Información, buenas prácticas y compromiso para usuarios de PRISIB



Pau Pericàs

Plataforma de Recerca en Informació en Salut de les Illes Balears (PRISIB)

Institut de Investigació Sanitària de les Illes Balears (Idisba)

Servei de Salut de les Illes Balears (IBSalut)

pau.pericas@ibsalut.es

28/02/2024

**Índice**

[Introducción 3](#__RefHeading___Toc794_2961146142)

[La PRISIB 3](#__RefHeading___Toc1177_2961146142)

[Presentación 3](#__RefHeading___Toc1179_2961146142)

[Fuentes de datos 3](#__RefHeading___Toc1183_2961146142)

[Participación en proyectos e instituciones 4](#__RefHeading___Toc1185_2961146142)

[Servicios 5](#__RefHeading___Toc1603_757473047)

[Uso secundario de datos clínicos 6](#__RefHeading___Toc1199_2961146142)

[El plan de gestión de datos 6](#__RefHeading___Toc810_2961146142)

[Principios FAIR 7](#__RefHeading___Toc806_2961146142)

[Los metadatos 8](#__RefHeading___Toc808_2961146142)

[Problemas del uso de datos clínicos 8](#__RefHeading___Toc3040_2961146142)

[Protección de datos 9](#__RefHeading___Toc796_2961146142)

[Ventajas del uso de datos clínicos 9](#__RefHeading___Toc3042_2961146142)

[Calidad de los datos 10](#__RefHeading___Toc3044_2961146142)

[Medidas para mejorar el uso de los datos clínicos 11](#__RefHeading___Toc3046_2961146142)

[Tipos de estudio 11](#__RefHeading___Toc1203_2961146142)

[Publicación de los datos 12](#__RefHeading___Toc812_2961146142)

[Modelo federado 13](#__RefHeading___Toc1330_2961146142)

[Annexos 14](#__RefHeading___Toc1088_4231230472)

[Ley de autonomía del paciente 14](#__RefHeading___Toc1207_2961146142)

[GDPR 15](#__RefHeading___Toc798_2961146142)

[LOPDGDD 17](#__RefHeading___Toc1834_1563066921)

[RD Estudios observacionales con medicamentos 20](#__RefHeading___Toc1209_2961146142)

[Ley de la Ciencia, la Tecnología y la Innovación 21](#__RefHeading___Toc1211_2961146142)

[Bibliografía 23](#__RefHeading___Toc363_1859376467)

# Introducción

El uso de la información sanitaria tiene diversas y profundas implicaciones éticas, legales y técnicas que este documento pretende resumir. Necesariamente la información aquí proporcionada es incompleta y susceptible de quedar obsoleta por los cambios en las leyes, normas y costumbres que atañen al uso de datos clínicos para la investigación, pero el Comité Científico Técnico de la PRISIB velará por mantener este documento relevante y actualizado

# La PRISIB

## Presentación

La PRISIB es la Plataforma de Investigación con Información Sanitaria del Institut d’Investigació Sanitària de les Illes Balears (IdISBa).

Esta plataforma se creó como colaboración entre el Servei de Salut de les Illes Balears (IBSalut) y el Instituto de investigación con la finalidad de disponer de una vía de acceso única para obtener datos válidos y fiables provenientes de la historia clínica electrónica y otras fuentes complementarias para crear conocimiento, apoyar a la investigación y facilitar la innovación y la evaluación en el ámbito sanitario.

Para garantizar el correcto uso de los datos, todos los proyectos de investigación son valorados por el Comité Científico – Técnico de la PRISIB. Éste aconseja a la plataforma y evalúa las solicitudes que se reciben, idealmente de manera previa o simultánea a la evaluación por parte del Comité de Ética para que las aportaciones que se puedan hacer en cuanto al manejo de datos sean aprobadas junto con el protocolo del proyecto.

## Fuentes de datos

En términos generales, cualquier fuente de datos recogidos durante la actividad del servicio de salud es susceptible de ser usada por la PRISIB para proyectos de investigación siempre que este uso haya sido aprobado por los Comités y Comisiones de Investigación y de Ética pertinentes.

En la actividad asistencial del IBSalut se hace uso de múltiples repositorios de información que no siempre se encuentran interconectados y cuya gobernanza puede ser compleja de navegar. Por este motivo resulta imposible realizar un catálogo completo de toda la información potencialmente accesible a la PRISIB.

La siguiente es una lista de las fuentes de datos más utilizadas por la plataforma:

* SIAP: Sistema de Información de Atención Primaria. Registra toda la actividad de los centros de salud y sus unidades complementarias.
* Receta Electrónica: Registra todas las prescripciones realizadas usando la tarjeta sanitaria y las recogidas de fármacos efectuadas en las farmacias.
* Laboratorio: Todos los laboratorios del sistema de salud están informatizados y sus datos son explotables de manera conjunta.
* CMBD: Conjunto Mínimo Básico de Datos. Repositorio con la codificación de todos los informes de alta hospitalaria.

Se ha accedido en el pasado de manera puntual a otras fuentes de información para su uso en investigación:

* PACS corporativo: Sistema de archivo y comunicación de imágenes del IBSalud donde se almacenan y se consultan en la práctica clínica las exploraciones radiológicas realizadas en los hospitales.
* RADELEC: Sistema de recogida y análisis de los electrocardiogramas realizados tanto en atención primaria como en los centros hospitalarios
* Biobanco IdISBa: Este Biobanco cuenta con diversas colecciones de muestras recogidas con fines de investigación que la PRISIB puede complementar con información clínica de los donantes.

## Participación en proyectos e instituciones

Dado el creciente interés del uso de los datos sanitarios en la comunidad científica, la plataforma participa en los siguientes proyectos orientados a la exploración de los datos para la investigación que pueden servir para mejorar el funcionamiento de la plataforma asimilándolo con el de otras instituciones equivalentes y participar en proyectos de investigación multicéntricos tanto si son originados en otras instituciones como si son impulsados desde la comunidad investigadora balear.

