

Ciencia Abierta.

Un paisaje en construcción con un horizonte de posibilidades

Teresa Gomez-Diaz

CNRS - Laboratoire d'informatique Gaspard-Monge

Colaboración con Prof. Tomas Recio, Universidad Antonio de Nebrija

Esta obra está compartida bajo una licencia Creative Commons
Reconocimiento-No Comercial-Sin Obra Derivada 4.0 Internacional (CC BY-NC-ND 4.0)

https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es_ES

Universidad Antonio de Nebrija, Madrid, 28 de septiembre de 2023



Contexto (1/3)

Colaboración con Prof. Tomas Recio, Universidad Antonio de Nebrija

- 2019 TGD, T. Recio. On the evaluation of research software: the CDUR procedure, F1000Research, Research on Research, <https://doi.org/10.12688/f1000research.19994.2>
- 2020-1 TGD, T. Recio. Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs, POLIS N. 19, 2020, <https://uet.edu.al/polis/wp-content/uploads/2022/01/polis-19.pdf>, V3 du 28/02/2021, <https://zenodo.org/record/4577066>
- 2021 TGD, T. Recio. Open comments on the Task Force SIRS report: Scholarly Infrastructures for Research Software (EOSC Executive Board, EOSCArchitecture), RIO 7:e63872, <https://doi.org/10.3897/rio.7.e63872>
- 2021 TGD, T. Recio. The future of Open Science asks for a common understanding, Poster EGI Virtual Conference 2021, <https://indico.egi.eu/event/5464/contributions/15729/>
- 2022 TGD, T. Recio. Research Software and Research Data: dissemination, evaluation and reusability in the Open Science context, Poster IDCC22, <https://zenodo.org/record/6778872>
- 2022 TGD, T. Recio. Research Software vs Research Data I: Towards a Research Data definition in the Open Science context, F1000Research, Research on Research, <https://f1000research.com/articles/11-118/v2>
- 2022 TGD, T. Recio. Research Software vs Research Data II: Protocols for Research Data dissemination and evaluation in the Open Science context, F1000Research, Research on Research, <https://f1000research.com/articles/11-117/v2>
- 2023 TGD, T. Recio. How to achieve FAIRER research data by studying evaluation assessment protocols, Conference Open Science FAIR, Madrid, 25-27 septiembre 2023, <http://opensciencefair.eu/posters/how-to-achieve-fairer-research-data-by-studying-evaluation-assessment-protocols>

Code for Thought Podcasts, Peter Schmidt

- 2022 P. Schmidt. Podcast Code for Thought: Open Science and Research Software, <https://codeforthought.buzzsprout.com/1326658/10822132-open-science-and-research-software>
- 2023 P. Schmidt. Podcast Code for Thought: Open Data, Open Software - with Teresa Gomez-Diaz, <https://codeforthought.buzzsprout.com/1326658/13216530-en-open-data-open-software-with-teresa-gomez-diaz>
- 2023 TGD, T. Recio. Podcast Code for Thought: Research Software and Research Data in Open Science, Zenodo preprint, <https://zenodo.org/record/8159906>

Contexto (2/3)

Algunas presentaciones y conferencias en España

- 09/2020 Ciclo de Conferencias sobre la Ciencia Abierta, Facultad de Ciencias, U. Cantabria
<http://igm.univ-mlv.fr/~teresa/logicielsLIGM/conferenciasSeptiembre2020/>
- 12/04/2021 A propósito de la Ciencia Abierta: compartir y difundir la producción científica, U. Nebrija
<http://igm.univ-mlv.fr/~teresa/logicielsLIGM/conferenciaUNebrija2021>
- 09/2021 Ciclo de Conferencias sobre la Ciencia Abierta, Facultad de Ciencias, U. Cantabria
<http://igm.univ-mlv.fr/~teresa/logicielsLIGM/conferenciasSeptiembre2021/>
- 20/09/2022 Ciencia Abierta: sobre el acceso abierto a las producciones científicas,
http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/Internacional/20220920SantancerCienciaAbierta_TGD.pdf
- 28/09/2023 Ciencia Abierta. Un paisaje en construcción con un horizonte de posibilidades, U. Nebrija
<http://igm.univ-mlv.fr/~teresa/logicielsLIGM/conferenciaUNebrija2023>
- Workshop "Open science, a landscape under construction with a horizon of possibilities", CIEM, Castro Urdiales, 11-13 noviembre 2022. <http://igm.univ-mlv.fr/~teresa/2022WorkshopOpenScienceAtCIEM>
- Jornadas de Ciencia abierta, CRUE, Universidad de Vigo, Vigo, 1 y 2 de junio de 2023,
<https://open-science.es/>, <https://tv.uvigo.es/series/647dc972b5099d79f2579453>
- Open Science FAIR, Madrid, 25-27 de septiembre de 2023, <http://opensciencefair.eu/>

Contexto (3/3)

Toma de conciencia, en curso, en la comunidad científica

El sangrante precio de publicar ciencia,

Luis González MacDowell, The Conversation, 19 septiembre 2023.

<https://theconversation.com/el-sangrante-precio-de-publicar-ciencia-212393>

*... las universidades españolas pagan por dar acceso libre a sus publicaciones alrededor de **2.500 euros por artículo** ...*

Por ejemplo, en la Universidad Complutense de Madrid se publican alrededor de 11.800 artículos al año, que al precio del acuerdo piloto supondrían en total 30 millones de euros en gastos de publicación: alrededor de 20 veces más que el actual desembolso de su biblioteca, que se sitúa aproximadamente en 1,3 millones de euros por año.

350 millones de euros anuales le supondrían a España sus publicaciones

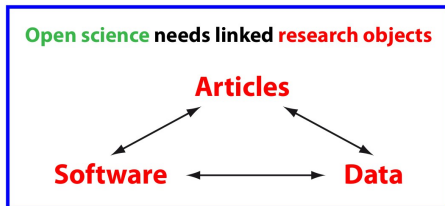
Escuchado en Open Science FAIR el 25 de septiembre:

Neil Coleman, Library Citizen Science Engagement Officer, University of Edinburgh

We have the infrastructure, the services, but researchers don't use the services as much as they should.

