

La Gestión de Datos de Investigación

Marisa Pérez Aliende

Universidad Autónoma de Madrid

mp.aliende@uam.es

Sevilla, 13 de Junio 2017





Horizonte 2020 y el por qué gestionar los datos de investigación

- La Unión Europea y H2020:
 - Piloto de Datos de Investigación en Abierto (ORD)
 - Datos FAIR
 - Piloto ORD y los datos personales
- Otros mandatos de investigación en abierto

La Unión Europea y Horizonte 2020

H2020 es el Programa Marco de Innovación e Investigación de la Unión Europea para el período 2014-2020. Es el instrumento financiero para llevar a cabo iniciativas y proyectos de investigación que contribuyen a:

- Abordar los principales retos sociales.
- Promover el liderazgo industrial en Europa.
- Reforzar la excelencia en su base científica.

Objetivos estratégicos del H2020:

- Crear una ciencia de excelencia para reforzar la posición de la UE en el panorama científico mundial. Se incrementa la dotación económica de las subvenciones y se fomenta la formación y la movilidad de los investigadores.
- Desarrollar tecnologías y sus aplicaciones para mejorar la competitividad europea, con una mayor inversión en tecnologías clave para la industria y en la financiación de las PYMES.
- Investigar en las grandes áreas que afectan a los ciudadanos europeos: salud, alimentación, energía, transporte, clima y seguridad.

La Unión Europea y Horizonte 2020

H2020 de un vistazo:

- 80 millones de euros de financiación durante siete años (2014-2020).
- Orientación hacia el mercado.
- Aúna investigación e innovación.
- Un único programa y un solo conjunto de normas de participación.
- Estructura simplificada que reduce la complejidad de las reglas, la burocracia, y los plazos de concesión de las subvenciones.
- Acceso más fácil para todos los participantes.
- Aspectos básicos: ciencia excelente, liderazgo industrial y retos sociales.
- Especial atención a la participación en las pequeñas y medianas empresas.

La estrategia europea 2020 para una economía sostenible e inclusiva subraya el papel central del conocimiento y la innovación para generar crecimiento.

La Unión Europea y Horizonte 2020

¿Por qué un mayor acceso a las publicaciones y a los datos de investigación en H2020?:

- Construir basándose en resultados de investigación previos
 - ↳ mejorando la calidad de los resultados
- Fomentar la colaboración y evitar la duplicidad de esfuerzos.
 - ↳ mayor eficiencia
- Acelerar la innovación.
 - ↳ mayor progreso en los mercados, lo que significa un mayor crecimiento
- Involucrar a los ciudadanos, a la sociedad.
 - ↳ mejora la transparencia del proceso científico

La Comisión Europea considera que no debiera de haber necesidad de pagar por la información financiada con fondos públicos, cada vez que se acceda o se usa.

En H2020 estos principios se convierten en requerimiento.

La Unión Europea y Horizonte 2020

Acuerdo de subvención

H2020 AGA – Annotated Model Grant Agreement, 21 de abril de 2017, v4.0

Destacamos el **art. 29** de difusión de resultados, acceso abierto y visibilidad de la financiación:

Artículo 29.1: Obligación de difundir los resultados, a no ser que vaya en contra de sus intereses legítimos, cada investigador debe difundir sus resultados tan pronto como sea posible.

Sin perjuicio de:

- Proteger los resultados (art. 27)
- Confidencialidad (art. 36)
- Seguridad (art. 37)
- Proteger los datos personales (art. 39)

Artículo 29.2: Acceso Abierto a las publicaciones científicas revisadas por pares:

- Tan pronto como sea posible, en un repositorio, de artículos, y también monografías, comunicaciones a congresos, etc... sobre todo versión publicada, pero también versión aceptada.
- Asegurarse el acceso abierto a la publicación vía repositorio, en los 6 meses siguientes (ciencias) y 12 meses (ciencias sociales y humanidades).
- Asegurar el acceso a los metadatos bibliográficos que identifican la publicación depositada.