* [IMPaCT-Data](https://impact-data.bsc.es/): es el Programa de Ciencia de Datos de la Infraestructura de Medicina de Precisión asociada a la Ciencia y la Tecnología promovida por el Instituto Carlos III que persigue apoyar el desarrollo de un sistema común, interoperable e integrado, de recogida y análisis de datos clínicos.
* [EHDEN](https://ehden.eu/):es la European Health Data & Evidence Network, promovida por el programa Horizonte 2020 y EFPIA para crear una red federada de centros de investigación con datos sanitarios que utilicen los mismos modelos y codificación de los datos (OMOP CDM) con la intención de generar evidencia científica sólida y fiable de manera rápida y eficiente.
* [OHDSI](https://www.ohdsi.org/): el programa Observational Health Data Sciences and Informatics es una colaboración internacional para el desarrollo de herramientas de análisis de información sanitaria a gran escala usando código abierto y el modelo federado de datos.
* [EOSC-A](https://eosc.eu/): La European Open Science Cloud Association es la entidad legal creada por la Comisión Europea para gobernar el portal de datos y servicios europeo para la investigación científica.

## Servicios

Los servicios que se pueden solicitar a la PRISIB son los siguientes:

* Extracción de datos de la historia clínica electrónica para proyectos de investigación. El proyecto que solicite datos a la PRISIB debe responder a un protocolo debidamente elaborado por el equipo investigador que lo desarrolla. Este protocolo debe haber sido aprobado por la correspondiente comisión de investigación y será evaluado por el Comité Científico-Técnico de la PRISIB y el CEIm.

En la evaluación por parte del CCT de la PRISIB se juzgarán los aspectos relacionados con la gestión de datos, como la adecuación de los datos solicitados a los objetivos declarados del estudio o los medios utilizados para registrar, almacenar y analizar los datos. Por ello se recomienda la redacción de un plan de gestión de datos como anexo al protocolo (ver apartado 'El plan de gestión de datos').

Como parte de este servicio se contempla desde la creación del modelo de datos del proyecto, la identificación de las fuentes de datos, el mapeo de variables, la creación de los algoritmos de extracción, la programación de extracciones periódicas, la georreferenciación anonimizada de los datos, el análisis de calidad de los datos, estadística descriptiva y anonimidad.

* Consultas para estudios de viabilidad o estudios exploratorios. Servicio gratuito de consulta para estudios de viabilidad que proporciona datos aproximados, agregados, sin control de calidad y cuyo objetivo es tener una primera aproximación para valorar si se puede desarrollar un estudio más ambicioso. Son estudios en los que se quiere saber la N que existe en las bases de datos de un determinado problema o cómo está registrada determinada información de interés.
* Visualización de datos vinculados a los proyectos de investigación. Estas visualizaciones pueden ir desde la generación de mapas vectoriales o gráficos estadísticos a la implementación de informes interactivos en formato web.
* Implementación de sistemas de recogida de datos por la investigación, de aleatorización de casos y pseudonimización.
* Custodia de datos de investigación, que se almacenarían en el repositorio de datos de la Plataforma en el centro de procesamiento de datos del Servicio de Salud.
* Asesoría en la redacción de solicitudes,protocolos, proyectos de investigación, planes de interoperabilidad, planes de gestión de datos y otra documentación relacionada con el uso de datos en el ámbito de la investigación sanitaria.

# Uso secundario de datos clínicos

## El plan de gestión de datos

Para la mayoría de convocatorias competitivas se está solicitando la presentación de un Plan de Gestión de Datos (PGD) junto con el protocolo del proyecto. El PGD es un documento en el que se especifican todos los aspectos de tratamiento y protección de datos durante y después de la realización de un proyecto de investigación.

El PGD es un documento que facilita la claridad de los aspectos legales del tratamiento de datos, la colaboración con otros investigadores, la evaluación de propuestas y la resolución de dudas referentes al uso de datos.

De manera resumida, un PGD debería especificar:

* De dónde se obtendrán los datos, que formato tendrán, si van a ser estáticos o cambiarán con el tiempo, qué cantidad de datos se espera recoger
* Cómo se van a documentar y describir estos datos, qué criterios se establecerán para validar su calidad y cómo se materializarán
* Dónde se van a almacenar y cómo se van a proteger estos datos frente a pérdidas o sustracciones
* Qué problemas de confidencialidad y privacidad se preveen y cómo se van a afrontar
* Cómo podrían otros investigadores encontrar, usar y acceder a esos datos
* Cómo, porqué y durante cuánto tiempo se van a archivar esos datos
* Quién será responsable de cada uno de los aspectos definidos en el PGD

Existen diferentes modelos de PGD y varias herramientas y guías que ayudan a su redacción, entre las que destacamos:

* [ARGOS (Openaire)](https://argos.openaire.eu/splash/)
* [DMPonline (Data Curation Center)](https://dmponline.dcc.ac.uk/)
* [Cora (Consoci de Serveis Universitaris de Catalunya)](https://dmp.csuc.cat/)
* [10 Pasos para elaborar un plan de gestion de datos (CRUE)](https://www.rebiun.org/acceso-abierto/10-pasos-para-elaborar-un-plan-de-gestion-de-datos)

La redacción y publicación del PGD se considera como una de las piedras angulares de la aplicación de los principios FAIR

## Principios FAIR

Estos principios se resumen en la publicación Mark D. Wilkinson et al., “The FAIR Guiding Principles for Scientific Data Management and Stewardship,” Scientific Data (March 15, 2016) de la siguiente manera

* Para ser localizable (Findable):

F1. A los datos y metadatos se les asigna un identificador global único y persistente (pe DOI).

F2. los datos se describen con metadatos enriquecidos (definidos por R1 a continuación)

F3. Los metadatos incluyen clara y explícitamente el identificador de los datos que describen.

F4. Los datos se registran o indexan en un recurso con capacidad de búsqueda.

* Para ser Accesible:

A1. Los datos y metadatos se pueden recuperar mediante su identificador utilizando un protocolo de comunicación estandarizado.

A1.1 el protocolo es abierto, gratuito y de implementación universal.

A1.2 el protocolo permite un procedimiento de autenticación y autorización, cuando sea necesario.

A2. Los metadatos son accesibles, incluso cuando los datos ya no están disponibles.

* Para ser Interoperable:

I1. Los datos y metadatos utilizan un lenguaje formal, accesible, compartido y ampliamente aplicable para la representación del conocimiento.

