común entre los diferentes integrantes del grupo de trabajo aportando la experiencia de cada institución, y definiendo tanto la titulación como la formación complementaria que deben tener.

Repositorios y plataformas analizados

Global Health Data Exchange (GHDx)
Descripción
El Global Health Data Exchange (GHDx) es un catálogo de datos creado y respaldado por el Institute for Health Metrics and Evaluation (IHME). El GHDx apoya directamente la misión del IHME de mejorar la salud de las poblaciones del mundo proporcionando la mejor información sobre la salud de la población. Fue creado como un lugar dedicado para cualquier persona interesada en la salud y la demografía global para encontrar y compartir rápidamente información sobre datos junto con conjuntos de datos reales.
Principales Características
Agrupación de catálogos de encuestas, censos, estadísticas vitales y otros datos relacionados con la salud más completo del mundo. Esta base de datos reúne documentos de salud de más de 200 países pudiéndose realizar búsquedas por tipo de documentos, países, tipo de organización, palabras claves, etc.
Desventaja
Es creado y financiado por un Instituto que pertenece a la Universidad de Whashington.

OpenNeuro
Descripción
Es una plataforma gratuita especializada y abierta para validar y compartir datos de MRI, PET, MEG, EEG e iEEG compatibles con BIDS con más de 29 mil participantes y más de 700 documentos públicos.
Principales Características
El nuevo material es agregado a medida que los investigadores abren sus propios datos. Los conjuntos de datos están disponibles públicamente para fomentar la investigación y conseguir mejores diagnósticos en formato de datos PET, MEG, EEG e iEEG compatible con Brain Imaging Data Structure (BIDS).
Desventajas
Es un repositorio específico de Neurociencia

GDC Data Portal GenomicData Commons
Descripción
<p>El Genomic Data Commons (GDC) es un programa de investigación del Instituto Nacional del Cáncer (NCI). La misión de la GDC es proporcionar a la comunidad de investigación del cáncer un repositorio unificado y una base de conocimientos sobre el cáncer que permita compartir datos en los estudios genómicos del cáncer en apoyo de la medicina de precisión.</p> <p>El GDC del Instituto Nacional del Cáncer (NCI) es una plataforma de intercambio de datos que promueve la medicina de precisión en oncología. No es solo una base de datos o una herramienta; es una red de conocimiento expandible que apoya la importación y estandarización de datos genómicos y clínicos de los programas de investigación del cáncer.</p> <p>El GDC contiene datos generados por el NCI a partir de algunos de los conjuntos de datos genómicos del cáncer más grandes y completos, incluido el Atlas del Genoma del Cáncer (TCGA) y la Investigación Terapéuticamente Aplicable para Generar Terapias Efectivas (TARGET). Por primera vez, estos conjuntos de datos se han procesado utilizando un conjunto común de canalizaciones bioinformáticas, de modo que los datos se pueden comparar directamente como un sistema de conocimiento creciente para el cáncer. El GDC también permite a los investigadores enviar datos, y el GDC procesa estos datos utilizando canalizaciones bioinformáticas para alinear los datos con un genoma de referencia común y generar datos de nivel superior, como llamadas de variantes y cuantificaciones de expresión. A medida que más investigadores agreguen datos clínicos y genómicos a la GDC, se convertirá en una herramienta aún más poderosa para hacer descubrimientos sobre las bases moleculares del cáncer que pueden conducir a una mejor atención para los pacientes.</p>
Principales Características
<p>Repositorio de datos unificado de los datos genómicos del Instituto Nacional del Cáncer (NCI). Permite compartir datos entre estudios genómicos del cáncer en apoyo de medicina de precisión.</p> <p>Apoya varios programas del genoma del cáncer en el Centro NCI para Cancer Genomics (CCG), como son:</p> <ul style="list-style-type: none">• The Cancer Genome Atlas (TCGA),• Terapéuticamente Investigación Aplicable para Generar Tratamientos Efectivos (TARGET),• Genoma del Cáncer Iniciativa de Caracterización (CGCI). <p>Proporciona una plataforma para consultar y descargar datos de alta calidad y completos. Proporciona un GDC Herramienta de transferencia de datos y una API de GDC para acceso de programación.</p>