Motivación (1/2)

- Ciencia Abierta: compartir y difundir la producción científica
- Movimientos de Open Access, Open Data: artículos, datos
- Los artículos son la producción más visible y accesible
- Los investigadores producen muchos programas y datos para obtener los resultados científicos que se publican en los artículos...
- ... pero sin tener aún un procedimiento claro de publicación
- Consecuencia: la mayor parte está en *unknown status*
- Cambios recientes: mayor importancia de los datos y de los programas
- La investigación necesita la producción científica y los enlaces entre los diferentes productos



Motivación (2/2)

- ***Wanted:***

- ▶ ***acciones científicas:***

- estudiar, colaborar, validar, verificar, reproducir... obtener nueva ciencia

- ▶ ***acciones legales*** (derechos de autor, licencias):

- escribir, utilizar, copiar, citar, modificar, (re)distribuir, explorar, traducir...

- Imposible o difícil de acceder al trabajo de otros sin ***(open) access***

- No es suficiente solamente con el acceso

- ¿Pero que quiere decir ***open*** ? Es decir,

- ¿en qué ***condiciones*** se difunden y se comparten las obras ?

- ▶ verificar: las definiciones, las licencias, las políticas

- Además...

***La Ciencia Abierta es imposible sin los datos (research data)
y sin los programas de ordenador (research software)***

Plan

1 Motivación

2 Las publicaciones científicas: artículos

- En la era del papel, en la era digital
- El gran negocio y la ciencia que no existe
- El acceso libre: la declaración de Budapest (BOAI) y algunos avances recientes

3 Ciencia Abierta

- Definición propuesta
- Recomendación de la UNESCO sobre Ciencia Abierta
- Guía del programa “Horizon Europa”

4 Los programas y los datos

- Definiciones, movimientos
- Las cuestiones legales: derecho de autor, derecho *sui generis*
- Las licencias de los programas y las licencias Creative Commons
- Procedimiento de difusión de programas y datos

5 Research Software (RS) y Research Data (RD)

- Definición de Research Software (RS)
- Protocolo(s) CDUR para la evaluación de RS
- Research Data (RD): definición, CDUR
- How to achieve FAIRER Research Data

6 Conclusiones

Las publicaciones científicas en la era del papel

[JCG 2001, JCG 2017, JCG 2019, EAL 2008, EDLC 2019, EDLC 2020]

- Los investigadores envían los *preprints* a las revistas científicas (*journals*)
- Roles: *publishers* (publican), *scientific editors* (edición científica), *referees* (validación)
- Una vez pasada la revisión por pares (*(blind) peer review*), el artículo se publica
- Los *publishers* venden las revistas a las bibliotecas universitarias o los organismos de investigación, donde otros investigadores acceden a los artículos
- *Journal*: registro público de contribuciones originales al conocimiento
quién hace qué + efecto de marca
- *Journal*: colaboración con intereses económicos (*commercial publishers*) o científicos
- Los mejores artículos son los más citados
- Los mejores científicos prefieren publicar en las mejores revistas
- Evaluación de la investigación: publicación en “buenas” revistas (prestigio)
- Hay que comprar las “buenas” revistas a todo coste
- En general no hay apreciación ninguna de la producción de datos o de programas
- Por ej. *Numerical Recipes in C++* (1988-92), CD con programas, seguido de página web

En la era digital

Sin ordenadores y redes, el acceso al conocimiento dependía principalmente de las bibliotecas. Otra forma de publicación científica y de acceso al conocimiento comienza a organizarse.

Ver OA Timeline, Peter Suber, <http://oad.simmons.edu/oadwiki/Timeline>

- En los años 80-90, ciertos equipos científicos establecen sus propias revistas sin *publishers*, por ej. Séminaire Lotharingien de Combinatoire <https://www.mat.univie.ac.at/~slc/> o SURFACES <https://www.erudit.org/fr/revues/surfaces/>
- (1991) Aparecen depósitos de *preprints* como arXiv <https://arxiv.org/>
- Aparecen organizaciones como Scielo <https://scielo.org/index.php?lang=es>, Research Papers in Economics (RePEc) <http://repec.org/>, Scholarly Publishing and Academic Resources Coalition (SPARC) <http://www.arl.org/sparc/> ...
- ...

- La publicación científica constituye un mercado inelástico: la demanda sufre una nula o muy poca variación cuando ocurre un cambio en el precio del producto.
Léase: **se puede aumentar el precio tanto como se quiere**
- Los *big publishers* compran editoriales más pequeñas y crean grandes monopolios
- Aparecen consorcios de bibliotecas que comienzan a negociar acuerdos con los *publishers*, por ej. Canadian National Site Licensing Project <https://www.crkn-rcdr.ca/en/keywords/canadian-national-site-licensing-project> o Couperin en Francia <https://www.couperin.org/presentation/historique>

Estamos en un modelo de intercambio de información caro, ineficiente, obsoleto.