I2. Los datos y metadatos utilizan vocabularios que siguen los principios FAIR.

I3. Los datos y metadatos incluyen referencias calificadas a otros datos y metadatos.

* Para ser Reutilizable:

R1. Los metadatos describen detalladamente los datos con atributos precisos y relevantes.

R1.1. Los datos y metadatos se publican con una licencia de uso de datos clara y accesible.

R1.2. Los datos y metadatos están asociados a una procedencia detallada.

R1.3. Los datos y metadatos cumplen con los estándares comunitarios relevantes para el dominio.

## Los metadatos

Para hacer practicable la aplicación de los principios FAIR se necesita una descripción detallada de los datos que se generan en un estudio. Esta descripción es la que llamamos metadatos, es decir, datos sobre los datos.

En la investigación sanitaria frecuentemente los datos no se pueden hacer públicos. Así la publicación de los metadatos es de aun mayor relevancia, ya que en ellos se explicará en que consiste el conjunto de datos y las personas y procedimientos con los que contactar para solicitar el acceso a los datos.

Existen diferentes modelos para documentar los metadatos específicos para diversos campos de investigación sin que exista una solución universal a todas las casuísticas posibles.

Para facilitar la comunicación entre investigadores y técnicos, así como la posterior publicación de al menos los metadatos, la PRISIB usa un modelo de datos propio para describir los conjuntos de datos que genera. Este modelo se puede encontrar en el repositorio institucional [Docusalut](http://hdl.handle.net/20.500.13003/18160) o en el repositorio abierto [Zenodo](http://10.5281/zenodo.7185046).

## Problemas del uso de datos clínicos

Los datos que se pueden encontrar en la historia clínica de los pacientes se recogen con una finalidad asistencial, no científica y eso hace que su uso para la investigación se deba hacer con conocimiento de las limitaciones que esto implica y las necesarias cautelas.

* Datos observacionales: Al tratarse de datos puramente observacionales no es posible atribuir causalidad en las hipótesis que se pretenden contrastar.
* Datos incompletos: Es difícil, si no imposible, saber por qué un dato no se encuentra en la historia clínica. Puede que no se midiera, o que se midiera y no se registrara y no suele haber manera de conocer que sesgo se introduce con esta omisión.
* Datos no fidedignos: Al no ser el objetivo del registro la propia recogida de datos, como se haría en un experimento científico, si no la atención a los pacientes, puede que se cometan errores en la introducción de los datos.
* Abundantes datos: El hecho de disponer de una gran cantidad de datos puede hacer aparecer asociaciones espurias a las que hay que evitar atribuir correlación relevante.
* Sesgos: Al recogerse los datos con una finalidad asistencial éstos están sesgados desde el inicio, ya que sólo se recoge información de aquellos usuarios que acuden al sistema sanitario público quedando excluidos aquellos que cuentan con asistencia privada o aquellos que por la causa que sea, no hacen uso de los recursos sanitarios (pe: situaciones de exclusión social)

## Protección de datos

El uso secundario de datos sanitarios implica siempre un equilibrio entre tres principios éticos de la máxima importancia:

* Confidencialidad

El personal sanitario está obligado a no revelar información confidencial del paciente sin su conocimiento y autorización

* Autonomía

Todas las personas tienen un valor intrínseco incondicional por lo que deberían tener el poder para tomar sus propias decisiones racionales y morales, y todos deberían poder ejercer su capacidad de manera autónoma

* Beneficencia

Es la obligación del personal médico de actuar en beneficio de sus pacientes y es la base de múltiples normas éticas que implican la protección y defensa de los derechos, la no maleficencia, la asistencia a personas en peligro, la eliminación de riesgos para la salud y la ayuda a personas con discapacidad

Hacer uso de la información del paciente sin un consentimiento previo a la recogida de información vulneraria el principio de confidencialidad y autonomía. No hacerlo supondría incumplir con el principio de beneficencia al no hacer uso de los recursos disponibles para la mejora de la salud de los pacientes. Este dilema ético está presente en la mayoría de estudios observacionales y su equilibrio es un tema en constante discusión y movimiento tanto en la esfera de la bioética como en la legal.

Como norma general, se debería obtener el consentimiento explícito y por escrito de todos los pacientes para usar sus datos. Pero de acuerdo al RD 957/2020 del 3 de noviembre, por el que se regulan los estudios observacionales con medicamentos de uso humano: "se podrá eximir de solicitar el consentimiento informado, siempre que el CEIm considere que la investigación observacional tiene un valor social importante, que su realización no sería factible o viable sin dicha dispensa, y que entraña riesgos mínimos para los participantes."

## Ventajas del uso de datos clínicos

Por otro lado, el uso de datos de la historia clínica presenta algunas ventajas con respecto al uso de datos recogidos de manera prospectiva en el contexto de un ensayo clínico.

* Se miden múltiples intervenciones
* Se valora sobre pacientes reales representativos de la práctica habitual
* Se realiza en el entorno real donde se aplican las medidas estudiadas
* Se cuenta con una mayor evolución de los datos
* Se puede ampliar el número de variables observadas
* Permite hacer seguimiento de los cambios en los resultados
* Es más barato y sencillo

Además, permiten analizar cuestiones relevantes que serían muy difíciles o imposibles de medir en un ensayo clínico aleatorizado:

* Impacto de la implantación de nuevas tecnologías
* Estudio de enfermedades raras
* Detección de variables relevantes en la toma de decisiones
* Monitorización del acceso y uso de los recursos sanitarios
* Detección de necesidades de salud no cubiertas
* Detección del uso inadecuado de recursos sanitarios
* Impacto de una intervención sobre los circuitos asistenciales

## Calidad de los datos

La calidad de los datos depende de su propósito. Un dato puede ser de alta calidad para su propósito primario pero si se le pretende dar un uso secundario puede que presente algún defecto en alguna de las diferentes dimensiones en que se puede describir la calidad del dato.