REPISALUD
Descripción
<p>REPISALUD es el Repositorio Institucional de SALUD del Instituto de Salud Carlos III (ISCIII) y sus Fundaciones, el Centro Nacional de Investigaciones Cardiovasculares (CNIC) y el Centro Nacional de Investigaciones Oncológicas (CNIO), en el que se recoge en un único archivo digital y de acceso libre, la producción científica y académica de estos centros de referencia en sus áreas.</p> <p>OBJETIVOS PRINCIPALES DE REPISALUD</p> <ul style="list-style-type: none">• Aumentar la visibilidad, el impacto y la transferencia del conocimiento generado en estos centros.• Asegurar el archivo y la preservación de los objetos depositados a texto completo así como su identificación mediante un identificador único persistente.• Facilitar a los investigadores de estos centros el cumplimiento de los mandatos de los organismos financiadores, tanto nacionales como internacionales, permitiéndoles depositar en Acceso abierto todos los documentos sujetos a dicha obligación.
Principales Características

El contenido de REPISALUD se estructura en torno a cinco comunidades, formadas a su vez por subcomunidades que agrupan los centros descritos y sus departamentos según proceda. Estas subcomunidades se dividen en las colecciones que se muestran a continuación, las cuales contienen los documentos susceptibles de formar parte del repositorio. Estos documentos pueden estar incluidos en más de una colección, a fin de facilitar su localización en base a distintos criterios de búsqueda.

INVESTIGACIÓN:

- Artículos
- Contribuciones a congresos
- Datos de investigación
- Libros y capítulos de libro
- Patentes
- Tesis y trabajos de investigación

INSTITUCIONAL:

- Informes y documentos de trabajo
- Material de divulgación
- Memorias

DOCENCIA:

- Material docente

EVENTOS CIENTÍFICOS:

- Conferencias
- Jornadas y congresos
- Seminarios

PROGRAMA EDITORIAL

- Monografías y Revistas

RECOLECTA
Descripción
<p>La Fundación Española para la Ciencia y Tecnología, F.S.P. (FECYT) y la Red de Bibliotecas Universitarias (REBIUN) de la CRUE han desarrollado a nivel nacional una Plataforma (agregador nacional de repositorios de acceso abierto) que agrupan a todas las infraestructuras digitales españolas en las que se publican y/o depositan resultados de investigación en acceso abierto (RECOLECTA, o Recolector de Ciencia Abierta).</p> <p>Entre los objetivos que persigue esta plataforma se encuentran:</p> <ul style="list-style-type: none">• Impulsar y coordinar la infraestructura nacional de repositorios científicos digitales de acceso abierto y garantizar que sean interoperables según los estándares de la comunidad mundial.• Promover, apoyar y facilitar la adopción del acceso abierto por todos los investigadores de las universidades y centros de I+D españoles, principales productores de conocimiento científico en nuestro país.• Dotar de una mayor visibilidad tanto nacional como internacional de los resultados de la investigación que se realiza en España. <p>En España existe una importante infraestructura de recursos en los que publicar y/o depositar en abierto los trabajos de investigación, formada por repositorios institucionales, repositorios temáticos y revistas en acceso abierto.</p> <p>RECOLECTA agrega todos esos recursos en una única plataforma que garantiza su interoperabilidad y permite el acceso a toda la producción científica nacional en abierto.</p> <p>Servicios de RECOLECTA a los repositorios de acceso abierto:</p> <ul style="list-style-type: none">• Validador: Permite a los repositorios autoevaluar, tantas veces como quieran, su grado de cumplimiento con las directrices internacionales de interoperabilidad e identificar los registros erróneos.• Recolector: Agrega los metadatos de los repositorios que forman parte de la comunidad de RECOLECTA con periodicidad semanal.• Buscador: Permite localizar y acceder de forma libre y gratuita a toda la producción científica depositada en abierto en los repositorios españoles a través de una interfaz única.• Atención a usuarios a través de la cuenta recolecta@fecyt.es
Principales Características
<p>En esta plataforma podemos encontrar 169 Repositorios de los cuales 108 son Repositorios Institucionales con más de 2.5 millones de documentos en acceso abierto y más de 28 mil set de datos de investigación.</p>