El gran negocio y la ciencia que no existe

Journal Impact Factor (JIF):

- En los años 70, Eugene Garfield en el Instituto de la Información Científica (ISI) lanza el *Science Citation Index (CSI)* que conduce al *Journal Impact Factor (JIF)*.
 - Actualmente el *Journal Citation Reports (JCR)*, empresa Clarivate, sigue +12.000 revistas, +80 países <https://clarivate.com/webofsciencengroup/>
 - Cálculo del JIF de una revista: según el n. de citas de los artículos de revistas en (JCR):
 $J =$ una revista científica en JCR (https://es.wikipedia.org/wiki/Factor_de_impacto)
 $na =$ número de artículos publicados en J en 2001/02
 $nac =$ número de artículos de J que ha sido citado en 2001/02 por revistas en JCR
 $JIF \text{ de } J \text{ en } 2003 = nac/na.$
 - JIF organiza la competición entre las revistas (se miran 3 decimales)
 - JIF ha producido **una evaluación de los científicos según las revistas donde publican**, pero el impacto científico de un artículo no es lo mismo que el JIF de un Journal
 - ¿criterios imparciales? (*unbiased criteria*)?: ¿Cómo se seleccionan las revistas, los países?
Segregación: pago, lengua, temas científicos
 - Las revistas pagan (caro) para ser incluidas en JCR
- Los países y los establecimientos pagan (caro) para obtener servicios/informaciones
 - ¿Se paga por la información? ¿Se paga por **la evaluación de la ciencia**?, pero con el dinero que tiene que financiar la ciencia...
 - [EAL 2008] (2003) Redalyc: **la ciencia que no se ve no existe** <https://www.redalyc.org/>
 - Hablemos de dinero (varios documentos consultados, distintas fechas): Elsevier controla 20% de los “core journals”. Las bibliotecas científicas de Canadá pagan 4.5 millones de US\$. 10 SPARC journals cuestan \$5.238 y los 10 comerciales equivalentes: \$40.677. Márgenes de beneficio: 20% a 30%. Publicar un artículo: (2013) 5.000 \$US en Cell Reports, (2020) 11.000 \$US en Nature. Márgenes de Elsevier: 37%, Hindawi: 50%. [CNRS 2020] La factura Elsevier en Francia: 35 millones € por año. La UE piensa que 70-80 mil millones € invertidos en investigación no llegan a su objetivo cuando la ciencia es sólo accesible a quien pueda pagar la suscripción a las revistas científicas.

El acceso abierto (o libre): la declaración de Budapest

La declaración de Budapest, más conocida como BOAI (2002):

<http://www.budapestopenaccessinitiative.org/>

Convergencia de una tradición antigua y una nueva tecnología:

las políticas de acceso abierto (*open access*, OA) son cada vez más importantes en la comunidad científica internacional.

Son ahora posibles gracias a Internet y sus consecuencias, están en parte inspiradas por los movimientos FOSS y motivadas por los problemas de reproducibilidad de los resultados publicados.

Por “acceso abierto” a esta literatura queremos decir su disponibilidad gratuita en Internet público, permitiendo a cualquier usuario leer, descargar, copiar, distribuir, imprimir, buscar o usarlos con cualquier propósito legal, sin ninguna barrera financiera, legal o técnica, fuera de las que son inseparables de las que implica acceder a Internet mismo. La única limitación en cuanto a reproducción y distribución y el único rol del copyright en este dominio, deberá ser dar a los autores el control sobre la integridad de sus trabajos y el derecho de ser adecuadamente reconocidos y citados.

¿La solución al problema?

Cambios recientes, en curso, llevará tiempo y atención:

- Cambios en las leyes

- Ley 17/2022, de 5 de septiembre, de la Ciencia, la Tecnología y la Innovación, artículo 2(c) + artículo 37, <https://www.boe.es/eli/es/l/2022/09/05/17/con>
- Ley Orgánica 2/2023, de 22 de marzo, del Sistema Universitario, artículo 12 <https://www.boe.es/eli/es/lo/2023/03/22/2/con>

- Cambios en las políticas científicas

- Estrategia Nacional de Ciencia Abierta (3 mayo 2023) <https://www.ciencia.gob.es/InfoGeneralPortal/documento/c30b29d7-abac-4b31-9156-809927b5ee49>

- Cambios en el sistema de publicaciones

- Open Research Europe (pilot), <https://open-research-europe.ec.europa.eu/>
- Free Journal Network, <https://freejournals.org/current-member-journals/>
- Open Journal Systems (OJS), PKP, <https://pkp.sfu.ca/software/ojs/>

- Infraestructuras, repositorios

- Recolecta, OpenAIRE, EOSC...

- Utilización de licencias

- Creative Commons, https://creativecommons.org/licenses/?lang=es_ES

- Cambios en la evaluación de la actividad científica

- (2012) San Francisco Declaration on Research Assessment (DORA), <https://sfdora.org/read/>
- (2019) [Foro Latinoamericano sobre Evaluación Científica \(FOLEC-CLACSO\)](#)
- (2021) [EC Scoping report](#) Towards a reform of the research assessment system
- (2022) [Coalition for Advancing Research Assessment \(COARA\)](#)

- ⚠ Toma de conciencia de la comunidad científica

Plan

1 Motivación

2 Las publicaciones científicas: artículos

3 **Ciencia Abierta**

- Definición propuesta
- Recomendación de la UNESCO sobre Ciencia Abierta
- Guía del programa “Horizon Europa”

4 Los programas y los datos

5 Research Software (RS) y Research Data (RD)

6 Conclusiones

Definición propuesta de Ciencia Abierta - artículo

[TGDR 2020-21] La Ciencia Abierta es:

[ESP] *el marco político y legal en el que las producciones científicas son compartidas y difundidas con el fin de hacerlas visibles, accesibles y reutilizables.*

[EN] the political and legal framework where research outputs are shared and disseminated in order to be rendered visible, accessible and reusable.

Versión	Título	Fecha	Publicación
V3	Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs	02/2021	https://zenodo.org/record/4577066 Incluye referencia Alma Swan 2012
V2	Towards an Open Science definition as a political and legal framework:...	12/2020	POLIS N. 19, pp. 1-25 - PDF
V1	A policy and legal Open Science ...	09/2020	https://zenodo.org/record/4075106

Objetivos: entender qué es la Ciencia Abierta, la motivación que nos lleva a proponer esta definición y ver cómo se estructura la información.