“Los datos son de alta calidad si son adecuados por los usos previstos en las operaciones, la toma de decisiones y la planificación. Los datos son aptos para su uso si carecen de defectos y tienen las características deseadas"

Data Quality: The Field Guide, Thomas C. Redman, Ph.D. Digital Press,2001

En la publicación *Sidi,Fatimah, et al. "Data quality: A survey of data quality dimensions." 2012 se describen las diferentes d*imensiones de la calidad de los datos:

Exactitud: ¿Representan los datos correctamente la entidad o evento real?

Consistencia: ¿Los datos no contienen contradicciones?

Disponibilidad: ¿Se puede acceder a los datos ahora y a lo largo del tiempo?

Integridad: ¿Incluyen los datos todos los elementos que representan la entidad o el evento?

Completitud: ¿Incluyen los datos todos los elementos que se pretenden representar?

Conformidad: ¿Siguen los datos las normas aceptadas?

Procesabilidad: ¿Son los datos legibles por las máquinas?

Credibilidad: ¿Se basan en fuentes fiables?

Pertinencia: ¿Incluyen una cantidad adecuada de datos?

Vigencia: ¿Representan la situación real y se publican con suficiente antelación?

## Medidas para mejorar el uso de los datos clínicos

Para evitar magnificar asociaciones espurias y medir realmente lo que se pretende en la pregunta científica, es necesario disponer de un modelo causal claro. Esto implica que antes de decidir qué información extraer de los registros clínicos hay que tener claros todos los factores de confusión y variables intermedias que pueden oscurecer y desdibujar las relaciones entre las variables que queremos investigar.

Para representar estas relaciones se recomienda el uso de diferentes diagramas que permiten exponer claramente las asociaciones que se espera encontrar en los datos y las que se quieren medir para responder a la pregunta del estudio, como los [grafos acíclicos dirigidos](http://10.1016/j.chest.2020.03.011).

Otra medida básica para paliar los problemas del uso secundario de datos es establecer unos criterios de validación y verificación de los datos antes de su extracción para mejorar su consistencia y credibilidad sin introducir nuevos sesgos en los que se podría concurrir al decidir la validez de los datos cuando ya se conocen.

Diferentes técnicas estadísticas pueden ayudar a comprender y paliar los sesgos, factores de confusión y heterogeneidades presentes en los datos que pueden introducir varianzas y desviaciones en nuestros resultados.

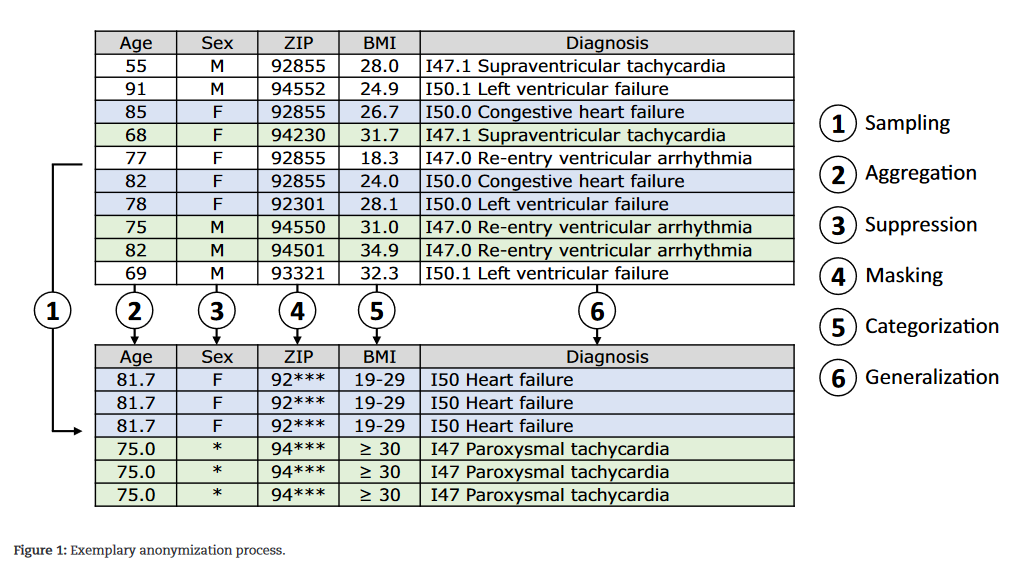
## Tipos de estudio

Pese a las limitaciones comentadas, el uso secundario de datos clínicos permite diferentes modelos de estudios en función de si se registra una intervención, si se comparan datos antes y después de dicha intervención o si se comparan resultados con un grupo de control:

* Puramente observacionales: Recogida de datos en un punto concreto del tiempo
* Post-exposición: Recogida de datos de pacientes seleccionados a partir de un evento
* Pre-Post exposición: Recogida de datos de pacientes seleccionados antes y después de un evento
* Ensayo clínico pragmático: Recogida de datos de pacientes seleccionados antes y después de un evento y de un grupo comparable sin la exposición a dicho evento siendo la exposición al evento aleatoria

## Publicación de los datos

La publicación de datos personales está fuertemente regulada por las diferentes leyes comunitarias y estatales. Para poder hacer públicos estos datos es necesario garantizar que el conjunto de datos no permite la identificación de ninguna persona y para ello se hacen uso de diferentes técnicas de anonimización que se ejemplifican en la siguiente figura:

Figura 1: A scalable software solution for anonymizing high-dimensional biomedical data, GigaScience, Volume 10, Issue 10, October 2021

Los datos correctamente anonimizados no se consideran datos personales y se pueden publicar abiertamente de acuerdo con las directivas europeas,

Para facilitar la aplicación de estas técnicas sobre grandes conjuntos de datos, existen herramientas de uso público como [ARX – Data Anonymization Tool](https://arx.deidentifier.org/) o [Amnesia (Openaire)](https://amnesia.openaire.eu/)