ELIXIR
Descripción
<p>ELIXIR es una organización intergubernamental que reúne recursos de ciencias de la vida de toda Europa. Estos recursos incluyen bases de datos, herramientas de software, materiales de capacitación, almacenamiento en la nube y supercomputadoras.</p> <p>El objetivo de ELIXIR es coordinar estos recursos para que formen una única infraestructura. Esta infraestructura facilita que los científicos encuentren y compartan datos, intercambien conocimientos y acuerden las mejores prácticas. En última instancia, les ayudará a obtener nuevos conocimientos sobre cómo funcionan los organismos vivos.</p> <p>TeSS es un ejemplo de un recurso ELIXIR. TeSS es un portal de formación en línea que reúne materiales y cursos de formación en ciencias de la vida de toda Europa y le permite buscarlos en un sitio web. Esto facilita que los científicos encuentren la formación que necesitan y da a los cursos de formación una mayor publicidad.</p> <p>ELIXIR incluye 22 miembros y un observador, que reúne a más de 220 organizaciones de investigación. Fue fundado en diciembre de 2013 y comenzó a implementar su primer programa científico en 2014. Actualmente está implementando su segundo programa científico de cinco años .</p> <p>ELIXIR ha compilado una lista de recursos que recomienda para el depósito de datos experimentales. La comunidad científica tiene la responsabilidad compartida de garantizar la conservación y accesibilidad de los datos a largo plazo. El propósito de esta lista de bases de datos de depósito es brindar orientación a quienes formulan políticas y prácticas de trabajo sobre los repositorios apropiados para publicar datos abiertos en las ciencias de la vida.</p>
Principales Características
<p>En esta plataforma que cuenta con un amplio listado de bases de datos en biomedicina así como de “core data resources” que puede utilizarse como una plataforma de referencia a nivel europeo para los IIS.</p>

Además de estos repositorios, puede ser interesante trabajar con ZENODO, repositorio de carácter multidisciplinar.

ZENODO
Descripción
<p>ZENODO ha sido construido y desarrollado por investigadores, para garantizar que todos puedan unirse a Open Science.</p> <p>El proyecto OpenAIRE, a la vanguardia de los movimientos de acceso abierto y datos abiertos en Europa, fue encargado por la CE para respaldar su incipiente política de datos abiertos al proporcionar un repositorio general para la investigación financiada por la CE. CERN, socio de OpenAIRE y pionero en código abierto, acceso abierto y datos abiertos, proporcionó esta capacidad y ZENODO se lanzó en mayo de 2013.</p> <p>En apoyo de su programa de investigación, el CERN ha desarrollado herramientas para la gestión de Big Data y ha ampliado las capacidades de la Biblioteca Digital para Open Data. A través de ZENODO, estas herramientas de Big Science podrían compartirse de manera efectiva con la cola larga de la investigación.</p>
Principales Características
<ul style="list-style-type: none">▪ Promovido por la Comisión Europea a través del proyecto OpenAire y alojado en CERN (lo que proporciona seguridad al futuro de los datos alojados).▪ Cumplen con el criterio FAIR.▪ Es gratuito.▪ Permite la reserva de DOI.▪ Integración con ORCID, permite integrarlo con la plataforma ARGOS (una herramienta para la gestión de PGD).▪ Permite subir 50GB de información por publicación.▪ Asignación de licencias Creative Commons, asociar los datos con una convocatoria, etc.▪ Publicando en ZENODO, se cumpliría con los requerimientos de las convocatorias internacionales (Horizonte Europa, etc) y nacionales.

Definición de perfiles profesionales necesarios para la Gestión de Datos

En la actualidad existen perfiles profesionales bien definidos que pueden realizar este tipo de funciones (Gestor de Datos) así como cursos específicos para desarrollar habilidades específicas en este campo.

Entre los perfiles profesionales que pueden actuar como Gestores de Datos encontramos entre muchos otros:

- Bibliotecario / documentalista especializado en Ciencias de la Salud
- Grado en Información y Documentación
- Grado en Ciencia de Datos
- Máster en Big Data

También pudieran establecerse grupos de trabajo interdisciplinarios que incluyan, Informáticos, Matemáticos, Biólogos, etc.

Lo que consideramos importante no es tanto el perfil como la formación que reciban las personas que llevarán este tema en cada IIS, estas acciones de formación debieran ser centralizadas por el ISCIII.

Por otra parte, como mismo existen convocatorias para la Contratación de personal técnico bioinformáticos, contratos de Gestión en Investigación, debiera lanzarse convocatorias para la contratación de Gestores de Datos.