Definición propuesta de Ciencia Abierta - poster

The future of Open Science
asks for a common understanding

*Open Science is
the political and legal framework
where research outputs are shared
and disseminated in order to be rendered
visible, accessible and reusable.*

I Three selected pillars

- BOAI (2002)
- Free Software Foundation (1985)
- CODATA (1966)

II Towards a political and legal framework

III Enablers:

- Institutional policies
- Infrastructures
- Research evaluation

**The future of Open Science
asks for a common understanding**

Teresa Gomez-Diaz, CNRS/LIGM, Est de Paris
Tomas Recio, University Nebrija, Madrid

A definition is missing

Definition of Open Science: "an ongoing, probably
iterative adaptation and change" (Gomez-Diaz 2016)
"the use of digital technologies to make research
an essential 'open' element" (Tomas 2018)
"There is a lack of consensus about what Open Science is,
despite the fact that it has been defined in various
ways" (Gomez-Diaz 2016)
"Open Science is a term that has been used in various
ways" (Gomez-Diaz 2016)

Goal
To contribute to the adoption of a common, unified vision.

Definition proposal
Open Science is the political and legal framework where
research outputs are shared and disseminated in order
to be rendered visible, accessible and reusable
(Gomez-Diaz & Recio, 2020-21).

Three steps supporting this proposal
1 - Three selected pillars for a common understanding
II - Towards a political and legal framework
III - Enablers: three cornerstones to get to a working framework

I - Three selected pillars for a common understanding

Free Software Foundation, 1985
The Free Software Foundation (FSF) was created in 1985 to defend the freedom to use, study, share, and improve software.
It is an independent non-profit organization.

Budgetary Open Access Initiative, 2002
The Open Access Initiative (OAI) was created in 2002 to promote the availability of research results in the public domain, ensuring they were freely available, shared, and reusable, without any barriers.

Committee on Data for Science and Technology, 1966
CODATA is an international organization for the dissemination of data in science and technology.

II - Towards a political and legal framework

"There is a lack of consensus about what Open Science is, despite the fact that it has been defined in various ways" (Gomez-Diaz 2016)

"There is a lack of consensus about what Open Science is, despite the fact that it has been defined in various ways" (Gomez-Diaz 2016)

"There is a lack of consensus about what Open Science is, despite the fact that it has been defined in various ways" (Gomez-Diaz 2016)

III - Enablers: three cornerstones to get to a working framework

Institutional policies: the required evaluation of policies of Universities and NGOs
Infrastructures: the development of Open Science-oriented infrastructures and services.
Research evaluation: the transformation of evaluation policies and practices.

References

Gomez-Diaz, T. (2016). Open Science: a new paradigm in research. *Open Science*, 5(1), 1-10.
Gomez-Diaz, T. (2018). Open Science: a new paradigm in research. *Open Science*, 7(1), 1-10.
Gomez-Diaz, T. (2020). Open Science: a new paradigm in research. *Open Science*, 9(1), 1-10.
Gomez-Diaz, T. (2021). Open Science: a new paradigm in research. *Open Science*, 10(1), 1-10.
Gomez-Diaz, T. (2022). Open Science: a new paradigm in research. *Open Science*, 11(1), 1-10.

EGI Virtual Conference 2021
19-21 October 2021

EGI Virtual Conference 2021, Lisbon, 19-21 october 2021

<https://padlet.com/gwenfranck/EGI2021Posters> - <https://zenodo.org/record/6433533>

(2020-21) TGD, T. Recio: Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs, POLIS N. 19 2020, V3 du 28/02/2021: <https://zenodo.org/record/4577066>

Recomendación de la UNESCO sobre Ciencia Abierta

- A petición de los 193 países miembros, 40th session, 11/2019
- Consultas globales, regionales, temáticas, 12/2019-07/2020
- Toma de conciencia importante, consecuencia de la pandemia COVID-19
- Recomendación de la UNESCO sobre la Ciencia Abierta, texto adoptado el 23 de noviembre 2021 (41st session, 11/2021)

https://unesdoc.unesco.org/ark:/48223/pf0000379949_spa

A los efectos de la presente Recomendación, la **ciencia abierta** se define como **un constructo inclusivo que combina diversos movimientos y prácticas** con el fin de que **los conocimientos científicos multilingües estén abiertamente disponibles y sean accesibles para todos, así como reutilizables por todos**, se incrementen las colaboraciones científicas y el intercambio de información en beneficio de la ciencia y la sociedad, y se abran los procesos de creación, evaluación y comunicación de los conocimientos científicos a los agentes sociales más allá de la comunidad científica tradicional.

La ciencia abierta comprende **todas las disciplinas científicas** y todos los aspectos de las prácticas académicas, incluidas las ciencias básicas y aplicadas, las ciencias naturales y sociales y las humanidades, y se basa en los siguientes pilares clave: conocimiento científico abierto, infraestructuras de la ciencia abierta, comunicación científica, participación abierta de los agentes sociales y diálogo abierto con otros sistemas de conocimiento.

En curso: 5 OS working groups for Implementation, they have produced UNESCO OS toolkit:

<https://www.unesco.org/en/open-science/toolkit>

Guía del programa “Horizon Europa” (HEurope) (1/2)

Horizon Europe Programme Guide, Version 3.0, 1 abril 2023 (V1 del 17/6/21) [PDF](#)

Guía de presentación de proyectos: *contains detailed guidance on the structure, budget and political priorities of Horizon Europe. It also includes details on how to prepare proposals.*

Ciencia Abierta en sección *Dissemination and exploitation of research results* (p.35)
y sección *Open Science* (p.40):

- Selección de proyectos: excellence, impact, and quality and efficiency of implementation
- Ciencia Abierta: se evalúa en *excellence*
- Ciencia Abierta: se evalúa en *quality and efficiency of implementation*
- Open access to peer-reviewed publications is **mandatory**
- Open access to generated research data is **required** but...
'as open as possible, as closed as necessary'
- Data Management Plans are **mandatory**
- Open access to other research outputs (software...) is **recommended**, as soon as possible, providing guidance to potentially interested users
- **Mandatory** practices are incentivised through their evaluation at proposal stage
- Proposers will have to provide information in **how** they plan to comply with mandatory O.S. practices, and explain the adoption of recommended practices, **justify otherwise...**

Guía del programa “Horizon Europa” (2/2)

Open Science in Horizon Europe (p.40)

Open science is an approach based on open cooperative work and systematic **sharing of knowledge and tools** as early and widely as possible in the process. It has the potential to increase the quality and efficiency of research and accelerate the advancement of knowledge and innovation by sharing results, making them more reusable and improving their reproducibility.