## Modelo federado

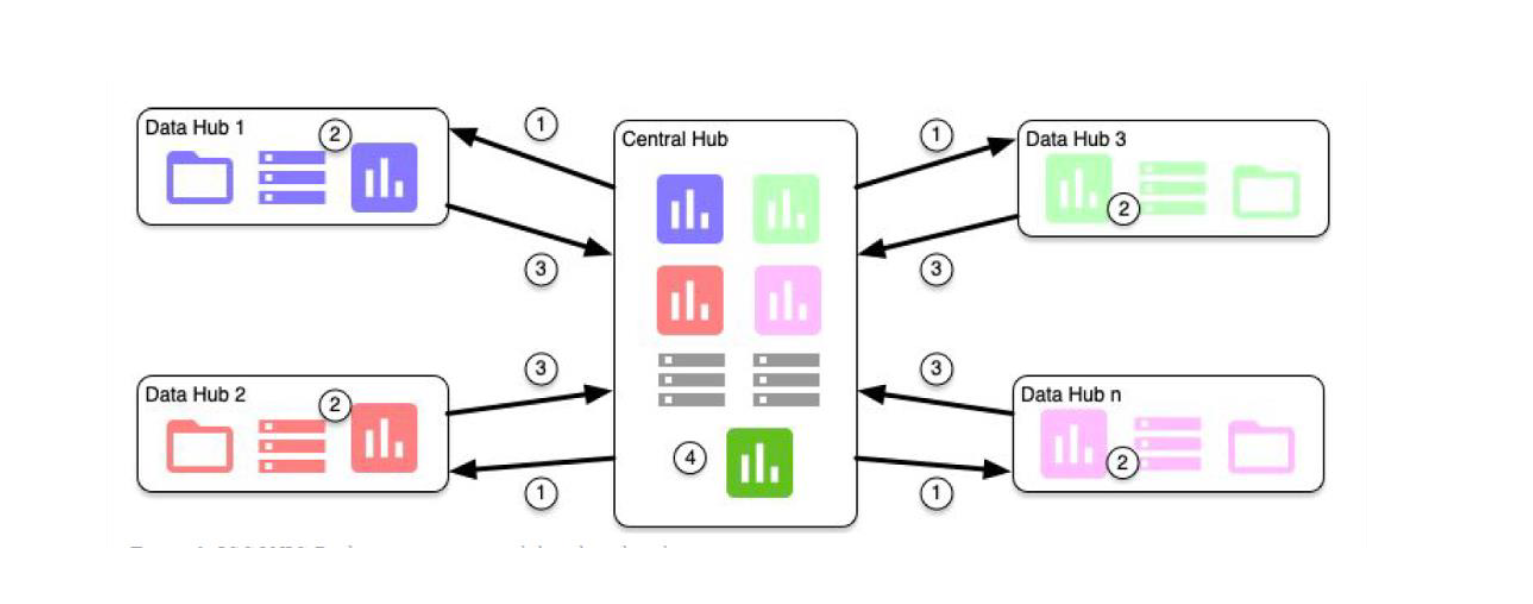
Una forma de generar una evidencia más sólida y eliminar algunos sesgos es la combinación de datos de diferentes fuentes para realizar un estudio multicéntrico.

Pero la cesión de datos personales para la realización de estos estudios, en especial si se tratan de datos de la población entera o de grupos vulnerables, genera muchas reticencias y problemas éticos y legales.

Para facilitar la realización de estos grandes estudios sin comprometer la seguridad y privacidad de los datos, existe la posibilidad de realizar estudios federados. En éstos, un centro coordinador comparte un modelo de datos y una definición clara de los datos a analizar y el algoritmo con el que se analizan los datos con múltiples nodos participantes. Cada nodo participante posteriormente extrae y harmoniza sus datos de acuerdo al modelo de datos y los analiza usando el mismo algoritmo usando sus recursos propios y devuelve al centro coordinador únicamente el resultado agregado de este análisis. Finalmente, el centro coordinador analiza los resultados de los diferentes nodos para dar lugar a los resultados finales.

Este modelo ofrece múltiples ventajas, principalmente la de que no se ceden datos personales en ningún caso, ya que estos los analiza cada nodo dentro de su propio espacio de datos. Además, permite distribuir la carga de trabajo computacional y de almacenamiento de los diferentes conjuntos de datos generados.

Si estos estudios se realizan entre centros que comparten una infraestructura compatible y un modelo de datos harmonizado, la realización de grandes estudios multicéntricos se reduce a la ejecución de un único paquete de análisis en cada nodo.

Iniciativas como EHDEN, [Darwin-EU](https://www.darwin-eu.org/) (liderada por la Agencia Europea del Medicamento) o IMPaCT, buscan implantar este modelo de cooperación que garantiza la gobernanza local de los datos pero permite crear grandes estudios poblacionales a escala internacional.

# Annexos

## Ley de autonomía del paciente

[La Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica](https://www.boe.es/eli/es/l/2002/11/14/41/con) establece lo siguiente en el punto 3 de su artículo 16:

“El acceso a la historia clínica con fines judiciales, epidemiológicos, de salud pública, de investigación o de docencia, se rige por lo dispuesto en la legislación vigente en materia de protección de datos personales, y en la Ley 14/1986, de 25 de abril, General de Sanidad, y demás normas de aplicación en cada caso. El acceso a la historia clínica con estos fines obliga a preservar los datos de identificación personal del paciente, separados de los de carácter clinicoasistencial, de manera que, como regla general, quede asegurado el anonimato, salvo que el propio paciente haya dado su consentimiento para no separarlos.

Se exceptúan los supuestos de investigación previstos en el apartado 2 de la Disposición adicional decimoséptima de la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales.

Asimismo, se exceptúan los supuestos de investigación de la autoridad judicial en los que se considere imprescindible la unificación de los datos identificativos con los clinicoasistenciales, en los cuales se estará a lo que dispongan los jueces y tribunales en el proceso correspondiente. El acceso a datos y documentos de la historia clínica se limita estrictamente a los fines específicos de cada caso.

Cuando ello sea necesario para la prevención de un riesgo o peligro grave para la salud de la población, las Administraciones sanitarias a las que se refiere la Ley 33/2011, de 4 de octubre, General de Salud Pública, podrán acceder a los datos identificativos de los pacientes por razones epidemiológicas o de protección de la salud pública. El acceso habrá de realizarse, en todo caso, por un profesional sanitario sujeto al secreto profesional o por otra persona sujeta, asimismo, a una obligación equivalente de secreto, previa motivación por parte de la Administración que solicitase el acceso a los datos.”

## GDPR

La [General Data Protection Regulation o Reglamento Europeo de Protección de Datos (GDPR/REPD)](https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32016R0679) de 2016 establece el derecho fundamental a la protección de datos de las personas físicas.