Conclusiones

Recomendamos no duplicar esfuerzos en la creación de nuevos repositorios ya que tanto a nivel Nacional como Internacional existen muchos repositorios agrupados por tipo de datos, enfocando los esfuerzos donde aún no existen repositorios para datos o disciplinas específicas.

Recomendar a los investigadores la necesidad de utilizar los repositorios actuales, y no tanto en intentar generar un repositorio nuevo que cumpla con unas características concretas, ya que los repositorios, dependiendo del campo de investigación deben aglutinar los datos de cada área científica. Podría ser de utilidad para los investigadores disponer de un listado de repositorios en función del tipo de datos que almacenan, como han hecho otras iniciativas como, por ejemplo, ELIXIR.

En el caso de que estratégicamente sea conveniente el desarrollo de un Repositorio nuevo porque se detecta una necesidad concreta no cubierta por los repositorios actuales, la sugerencia de este grupo de trabajo es que sea parte de la Plataforma de Datos REPISALUD. El desarrollo podría hacerse en un contexto nacional o internacional, dependiendo de la naturaleza y complejidad de este nuevo repositorio. Esta plataforma puede integrarse en la Plataforma Recoleta, de forma que los datos que se recojan sean interoperables según los estándares de la comunidad mundial.

En cuanto a los perfiles profesionales de los gestores de datos, éstos deberían ser profesionales de los siguientes perfiles:

- Bibliotecario /documentalista especializado en Ciencias de la Salud
- Grado en Información y Documentación
- Grado en Ciencia de Datos
- Máster en Big Data

Es importante la realización de acciones formativas centralizadas por el ISCIII dirigidas fundamentalmente a perfiles afines a la Gestión de Datos con el objetivo de que el personal que esté al frente de la Gestión de Datos en los IIS adquiera las habilidades necesarias según el repositorio o plataforma que el ISCII haya seleccionado y/o desarrollado para la Gestión de Datos de los IIS.

Desde el ISCIII debieran abrirse convocatorias de ayudas destinadas a la contratación en los IIS de Gestores de Datos.

Necesidades formativas para la realización de Planes de Gestión de Datos

Objetivo: Identificar las necesidades formativas para el correcto desarrollo de los planes de gestión de datos.

Introducción y Metodología

Se ha realizado una encuesta en el ámbito de los Instituto de Investigación Sanitaria, abriendo la posibilidad a que estos lo reenviaran otras instituciones que consideraran relevantes.

La encuesta ha sido anónima, ha constado de 12 preguntas de fácil cumplimentación y ha tenido 26 respuestas.

Se adjuntan las 12 cuestiones incluidas en la encuesta realizada:

1. ¿Cuál es el perfil de la persona que está cumplimentando la encuesta?
 - Investigador
 - Personal de gestión
 - Personal de apoyo (datamanager, ...)
 - Otro (especificar)
2. ¿Existe una persona responsable de la gestión de datos en la institución?
 - Sí
 - No
3. En caso afirmativo, ¿cuál es el perfil de este responsable?
 - Investigador
 - Personal de gestión
 - Personal de apoyo (datamanager, ...)
 - Otro (especificar)
4. ¿Incluye la oferta formativa de la institución formación en gestión de datos?
 - Sí
 - No
 - No lo sé
5. En caso afirmativo, ¿a qué personal está destinada? (se puede seleccionar más de uno)
 - Investigador Predoctoral
 - Investigador Postdoctoral
 - Investigador Senior
 - Técnicos de Apoyo
 - Personal de Gestión
 - Otros (especificar): _____
6. ¿En qué áreas imparte el centro formación en relación con la gestión de datos? (se puede seleccionar más de uno)
 - Formación general sobre gestión de datos
 - Formación específica sobre los aspectos éticos de la gestión de datos
 - Formación específica para realizar Planes de Gestión de Datos
 - Formación específica sobre herramientas para realizar Planes de Gestión de Datos