La Unión Europea y Horizonte 2020

Acuerdo de subvención

H2020 AGA – Annotated Model Grant Agreement, 21 de abril de 2017, v4.0

Artículo 29.3: Acceso abierto a los datos de investigación.

La Comisión Europea pone en marcha una iniciativa piloto de datos de investigación denominada “**Open Research Data Pilot in H2020**” (ORD Pilot) para depositar los datos en abierto en un repositorio de datos, permitiendo validar los resultados de las publicaciones científicas. Los beneficiarios deben:

- Depositar los datos en un repositorio de datos permitiendo que terceros puedan: acceder, explotar, reproducir y difundir:
 - Los datos, incluidos los metadatos asociados permiten validar los resultados de investigación.
 - Otros datos, incluidos los metadatos asociados y definidos en el Plan de Gestión de Datos.
- Proporcionar información (vía repositorio) de las herramientas e instrumentos necesarios para validar esos resultados, y si es posible proporcionar esas herramientas.

H2020: Piloto de Datos de Investigación en Abierto

Tiene como objetivo mejorar y maximizar el acceso y la reutilización de los datos de la investigación generados por los proyectos.

Áreas participantes en el Programa 2014-2016
Además de las áreas indicadas, otras pueden participar de forma voluntaria.

- Infraestructuras de investigación, parte de e-infraestructuras.
- Liderazgo en tecnologías industriales y de capacitación.
- Tecnologías Futuras y Emergentes.
- Reto Social: “Energía eficiente, segura y limpia” - sección ciudades y comunidades inteligentes.
- Reto Social: “Acción por el clima, medio ambiente, eficiencia de los recursos y materias primas”, exceptuando las líneas relacionadas con las materias primas.
- Reto Social: “Europa en un mundo cambiante, sociedades inclusivas, innovadoras y reflexivas”.
- Ciencia con y para la Sociedad.

Programa 2017

A partir del 2017 el Proyecto Piloto se extenderá a todas las áreas temáticas del programa H2020. La norma por defecto será el acceso abierto a los datos de investigación.

H2020: Piloto de Datos de Investigación en Abierto

Opting out

En 2017, extendiendo el programa Piloto a todas las áreas de investigación, el acceso abierto se convierte en la opción por defecto. Sin embargo, no todos los datos de investigación pueden estar en abierto.