Horizon Europe moves beyond open access to open science for which it features a comprehensive **policy** implemented from the proposal stage to project reporting. The Horizon Europe Regulation sets the **legal basis** for the open science obligations and incentives that apply to Horizon Europe beneficiaries.

In Horizon Europe, open science practices are considered in the **evaluation of proposals**, under ‘excellence’ and under the ‘quality and efficiency of implementation’.
There are mandatory open science practices...

Plan

1 Motivación

2 Las publicaciones científicas: artículos

3 Ciencia Abierta

4 Los programas y los datos

- Definiciones, movimientos
- Las cuestiones legales: derecho de autor, derecho *sui generis*
- Las licencias de los programas y las licencias Creative Commons
- Procedimiento de difusión de programas y datos

5 Research Software (RS) y Research Data (RD)

6 Conclusiones

Definición de programa libre - *free software*

Según la Free Software Foundation (FSF), 1985, fundada por R. Stallman, *Software libre* es el software que respeta la libertad de los usuarios y la comunidad. Las **cuatro libertades esenciales** que habría que respetar son:

<http://www.gnu.org/philosophy/free-sw.es.html>

- libertad de ejecutar el programa como se desee (todo uso, sin restricción),
- (*) libertad de **estudiar** su funcionamiento y de modificarlo,
- libertad de redistribuir copias,
- (*) libertad de distribuir copias de las versiones modificadas.

⇒ (*) **condición necesaria: acceso al código fuente**

Ejemplos: programas muy conocidos que corresponden a la definición en fechas anteriores: T_EX de D. Knuth (1978), la Berkeley Software Distribution (BSD) de la University of California (1977-1995).

La garantía de las 4 libertades es un texto legal: la **licencia**.

El origen de esta definición se sitúa en un contexto académico (MIT).

Más información en [JMGB 2011, JMGB 2021].

Definición de programa de código abierto - *open source software*

Según la Open Source Initiative (OSI), 1998, un programa es de código abierto (*open source software*) si su **licencia** respeta estas condiciones: <http://www.opensource.org/docs/osd>

1. Free Redistribution
The license shall not restrict any party from **selling** or giving away...
2. Source Code (⇒ **acceso al código fuente**)
3. Derived Works
4. Integrity of The Author's Source Code
5. No Discrimination Against Persons or Groups
6. No Discrimination Against Fields of Endeavor
7. Distribution of License
8. License Must Not Be Specific to a Product
9. License Must Not Restrict Other Software
10. License Must Be Technology-Neutral

El origen de la definición se sitúa en un contexto cercano a empresas.

Se necesita una **licencia** para verificar las 10 condiciones.

FOSS, FLOSS: free (libre) y/o open source software (las mismas licencias).

Más información en [JMGB 2011, JMGB 2021].

Datos abiertos (*open data*): CODATA, RDA, FAIR...

- El movimiento de Open Data: CODATA (Committee on Data for Sciences and Technology) constituido en 1966 por International Council of Scientific Unions (ICSU), (1971) http://www.codata.info/resources/newsletters/Newsletter_7.pdf:

CODATA est un Comité au niveau scientifique international le plus élevé [...] à cause de l'importance qui s'attache à l'évaluation des données [...] c'est un comité de coordination et sa principale tâche est de prendre des initiatives et de souligner l'importance des aspects communs à plusieurs domaines de la science et de la technologie, ce qui comprend les activités suivantes:

- a) l'évaluation des méthodes de contrôle de la qualité,*
- b) la définition des besoins des utilisateurs,*
- c) les standards divers...*

- Visión + reciente: Research Data Alliance (RDA) <https://www.rd-alliance.org/about-rda>, lanzada en 2013 por la Comisión europea, la Fundación Nacional de la Ciencia (NSF, USA) el Instituto Nacional de Standards y Tecnología (USA), y el Departamento de Innovación del Gobierno de Australia, con el objetivo de construir la infraestructura técnica y social que facilite compartir y reutilizar los datos.
- [MW 2016] *The FAIR Guiding Principles for scientific data*: Findable, Accessible, Interoperable, Reusable.

La base legal: los derechos de autor, el derecho *sui generis*

BOE número 97, 22 de abril de 1996

<https://www.boe.es/buscar/pdf/1996/BOE-A-1996-8930-consolidado.pdf>

Los **artículos y los programas de ordenador** son obras literarias, artísticas o científicas, y tienen asociados:

- derechos *morales*, que son irrenunciables e inalienables, están asociados a personas físicas, se pueden heredar... por ej. firma (asociar autor a la obra)
- derechos *patrimoniales*, que se pueden ceder o vender. Las acciones de reproducción, distribución, comunicación pública, transformación no podrán ser realizadas sin autorización.

Los **programas** tienen un tratamiento especial, por ej. los derechos de explotación corresponderán al empleador, salvo pacto en contrario.

En las **bases de datos**, el derecho de autor puede aplicarse a los datos, a la estructura de la base, a los útiles que la hacen funcionar (por ej. programas).

El derecho *sui generis* protege principalmente al productor de la base de datos y puede prohibir la extracción y/o reutilización de la totalidad o de una parte sustancial del contenido de ésta.

Las licencias

¿Quién puede utilizar un programa de ordenador?

Directiva CE (13) y (15)

La utilización, reproducción, traducción, adaptación o transformación no autorizadas de la forma del código en el que se suministra la copia del programa de ordenador constituyen una **infracción** de los derechos exclusivos del autor.