- Formación específica en almacenamiento y conservación de datos
 - Otros: _____
7. ¿Qué formato tiene la formación ofertada? (se puede seleccionar más de uno)
- Curso de 10 horas o menos
 - Curso de más de 10 horas
 - Seminario
 - Talleres prácticos
 - Otros(especificar): _____
8. ¿Hay alguna formación que hayas recibido y que quieras destacar (del instituto o de cualquier otra institución)?
- Texto libre
9. ¿Qué formación consideras prioritaria impartir en relación con la gestión de datos? (se puede seleccionar más de uno)
- Formación general sobre gestión de datos
 - Formación específica sobre los aspectos éticos de la gestión de datos
 - Formación específica para realizar Planes de Gestión de Datos
 - Formación específica sobre herramientas para realizar Planes de Gestión de Datos
 - Formación específica en almacenamiento y conservación de datos
 - Otros: _____
10. ¿A quién iría dirigida la formación? (se puede seleccionar más de uno)
- Investigador Predoctoral
 - Investigador Postdoctoral
 - Investigador Senior
 - Técnicos de Apoyo
 - Personal de Gestión
 - Otros: _____
11. ¿Qué formato debería tener la formación? (se puede seleccionar más de uno)
- Curso de 10 horas o menos
 - Curso de más de 10 horas
 - Seminario
 - Talleres prácticos
 - Otros(especificar): _____
12. Agradeceremos cualquier otra consideración u observación.
- Texto libre

Resultados

1. **¿Cuál es el perfil de la persona que está cumplimentando la encuesta?**
 - Investigador: 3,8%
 - Personal de gestión: 61,5%
 - Personal de apoyo (datamanager, ...): 11,5%
 - Directores Científicos: 23,2%
2. **¿Existe una persona responsable de la gestión de datos en la institución?**
 - Sí: 53,8%
 - No: 46,2%
3. **En caso afirmativo, ¿cuál es el perfil de este responsable?**
 - Investigador: 7,7%
 - Personal de gestión: 46,2%
 - Personal de apoyo (datamanager, ...): 23,1%
 - Dirección (temporal): 7,7%
 - Investigador: 7,7%
 - Varias personas: 7,7%
 - Ninguno: 7,7%
4. **¿Incluye la oferta formativa de la institución formación en gestión de datos?**
 - Sí: 43,2%
 - No: 53,8%
 - No lo sé: 3,8%
5. **En caso afirmativo, ¿a qué personal está destinada? (se puede seleccionar más de uno)**
 - Investigador Predoctoral: 10 (76,9%)
 - Investigador Postdoctoral: 10 (76,9%)
 - Investigador Senior: 11 (84,6%)
 - Técnicos de Apoyo: 8 (61,5%)
 - Personal de Gestión: 10 (76,9%)
 - Todo el personal: 1 (7,7%)
6. **¿En qué áreas imparte el centro formación en relación con la gestión de datos? (se puede seleccionar más de uno)**
 - Formación general sobre gestión de datos: 8 (57,1%)
 - Formación específica sobre los aspectos éticos de la gestión de datos: 3 (21,4%)
 - Formación específica para realizar Planes de Gestión de Datos: 5 (35,7%)
 - Formación específica sobre herramientas para realizar Planes de Gestión de Datos: 3 (21,4%)
 - Formación específica en almacenamiento y conservación de datos: 2 (14,3%)
 - Formación sobre LOPD: 1 (7,1%)
 - REDCAP: 1 (7,1%)
7. **¿Qué formato tiene la formación ofertada? (se puede seleccionar más de uno)**

- Curso de 10 horas o menos: 6 (46,2%)
 - Curso de más de 10 horas: 2 (15,4%)
 - Seminario: 4 (30,8%)
 - Talleres prácticos: 1 (7,7%)
 - Jornada informativa: 1 (7,7%)
 - Otros: 1 (7,7%)
8. ¿Hay alguna formación que hayas recibido y que quieras destacar (del instituto o de cualquier otra institución)?

Cursos del UIMP /ISCIII y se basa en mi experiencia en grupo ELSI del Proyecto B1Mgenomes, IMpact Genómica, etc.
PGD para gestores ISCiii
"Gestión de datos de investigación y Planes de Gestión de datos (PGD) para gestores de proyectos de los Institutos de Investigación Sanitaria", organizado por la Biblioteca Nacional de Ciencias de la Salud.
Píldoras formativas del CSUC sobre curación de datos y un curso de gestión de datos de investigación para gestores, del ISCIII.
GESTIÓN DE DATOS DE INVESTIGACIÓN Y PLANES DE GESTIÓN DE DATOS (PARA GESTORES) ofertado por el ISCIII
Jornada sobre Planes de Gestión del ISCIII
Aplicación ética y DM (impartida por BIB Barcelona).
DATOS DE INVESTIGACIÓN: PLANIFICACIÓN, GESTIÓN, DISEMINACIÓN, PRESERVACIÓN Y EXPLOTACIÓN EN EL MARCO DE LA CIENCIA ABIERTA - https://upo.es/olavideencarmona/cursos-de-verano/cursos/cc17/index.html
Formando para Formar en Ciencia Abierta - ISCIII
Seminarios de REDCap con REDCapCon Hispano