Artículo 9

Tratamiento de categorías especiales de datos personales

1. Quedan prohibidos el tratamiento de datos personales que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o las orientación sexuales de una persona física.

2. El apartado 1 no será de aplicación cuando concurra una de las circunstancias siguientes: [...]

i) el tratamiento es necesario por razones de interés público en el ámbito de la salud pública, como la protección frente a amenazas transfronterizas graves para la salud, o para garantizar elevados niveles de calidad y de seguridad de la asistencia sanitaria y de los medicamentos o productos sanitarios, sobre la base del Derecho de la Unión o de los Estados miembros que establezca medidas adecuadas y específicas para proteger los derechos y libertades del interesado, en particular el secreto profesional,

j) el tratamiento es necesario con fines de archivo en interés público, fines de investigación científica o histórica o fines estadísticos, de conformidad con el artículo 89, apartado 1, sobre la base del Derecho de la Unión o de los Estados miembros, que debe ser proporcional al objetivo perseguido, respetar en lo esencial el derecho a la protección de datos y establecer medidas adecuadas y específicas para proteger los intereses y derechos fundamentales del interesado.

Se dice que “cuando el responsable del tratamiento proyecte el tratamiento ulterior de datos personales para un fin que no sea aquel para el que se recogieron, proporcionará al interesado, con anterioridad a dicho tratamiento ulterior, información sobre ese otro fin y cualquier información adicional pertinente […]”. Aunque exime de la obligación de informar específicamente de este tratamiento ulterior “en la medida en que el interesado ya disponga de la información” sobre este tratamiento.

Esta ley además regula las responsabilidades de los responsables y encargados del tratamiento de datos personales partiendo de la “protección de datos desde el diseño y por defecto” y “teniendo en cuenta el estado de la técnica, el coste de la aplicación y la naturaleza, ámbito, contexto y fines del tratamiento, así como los riesgos […] el responsable del tratamiento aplicará medidas técnicas y organizativas apropiadas, como la seudonimización [… y ] la minimización de datos”.

Sobre el Encargado del tratamiento dice, entre otras cosas que “tratará los datos personales únicamente siguiendo instrucciones documentadas del responsable”, “garantizará que las personas autorizadas para tratar datos personales se hayan comprometido a respetar la confidencialidad o estén sujetas a una obligación de confidencialidad de naturaleza estatutaria” y “a elección del responsable, suprimirá o devolverá todos los datos personales una vez finalice la prestación de los servicios de tratamiento, y suprimirá las copias existentes”. “Cuando un encargado del tratamiento recurra a otro encargado para llevar a cabo determinadas actividades de tratamiento por cuenta del responsable, se impondrán a este otro encargado, mediante contrato u otro acto jurídico [...] las mismas obligaciones de protección de datos que las estipuladas en el contrato u otro acto jurídico entre el responsable […]. Si ese otro encargado incumple sus obligaciones de protección de datos, el encargado inicial seguirá siendo plenamente responsable”. Así pues “El encargado del tratamiento notificará sin dilación indebida al responsable del tratamiento las violaciones de la seguridad de los datos personales de las que tenga conocimiento.”

Evaluación de impacto relativa a la protección de datos

1. Cuando sea probable que un tipo de tratamiento, en particular si utiliza nuevas tecnologías, por su naturaleza, alcance, contexto o fines, entrañe un alto riesgo para los derechos y libertades de las personas físicas, el responsable del tratamiento realizará, antes del tratamiento, una evaluación del impacto de las operaciones de tratamiento en la protección de datos personales. Una única evaluación podrá abordar una serie de operaciones de tratamiento similares que entrañen altos riesgos similares.

Artículo 40

Códigos de conducta

1. Los Estados miembros, las autoridades de control, el Comité y la Comisión promoverán la elaboración de códigos de conducta destinados a contribuir a la correcta aplicación del presente Reglamento, teniendo en cuenta las características específicas de los distintos sectores de tratamiento y las necesidades específicas de las microempresas y las pequeñas y medianas empresas.

2. Las asociaciones y otros organismos representativos de categorías de responsables o encargados del tratamiento podrán elaborar códigos de conducta o modificar o ampliar dichos códigos con objeto de especificar la aplicación del presente Reglamento, como en lo que respecta a:

a) el tratamiento leal y transparente;

b) los intereses legítimos perseguidos por los responsables del tratamiento en contextos específicos;

c) la recogida de datos personales;

d) la seudonimización de datos personales;

e) la información proporcionada al público y a los interesados;

f) el ejercicio de los derechos de los interesados;

g) la información proporcionada a los niños y la protección de estos, así como la manera de obtener el consentimiento de los titulares de la patria potestad o tutela sobre el niño;

h) las medidas y procedimientos a que se refieren los artículos 24 y 25 y las medidas para garantizar la seguridad del tratamiento a que se refiere el artículo 32;

i) la notificación de violaciones de la seguridad de los datos personales a las autoridades de control y la comunicación de dichas violaciones a los interesados;

j) la transferencia de datos personales a terceros países u organizaciones internacionales, o

k) los procedimientos extrajudiciales y otros procedimientos de resolución de conflictos que permitan resolver las controversias entre los responsables del tratamiento y los interesados relativas al tratamiento, sin perjuicio de los derechos de los interesados en virtud de los artículos 77 y 79

## LOPDGDD

Estos derechos se desarrollan y regulan en España por la [Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales](https://www.boe.es/eli/es/lo/2018/12/05/3/con).

De especial interés es la Disposición adicional decimoséptima sobre el tratamientos de datos de salud que en su punto 2 dice:

El tratamiento de datos en la investigación en salud se regirá por los siguientes criterios:

a) El interesado o, en su caso, su representante legal podrá otorgar el consentimiento para el uso de sus datos con fines de investigación en salud y, en particular, la biomédica. Tales finalidades podrán abarcar categorías relacionadas con áreas generales vinculadas a una especialidad médica o investigadora.

b) Las autoridades sanitarias e instituciones públicas con competencias en vigilancia de la salud pública podrán llevar a cabo estudios científicos sin el consentimiento de los afectados en situaciones de excepcional relevancia y gravedad para la salud pública.

c) Se considerará lícita y compatible la reutilización de datos personales con fines de investigación en materia de salud y biomédica cuando, habiéndose obtenido el consentimiento para una finalidad concreta, se utilicen los datos para finalidades o áreas de investigación relacionadas con el área en la que se integrase científicamente el estudio inicial.