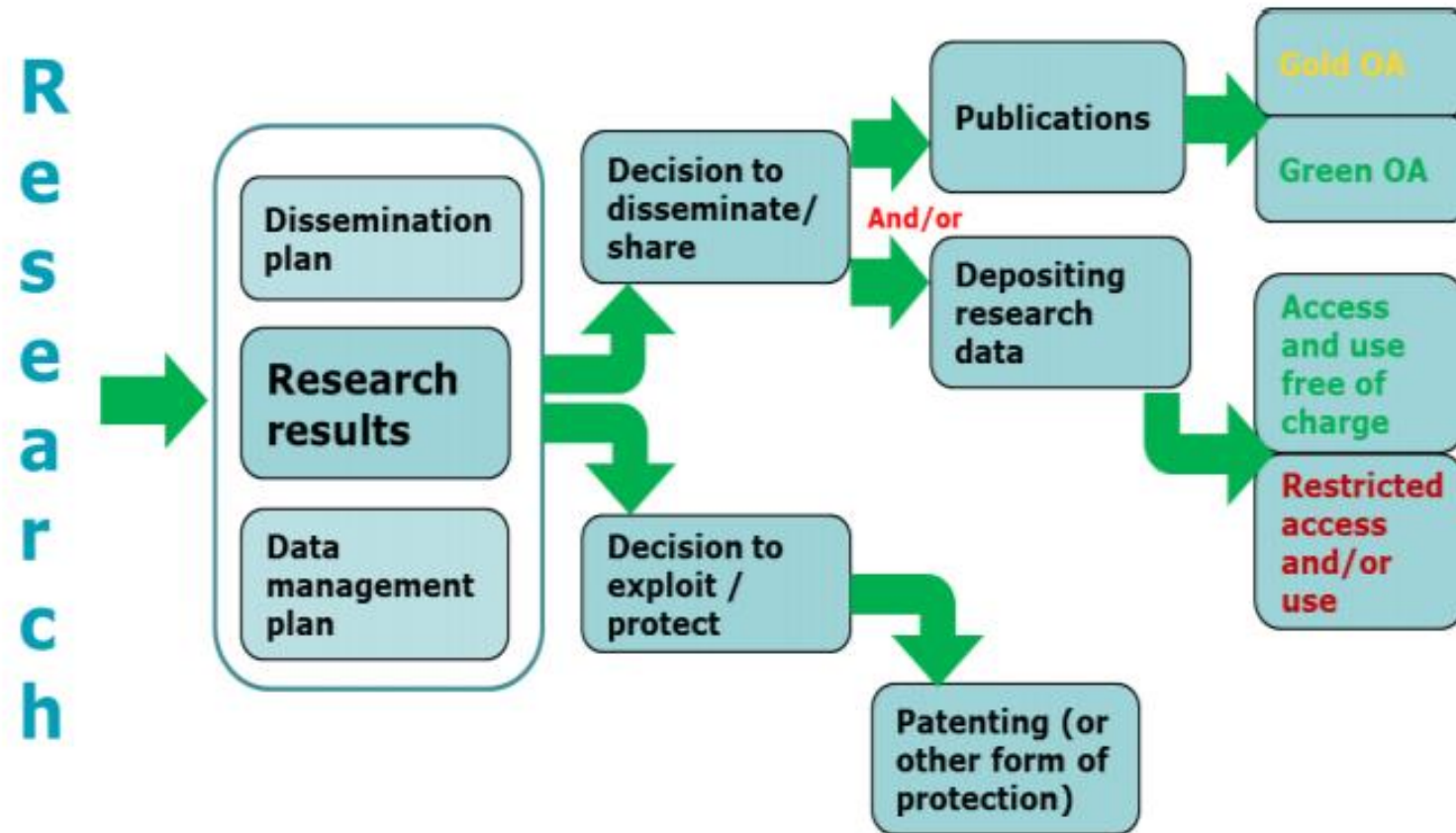
Los proyectos pueden optar en cualquier etapa a salir del Piloto, por alguno de los siguientes motivos:

- La participación sea incompatible con:
 - La obligación de proteger los resultados de investigación (previsión de explotarse comercial o industrialmente).
 - La confidencialidad, en relación con cuestiones de seguridad.
 - Las normas sobre la protección de datos personales.
 - El objetivo principal del proyecto
- El proyecto no genera o no contempla datos de investigación.
- Otras razones legítimas que lo justifiquen.

El lema de la Comisión en relación con los datos de investigación es:

“As open as possible, as closed as necessary”

H2020: Piloto de Datos de Investigación en Abierto



Graph: Open access to scientific publication and research data in the wider context of dissemination and exploitation

“H2020 Programme: Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in H2020”. 21 marzo 2017 v3.2

H2020: Datos FAIR

Los beneficiarios de H2020 deben hacer que los datos de investigación sean “FAIR” :

FINDABLE (Localizables):

- Asignarles un identificador único y persistente (DOI).
- Describirlos con metadatos enriquecidos.
- Registrarlos e indexarlos en recursos de búsqueda.

ACCESIBLE (Accesibles):

- Recuperarlos a través de protocolos normalizados y abiertos, implementables internacionalmente.
- Licencias, posibles embargo.
- Metadatos accesibles, incluso cuando los datos no estén disponibles.