Directiva 2009/24/CE del Parlamento europeo y del consejo, 23 de abril de 2009 sobre la protección jurídica de programas de ordenador, <https://www.boe.es/doue/2009/111/L00016-00022.pdf>

También: A Practical Guide to Using Free Software in the Public Sector, T. Aimé (2010)
<https://zenodo.org/record/7096100>

- Las licencias completan el marco jurídico de la ley: si no hay un derecho explícitamente dado, utilizar un programa constituye una infracción.
- Las licencias son **contratos** y protegen a los autores, los usuarios y los posibles colaboradores.
- Las licencias conceden derechos (y libertades) y pueden contener cláusulas de reciprocidad o imponer obligaciones que se tienen que respetar.
- Ejemplo en un libro: una traducción es una obra derivada.
J. K Rowling es la autora de la obra original y de todas las traducciones.

Tipos de licencias libres/open source para los programas

● Copyleft fuerte

- Licencia inicial se impone a todo.
- Obligación de reciprocidad, evita cerrar programas libres.

● Copyleft débil

- Licencia inicial permanece.
- Añadidos pueden tener otra licencia.

● Sin Copyleft

- La licencia inicial no se impone.
- Los derivados pueden tener otra lic.

- Copyleft fuerte : GNU GPL, CeCILL v2
- Copyleft débil : MPL, GNU LGPL, CeCILL-C
- Sin Copyleft : Apache, BSD, MIT, CeCILL-B

Software \ Licence		Non Copyleft	Weak Copyleft	Strong Copyleft
Original	Original			
	Modified			

GPLv2: *“You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.”*

Ver: T. Aimé, A Practical Guide to Using Free Software in the Public Sector (2010)

<https://zenodo.org/record/7096100>

J.-L. Archimbaud, T. Gomez-Diaz, PLUME (2009)

FAQ : Licence & copyright pour les développements de logiciels libres de laboratoires de recherche

<https://zenodo.org/record/7063146>

La licencia “European Union Public License” (EURL)

- Instrumento de derecho europeo, v1.2 del 9 mayo 2017, conforme a las leyes de la Union Europea y de los Estados Miembros
<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017D0863>
<https://joinup.ec.europa.eu/collection/eupl/eupl-text-eupl-12>
- Multilingüe, 23 idiomas, con el mismo valor jurídico, competencia del país productor del programa, que da la licencia
- F/OSS, con obligación de reciprocidad, evitar la apropiación exclusiva (GAFAM) y garantizar un acceso compartido a las posibles mejoras
- Con mecanismos de compatibilidad, lista de licencias compatibles incluida
- Utilizada en los programas de la Comisión Europea, Decisión C(2021)8759 del 8/12/2021 [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021D1209\(01\)](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021D1209(01))

Presentación de EURL (PowerPoint presentation)

<https://joinup.ec.europa.eu/collection/eupl/present-eupl-powerpoint-presentation>

Video : Patrice-Emmanuel Schmitz (Consultant Joinup, EC), Atelier BlueHats, 25/02/2022

<https://communs.numerique.gouv.fr/ateliers/eupl/>

Joinup Licensing Assistant - Compatibility Checker,

<https://joinup.ec.europa.eu/collection/eupl/solution/joinup-licensing-assistant/jla-compatibility-checker>

Las licencias Creative Commons (CC)

Otorgan permisos legales a una obra e indican si puede ser **copiada, distribuida, editada, remezclada y desarrollada**, dentro de los límites de la ley de propiedad intelectual.

La versión V4.0 está adaptada al derecho *sui generis* de las bases de datos.

No se utilizan con los programas de ordenador.

https://creativecommons.org/licenses/?lang=es_ES, https://creativecommons.org/choose/?lang=es_ES

-  CC-BY: libre sin copyleft
-  CC-BY-SA: libre, copyleft
-  CC-BY-ND
-  CC-BY-NC
-  CC-BY-NC-SA
-  CC-BY-NC-ND

-  Public domain mark: si no hay propietario conocido

<https://creativecommons.org/share-your-work/public-domain/pdm/>

-  CC-0: renunciar a los derechos de propiedad intelectual

<https://creativecommons.org/share-your-work/public-domain/cc0>

-  CC-0 $\neq$ 

Procedimiento de difusión de programas y de datos

Se adapta a la situación en curso, válido para datos.

- Elegir título, evitar las marcas o nombres utilizados.
- (*) Establecer la lista de autores (% participación) y afiliaciones.
- (*) Establecer la lista de las funcionalidades más importantes.
- (*) Establecer la lista de los programas o datos incluidos, sus licencias.
- **Elegir una licencia**, con acuerdo de autores y titulares de los derechos de autor.
Si se puede: acuerdo firmado. Atención a compatibilidad y herencia de las licencias.
- Elegir un sitio web, forge o depósito para la difusión, indicar licencias y condiciones de uso, copia, modificación etc., con la forma de citar la obra.
Utilizar **identificadores persistentes (PID)** si posible.
- (*) Archivar regularmente ficheros tar.gz para identificar nuevas funcionalidades.
- Informar a los laboratorios e instituciones (si no se hizo en el paso licencia).
- Crear e indicar claramente una dirección de contacto.
- **Difundir** el software o los datos.
- Informar a la comunidad, considerar la publicación de Software o Data Papers.

(*) Revisar en cada nueva versión.

Ver: T. Gomez-Diaz (2015),

Software libre, software de código abierto, licencias. Donde se propone un procedimiento de distribución de software y datos de investigación. <https://hal.archives-ouvertes.fr/hal-01261064>

Plan

- 1 Motivación
- 2 Las publicaciones científicas: artículos
- 3 Ciencia Abierta
- 4 Los programas y los datos
- 5 **Research Software (RS) y Research Data (RD)**
 - Definición de Research Software (RS)
 - Protocolo(s) CDUR para la evaluación de RS
 - Research Data (RD): definición, CDUR
 - How to achieve FAIRER Research Data
- 6 Conclusiones

Definición de Research Software (RS)

Research software (2019) *is a well identified set of code that has been written by a (again, well identified) research team.*

It is software that has been built and used to produce a result published or disseminated in some article or scientific contribution.