En tales casos, los responsables deberán publicar la información establecida por el artículo 13 del Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de sus datos personales y a la libre circulación de estos datos, en un lugar fácilmente accesible de la página web corporativa del centro donde se realice la investigación o estudio clínico, y, en su caso, en la del promotor, y notificar la existencia de esta información por medios electrónicos a los afectados. Cuando estos carezcan de medios para acceder a tal información, podrán solicitar su remisión en otro formato.

Para los tratamientos previstos en esta letra, se requerirá informe previo favorable del comité de ética de la investigación.

d) Se considera lícito el uso de datos personales seudonimizados con fines de investigación en salud y, en particular, biomédica. El uso de datos personales seudonimizados con fines de investigación en salud pública y biomédica requerirá:

1.º Una separación técnica y funcional entre el equipo investigador y quienes realicen la seudonimización y conserven la información que posibilite la reidentificación.

2.º Que los datos seudonimizados únicamente sean accesibles al equipo de investigación cuando:

i) Exista un compromiso expreso de confidencialidad y de no realizar ninguna actividad de reidentificación.

ii) Se adopten medidas de seguridad específicas para evitar la reidentificación y el acceso de terceros no autorizados.

Podrá procederse a la reidentificación de los datos en su origen, cuando con motivo de una investigación que utilice datos seudonimizados, se aprecie la existencia de un peligro real y concreto para la seguridad o salud de una persona o grupo de personas, o una amenaza grave para sus derechos o sea necesaria para garantizar una adecuada asistencia sanitaria.

e) Cuando se traten datos personales con fines de investigación en salud, y en particular la biomédica, a los efectos del artículo 89.2 del Reglamento (UE) 2016/679, podrán excepcionarse los derechos de los afectados previstos en los artículos 15, 16, 18 y 21 del Reglamento (EU) 2016/679 cuando:

1.º Los citados derechos se ejerzan directamente ante los investigadores o centros de investigación que utilicen datos anonimizados o seudonimizados.

2.º El ejercicio de tales derechos se refiera a los resultados de la investigación.

3.º La investigación tenga por objeto un interés público esencial relacionado con la seguridad del Estado, la defensa, la seguridad pública u otros objetivos importantes de interés público general, siempre que en este último caso la excepción esté expresamente recogida por una norma con rango de Ley.

f) Cuando conforme a lo previsto por el artículo 89 del Reglamento (UE) 2016/679, se lleve a cabo un tratamiento con fines de investigación en salud pública y, en particular, biomédica se procederá a:

1.º Realizar una evaluación de impacto que determine los riesgos derivados del tratamiento en los supuestos previstos en el artículo 35 del Reglamento (UE) 2016/679 o en los establecidos por la autoridad de control. Esta evaluación incluirá de modo específico los riesgos de reidentificación vinculados a la anonimización o seudonimización de los datos.

2.º Someter la investigación científica a las normas de calidad y, en su caso, a las directrices internacionales sobre buena práctica clínica.

3.º Adoptar, en su caso, medidas dirigidas a garantizar que los investigadores no acceden a datos de identificación de los interesados.

4.º Designar un representante legal establecido en la Unión Europea, conforme al artículo 74 del Reglamento (UE) 536/2014, si el promotor de un ensayo clínico no está establecido en la Unión Europea. Dicho representante legal podrá coincidir con el previsto en el artículo 27.1 del Reglamento (UE) 2016/679.

g) El uso de datos personales seudonimizados con fines de investigación en salud pública y, en particular, biomédica deberá ser sometido al informe previo del comité de ética de la investigación previsto en la normativa sectorial. En defecto de la existencia del mencionado Comité, la entidad responsable de la investigación requerirá informe previo del delegado de protección de datos o, en su defecto, de un experto con los conocimientos previos en el artículo 37.5 del Reglamento (UE) 2016/679.

## RD Estudios observacionales con medicamentos

[El Real Decreto 957/2020 del 3 de noviembre, por el que se regulan los estudiosobservacionales con medicamentos de uso humano](https://www.boe.es/eli/es/rd/2020/11/03/957):

Artículo 4. Requisitos previos al inicio de los estudios observacionales con medicamentos.

1. De acuerdo con lo previsto en el artículo 12 del Real Decreto 1090/2015, de 4 de diciembre, así como el capítulo III del presente real decreto, con carácter previo a su inicio, todos los estudios observacionales con medicamentos requerirán el dictamen favorable de un CEIm acreditado en España. Dicho dictamen será único, vinculante y reconocido en todo el territorio nacional. La obtención del dictamen del CEIm seguirá el procedimiento regulado en el capítulo III.

2. En el caso de estudios observacionales con medicamentos de seguimiento prospectivo, una vez obtenido el dictamen favorable del CEIm, las autoridades sanitarias competentes podrán establecer requisitos adicionales para que puedan iniciarse dichos estudios en los centros de su competencia. El establecimiento de estos requisitos deberá justificarse en criterios de factibilidad o pertinencia, pero no en aspectos del estudio ya evaluados por el CEIm correspondiente. No podrán establecerse requisitos adicionales en aquellos estudios cuyo promotor sea una Administración Pública o en los que quede acreditado que se trata de una investigación clínica sin ánimo comercial.

3. Los estudios que impliquen la obtención de información directamente del sujeto participante o del profesional sanitario que lo atiende en un centro, servicio o establecimiento sanitario, requerirán la conformidad previa del responsable del mismo al protocolo y al resto de documentación que ha obtenido el dictamen favorable del CEIm. La conformidad se expresará mediante la firma de un contrato con el promotor. Dicho contrato no será necesario en aquellos casos en los que el promotor pertenezca al centro, servicio, o establecimiento sanitario donde se realiza el estudio, siendo suficiente con recabar la conformidad expresa del responsable del mismo.

Artículo 5. Consentimiento informado y protección de los datos personales de los sujetos participantes.