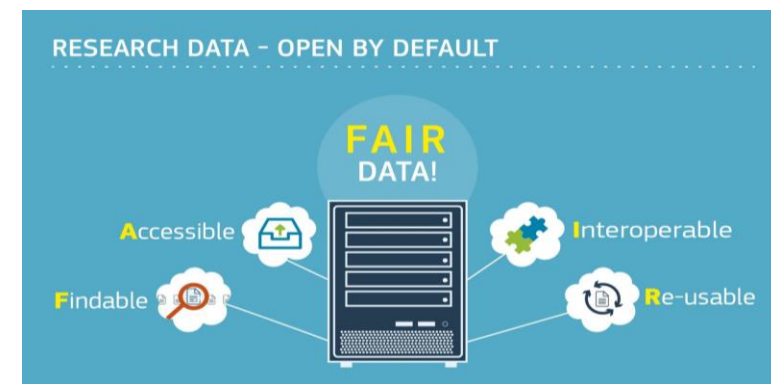
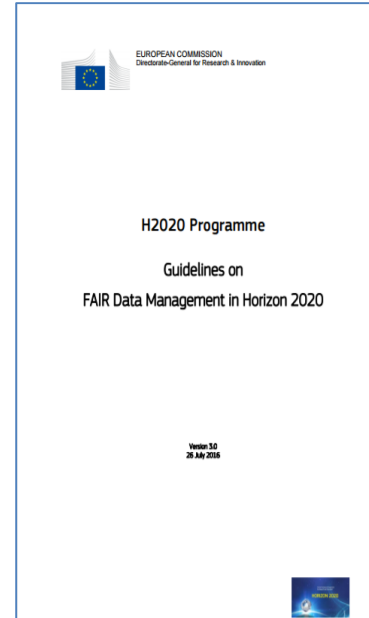
INTEROPERABLE (Interoperables):

- Utilicen un lenguaje formal y ampliamente aplicable.
- Vocabularios estándar.
- Incluir referencias a otros metadatos.

RE-USABLE (Reutilizables):

- Precisos y relevantes
- Licencias de uso claras
- Información sobre su procedencia

Principios FAIR (www.force11.org/group/fairgroup/fairprinciples)



(Gwen, F. et al. “Open Services & Tools: for Open Research Data in H2020”. IDCC workshop 2017)

Piloto de ORD y los datos personales

OpenAIRE proporciona al investigador soporte y servicios para el ORD Pilot e investiga sus aspectos legales, entre ellos la protección de datos.

En abril de 2017 se publica “Study on the EC ORD Pilot and Personal Data Rules”.

Dentro de la UE, el Charter of Fundamental Rights garantiza la protección de datos personales, junto con leyes secundarias como Directiva 95/46/EC que se remplazarán por GDPR (General Data Protection Regulation) con entrada en vigor en mayo 2018.

La protección de datos se aplica a cualquier información calificada como “datos personales”:

“any information relating to an identified or identifiable natural person (data subject); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as name, an identification number, location data, online identifier or to one or more factors specific to the physical, psychological, genetic, mental, economic, cultural or social identity of that person.”

Hay que tener en cuenta:

- Consentimiento. No legitima su uso.
- Anonimización. Es la mejor opción antes del depósito en el repositorio.

Mandatos de Datos de Investigación en Abierto

Existen sitios web en los que podemos acceder a las políticas y mandatos de publicar en abierto los datos de investigación de diferentes agencias de financiación. Entre ellos podemos hablar:

- Sherpa-Juliet

<http://www.sherpa.ac.uk/juliet/index.php>

- ROARMAP: Registry of Open Access Repository Mandates and Policies

http://roarmap.eprints.org/view/policymaker_type/funder.html

- Johns Hopkins University. Data Management Resources. Funder Data Related Mandates

<http://dms.data.jhu.edu/data-management-resources/plan-research/funders-data-sharing-requirement/funder-data-related-mandates-and-public-access-plans/>

- Digital Curation Centre (DCC)

<http://www.dcc.ac.uk/resources/policy-and-legal/funders-data-policies>

<http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>

La Gestión de Datos de Investigación

Marisa Pérez Aliende

Universidad Autónoma de Madrid

mp.aliende@uam.es

Sevilla, 13 de Junio 2017