Each research software encloses a set (of files) that contains the source code and the compiled code. It can also include other elements as the documentation, specifications, use cases, a test suite, examples of input data and corresponding output data, and even preparatory material.

- Esta definición se apoya sobre el concepto legal de *software* y el derecho de autor.
- Otros conceptos:
autores y contribuidores, publicación de RS, referenciar y citar RS.
- Se estudian los contextos de evaluación de la investigación.
- Se proponen Protocolo(s) CDUR para la evaluación de RS.

(2019) TGD, T. Recio. On the evaluation of research software: the CDUR procedure

Protocolo(s) CDUR para la evaluación de RS

El objetivo es ayudar a: investigadores, comités de evaluación, responsables de las decisiones.

(C) Citation mide si el RS está correctamente identificado:
referencia correcta, y también metadata, citas correctas a otros RS...

perspectiva legal: autores, afiliación, % de participación

(D) Dissemination prácticas de difusión, según las normas de la evaluación
(2014) TGD: Free software, Open source software, licenses...

perspectiva política: Ciencia Abierta, **perspectiva legal:** licencias

(U) Use aspectos “software” **del RS:** resultados correctos, fácil de utilizar,
software: doc., test, instal., ..., leer el código, lanzar ejemplos...

perspectiva reproducibilidad: validación de los resultados científicos

(R) Research aspectos “investigación” **del RS:** calidad del trabajo científico,
algoritmos y estructuras de datos utilizados/implementados,
publicaciones, colaboraciones...

perspectiva investigación: impacto

Aplicación flexible: los comités de decisión y de evaluación **establecen** su propio
protocolo **CDUR** adaptado al contexto y a los objetivos de la evaluación en curso.

Definición de Research Data (RD)

Proponemos formulaciones similares para las definiciones de RS y RD.

Research Data (RD) *is a well identified set of data that has been produced (collected, processed, analyzed, shared & disseminated) by a research team.*

The data has been collected, processed and analyzed to produce a result published or disseminated in some article or scientific contribution.

Each research data encloses a set (of files) that contains the dataset maybe organized as a database, and it can also include other elements as the documentation, specifications, use cases, and any other useful material as provenance information, instrument information...

It can include the research software that has been developed to manipulate the dataset (from short scripts to research software of larger size) or give the references to the software that is necessary to manipulate the data (RS or other).

- Esta definición no se apoya sobre el concepto legal de *data*.
- Se adapta el procedimiento de difusión de RS a RD.
- Se proponen respuestas a los Borgman's conundrum challenges [CLB 2012].
- Se adaptan los Protocolo(s) CDUR para la evaluación de RS a RD.

(2022) TGD, T. Recio. Research Software vs Research Data I: Towards a Research Data definition...

(2022) TGD, T. Recio. Research Software vs Research Data II: Protocols for Research Data dissemination...

Protocolo(s) CDUR para la evaluación de RD

Ayudar a: investigadores, comités de evaluación, responsables de las decisiones.

(C) Citation mide si el RD está correctamente identificado:
referencia correcta, y también metadata, citas correctas a otros RD...
perspectiva legal: autores (si hay derechos de autor), afiliación, % de participación

(D) Dissemination prácticas de difusión, según las normas de la evaluación
(2014) TGD: Free software, Open source software, licenses...
perspectiva política: Ciencia Abierta, perspectiva legal: ★ otros temas legales, licencias

★ **(U) Use** aspectos “data” **del RD**: calidad, documentación, ejemplos de
utilización, facilitar la reutilización, *best data practices*...
perspectiva reproducibilidad: validación de los resultados científicos

(R) Research aspectos “investigación” **del RD**: calidad del trabajo científico,
algorithms y estructuras de datos utilizados/implementados,
publicaciones, colaboraciones...
perspectiva investigación: impacto

Aplicación flexible: los comités de decisión y de evaluación **establecen** su propio
protocolo **CDUR** adaptado al contexto y a los objetivos de la evaluación en curso.

How to achieve FAIRER Research Data...

Tabla de comparación entre los The FAIR Guiding Principles for scientific data [MW et al] y los Protocolos CDUR para los datos de investigación.

CDUR [2,7]	The FAIR Guiding Principles [1]
(C) Citation: The RD is well identified, involving issues concerning: - citation form or reference - metadata (including PIDs)	To be Findable: F1. (meta)data are assigned a globally unique and persistent identifier F2. data are described with rich metadata (defined by R1 below) F3. metadata clearly and explicitly include the identifier of the data it describes To be Interoperable: I3. (meta)data include qualified references to other (meta)data To be Reusable: R1.2. (meta)data are associated with detailed provenance
(D) Dissemination: The RD is well disseminated, involving issues concerning: - list of included components - RD licence - RD deposit	To be Findable: F4. (meta)data are registered or indexed in a searchable resource To be Accessible: A1. (meta)data are retrievable by their identifier using a standardized communications protocol A1.1 the protocol is open, free, and universally implementable A1.2 the protocol allows for an authentication and authorization procedure, where necessary A2. metadata are accessible, even when the data are no longer available To be Reusable: R1.1. (meta)data are released with a clear and accessible data usage license
(U) Use: The RD facilitates its reuse, involving issues like: - documentation, tutorials, examples... - reproducibility and replicability issues	To be Interoperable: I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation I2. (meta)data use vocabularies that follow FAIR principles To be Reusable: R1. meta(data) are richly described with a plurality of accurate and relevant attributes R1.3. (meta)data meet domain-relevant community standards
(R) Research: Measures the impact of the RD related scientific work	Not applicable

Table 1. Relationships between the FAIR principles and the CDUR RD evaluation protocols.

(2022) TGD, T. Recio. Research Software vs Research Data II: Protocols for Research Data dissemination...