1. Los estudios observacionales con medicamentos que conlleven entrevistar al sujeto participante, requerirán su consentimiento informado. No obstante, siguiendo las disposiciones aplicables de la normativa vigente y los principios éticos para las investigaciones médicas en seres humanos, se podrá eximir de solicitar el consentimiento informado, siempre que el CEIm considere que la investigación observacional tiene un valor social importante, que su realización no sería factible o viable sin dicha dispensa, y que entraña riesgos mínimos para los participantes.

2. Para la solicitud del consentimiento informado se tendrán en cuenta, en su caso, las normas y principios éticos relativos a las previsiones en relación con la recogida, almacenamiento y posible uso futuro de las muestras biológicas de los sujetos.

3. Los promotores de los estudios que utilicen alguna fuente de información que incluya el tratamiento de datos personales deberán tener en cuenta lo previsto en el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE, y en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales y, en particular, lo siguiente:

a) El promotor deberá haber evaluado y mitigado, mediante las medidas apropiadas en cada caso, el impacto que la realización del estudio puede tener en la protección de datos personales.

b) El promotor y los investigadores del estudio deben garantizar la confidencialidad de los datos de los sujetos participantes.

c) Sin perjuicio de lo dispuesto en el apartado 1, será necesario el consentimiento del sujeto participante a no ser que sea de aplicación otra base legítima para el tratamiento de sus datos personales de entre las referidas en los artículos 6.1 y 9.2 del Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016. Además, el promotor y los investigadores deberán aplicar los criterios que rigen el tratamiento de datos en la investigación en salud de conformidad con la disposición adicional decimoséptima de la Ley Orgánica 3/2018, de 5 de diciembre.

d) Las condiciones de acceso a los datos personales deberán detallarse en el protocolo, incluyendo las condiciones de su transmisión internacional fuera del ámbito del Espacio Económico Europeo, si ello está previsto.

4. Cuando se trate de un estudio con datos anónimos o que hayan sido sometidos a un tratamiento de seudonimización, el protocolo recogerá el procedimiento seguido para conseguir dicha anonimización o seudonimización.

5. El acceso a los datos de los sujetos participantes se realizará, en todo caso, en las condiciones que establezca el responsable del tratamiento de dichos datos, de modo que se garantice el cumplimento de la normativa sobre protección de datos personales

## Ley de la Ciencia, la Tecnología y la Innovación

La [Ley 17/2022, de 5 de septiembre, por la que se modifica la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación](https://www.boe.es/eli/es/l/2022/09/05/17) define entre sus objetivos generales “Impulsar la ciencia abierta al servicio de la sociedad y promover iniciativas orientadas a facilitar el libre acceso a los datos, documentos y resultados generados por la investigación, desarrollar infraestructuras y plataformas abiertas, y fomentar la participación abierta de la sociedad civil en los procesos científicos.”

Este objetivo se desarrolla con más profundidad en el artículo 37 donde se dice:

“2. El personal de investigación del sector público o cuya actividad investigadora esté financiada mayoritariamente con fondos públicos y que opte por diseminar sus resultados de investigación en publicaciones científicas, deberá depositar una copia de la versión final aceptada para publicación y los datos asociados a las mismas en repositorios institucionales o temáticos de acceso abierto, de forma simultánea a la fecha de publicación.”

Tanto la [Universitat de les Illes Balears](https://dspace.uib.es/xmlui/) como el [IBSalut](https://docusalut.com/) cuentan con sus propios repositorios institucionales.

“5. […] Además del acceso abierto, y siempre con el objetivo de hacer la ciencia más abierta, accesible, eficiente, transparente y beneficiosa para la sociedad, los Ministerios de Ciencia e Innovación y de Universidades, cada uno en su ámbito de actuación, así como las Comunidades Autónomas en el marco de sus competencias, promoverán también otras iniciativas orientadas a facilitar el libre acceso y gestión de los datos generados por la investigación (datos abiertos), de acuerdo a los principios internacionales FAIR (sencillos de encontrar, accesibles, interoperables y reutilizables), a desarrollar infraestructuras y plataformas abiertas, a fomentar la publicación de los resultados científicos en acceso abierto, y la participación abierta de la sociedad civil en los procesos científicos, tal como se desarrolla en el artículo 38.”

# Bibliografía

1. <https://impact-data.bsc.es/>
2. <https://eosc.eu/>
3. <https://ehden.eu/>
4. <https://www.ohdsi.org/>
5. Suttorp, M. M et Al. Graphical presentation of confounding in directed acyclic graphs. 2015 Nephrology Dialysis Transplantation
6. Etminan M. Using Causal Diagrams to Improve the Design and Interpretation of Medical Research. Chest. 2020 DOI: [10.1016/j.chest.2020.03.011](https://doi.org/10.1016/j.chest.2020.03.011)
7. <https://www.darwin-eu.org/>
8. <https://www.phiri.eu/>
9. <https://www.boe.es/eli/es/l/2002/11/14/41/con>
10. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32016R0679>
11. <https://www.boe.es/eli/es/lo/2018/12/05/3/con>
12. <https://www.boe.es/eli/es/rd/2020/11/03/957>
13. <https://www.boe.es/eli/es/l/2022/09/05/17>
14. <https://www.go-fair.org/fair-principles/>
15. 10.1038/sdata.2016.18
16. <https://dspace.uib.es/xmlui/>
17. https://docusalut.com/
18. https://datamanagement.hms.harvard.edu/plan-design/data-management-plans
19. El Emam, K., & Dankar, F. K. (2008). Protecting privacy using k-anonymity. *Journal of the American Medical Informatics Association : JAMIA*, *15*(5), 627–637. https://doi.org/10.1197/jamia.M2716
20. Thierry Meurers, Raffael Bild, Kieu-Mi Do, Fabian Prasser, A scalable software solution for anonymizing high-dimensional biomedical data, GigaScience, Volume 10, Issue 10, October 2021, giab068, <https://doi.org/10.1093/gigascience/giab068>
21. https://www.rebiun.org/acceso-abierto/10-pasos-para-elaborar-un-plan-de-gestion-de-datos
22. https://www.rebiun.org/acceso-abierto/los-datos-de-investigacion-importan-gestiona-comparte-publica-y-reutiliza