9. ¿Qué formación consideras prioritaria impartir en relación con la gestión de datos? (se puede seleccionar más de uno)
- Formación general sobre gestión de datos: 16 (61,5%)
 - Formación específica sobre los aspectos éticos de la gestión de datos: 11 (42,3%)
 - Formación específica para realizar Planes de Gestión de Datos: 22 (84,6%)
 - Formación específica sobre herramientas para realizar Planes de Gestión de Datos: 15 (57,5%)
 - Formación específica en almacenamiento y conservación de datos: 17 (65,4%)
 - Formación en el uso de repositorios de datos: 1 (3,8%)

10. ¿A quién iría dirigida la formación? (se puede seleccionar más de uno)
- Investigador Predoctoral: 17 (65,4%)
 - Investigador Postdoctoral: 21 (80,8%)
 - Investigador Senior: 21 (80,8%)
 - Técnicos de Apoyo: 15 (57,7%)
 - Personal de Gestión: 19 (73,1%)
 - Responsables de gestión: 1 (3,8%)
 - Todo el personal: 1 (3,8%)
 - Datamanager: 1 (3,8%)
11. ¿Qué formato debería tener la formación? (se puede seleccionar más de uno)
- Curso de 10 horas o menos: 20 (76,9%)
 - Curso de más de 10 horas: 7 (26,9%)
 - Seminario: 8 (30,8%)
 - Talleres prácticos: 13 (50%)
12. Agradeceremos cualquier otra consideración u observación.

Es necesario ir formando de modo general a todas las personas implicadas. Y luego dirigirse a colectivos y profesionales específicos implicados en cada una de las fases de la gestión de los datos.
--

Idealmente un formato híbrido online en directo + material de trabajo en diferido.
--

Aunque no disponemos de personal específico para la gestión del dato, existen figuras referentes para distintas cuestiones relacionadas con el tema. En estos momentos nos encontramos en una fase de reestructuración y organización en torno a la digitalización y gestión del dato, por lo que en un periodo a corto plazo se podrán cubrir todas estas necesidades.

Los planes de gestión de datos tienen que ir alineados con la normativa de protección de datos y de propiedad industrial/intelectual.

Consideramos necesario una formación más intensa para los responsables de la gestión de datos de ISABIAL para poder apoyar a los investigadores en todo tipo de situaciones. Además, la formación en cursos de aspectos éticos no lo consideramos necesario debido al CEIm de nuestra institución.
--

Conclusiones

Casi un 50% de las instituciones no tienen un responsable de gestión de datos. Se considera prioritario promocionar esta figura dentro de los IIS y ofertar formación que permita desempeñar adecuadamente su labor, tanto de responsable como de promotor dentro de la institución.

La mayoría de las instituciones no ofertan formación en gestión de datos, y si la ofertan es genérica orientada a todo el personal. En las prioridades de formación demandadas por las instituciones, se recoge una formación genérica, pero también una formación más práctica y concreta, con talleres o cursos cortos, que capacite para la realización efectiva de Planes de Gestión de Datos.

La mayoría de la formación recibida ha sido impartida por el ISCIII. Sería recomendable que el ISCIII siguiera desarrollando formación en esta línea, tanto genérica como más práctica, independientemente de que las instituciones lo integren en sus planes de formación. De cara a promocionar y armonizar la formación en los centros, tanto básica como especializada, sería recomendable actividades de "Formación de formadores" por parte del ISCIII, con el objetivo de que los institutos acaben siendo autónomos.

Adicionalmente, se podría promover formación de otros perfiles encargados de la extracción, anonimización y preparación de los datos para incluirlos en los repositorios, así como en los aspectos relacionados con la protección de datos personales, dada su especial relevancia en el ámbito de la salud.