(2023) TGD, T. Recio. How to achieve FAIRER research data by studying eval..., Open Science FAIR (25/09/23)

Conclusiones (1/2)

- 2015: imposible de parar,
- 2017: se va a fondo
- 2020-23: la comunidad científica tiene que estar atenta

- El paisaje de la Ciencia Abierta sigue en construcción
- Se trabaja mucho para los artículos, para las publicaciones
- También se avanza en los programas, en los datos
- Atención con los cambios en el sistema de publicación

- ▶ [JCG 2019] EC Expert Group report: Future of Scholarly Publishing... (2019)

*The best way to make the scholarly communication and publishing system evolve in a way satisfying the research-centred perspective is to **maximise cooperation and collaboration** among the actors willing to act in such a direction.*

- ▶ CODATA coordinated report for a UNESCO Recommendation: Open Science for a Global Transformation (07/2020) <https://zenodo.org/record/3935461>

*Scholarly publishing needs fundamentally reenvisioning such that it is unambiguously **put to the service of science as a global public good**.*

- ▶ CLACSO, Semana Internacional del acceso abierto (10/2020) <https://www.clacso.org/semana-internacional-del-acceso-abierto/>

Equity and inclusion: community-owned infrastructures for open science.

- El sistema de publicaciones está evolucionando, pero muy lentamente
- Los editores no tienen interés en cambiar el sistema

Conclusiones (2/2)

- El acceso abierto se construye en un contexto legal, que implica la utilización de licencias y la evolución de las leyes (EC, países).
- No se entienden correctamente los derechos de autor
- La comunidad científica paga muy caro las consecuencias
- La sociedad en general paga muy caro las consecuencias
- Todavía hay una disparidad enorme entre ciertos discursos y
 - ▶ la política científica de nuestras instituciones
 - ▶ la evaluación de la investigación: factores de impacto, datos y programas...
 - ▶ la actividad cotidiana de investigadores y laboratorios
- En algún momento habrá que lanzar los cambios: evaluación...
- COARA: ¿cómo evaluar la producción de datos y/o programas, ie. RS & RD?
- Todo cambio tiene consecuencias, positivas y negativas
- Hay que evaluar las consecuencias a largo plazo
- Hay que preparar los cambios para las generaciones futuras
- **El dinero de la ciencia tiene que financiar la ciencia**
Desafíos: clima, pandemias...
- Además... ***La Ciencia Abierta es imposible sin los datos (research data) y sin los programas de ordenador (research software)***

Referencias - Sitios en Internet, presentaciones...

- Amelica: Conocimiento Abierto sin fines de lucro <http://amelica.org/>
- CLACSO: Consejo Latinoamericano de Ciencias Sociales <https://www.clacso.org/>
- Dialnet: hemeroteca de artículos científicos hispanos <https://dialnet.unirioja.es/>
- DORA: San Francisco, evaluación investigación <https://sfpora.org/read/read-the-declaration-espanol/>
- Predatory journals <https://predatoryjournals.com/about/>
- Recolecta: Recolector de Ciencia Abierta <https://www.recolecta.fecyt.es/>
- Redalyc: revistas con publicación sin fines de lucro <https://www.redalyc.org/>
- Scielo: Scientific Electronic Library Online <https://scielo.org/>
- Zenodo: repositorio de Ciencia Abierta <https://zenodo.org/>
- Herramientas y estándares para la ciencia abierta, Isabel Bernal (2022),
https://digital.csic.es/bitstream/10261/271944/1/Openscience_standards_tools_2022_Bernal.pdf
- Estrategias de publicación y difusión de la investigación en Ciencias de la Educación, Emilio Delgado López-Cózar,
https://www.researchgate.net/profile/Emilio-Delgado-Lopez-Cozar/publication/339325720_Estrategias_de_publicacion_y_difusion_de_la_investigacion_en_Ciencias_de_la_Educacion/
- Hacer una tesis doctoral: Informarse, leer, indagar, escribir... publicar...
https://www.researchgate.net/publication/313218460_Hacer_una_tesis_doctoral_Informarse_leer_indagar_escribirpublicar

Referencias (1/2)

- JCG 2001 Jean Claude Guéron (2001), In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the control of scientific publishing, <https://arll.secure.nonprofitsoapbox.com/storage/documents/publications/in-oldenburgs-long-shadow.pdf>
- OECD 2007 OECD (2007), OECD Principles and Guidelines for Access to Research Data from Public Funding, <https://www.oecd-ilibrary.org/content/publication/9789264034020-en-fr>
- EAL et al, 2008 Emilio Aguado López et al. (2008), Redalyc: una alternativa a las asimetrías en la distribución del conocimiento, <https://www.redalyc.org/pdf/145/14511370002.pdf>
- JMGB, 2011 JM González-Barahona (2011), El concepto de software libre, <https://revistes.uab.cat/tradumatica/article/download/10/pdf>
- CLB 2012 Christine L. Borgman. The conundrum of sharing research data. Journal of the American Society for Information Science and Technology, vol. 63, no 6, p. 1059-1078.
- ND et al, 2013 N. Dietrich et al. (2013), Safe to Be Open: Study on the Protection of Research Data and Recommendations for Access and Usage, <https://library.oapen.org/handle/20.500.12657/32411>
- MW et al, 2016 M. Wilkinson et al. (2016), The FAIR Guiding Principles for scientific data management and stewardship, <https://doi.org/10.1038/sdata.2016.18>
- JCG 2017 Jean Claude Guéron (2017), Open Access: Toward the Internet of the Mind (BOA15), <https://www.budapestopenaccessinitiative.org/open-access-toward-the-internet-of-the-mind>
- LAEA 2018 Lluís Anglada, Ernest Abadal (2018), ¿Qué es la ciencia abierta?, <https://dialnet.unirioja.es/servlet/articulo?codigo=6842308>
- JCG et al, 2019 Jean Claude Guéron et al. (2019), Future of scholarly publishing and scholarly communication, https://eosc-portal.eu/sites/default/files/KI0518070ENN.en_.pdf

Referencias (2/2)

- EDLC 2019** Emilio Delgado López-Cózar, Alberto Martín-Martín (2019), El Factor de Impacto de las revistas científicas sigue siendo ese número que devora la ciencia española: ¿hasta cuándo?
<https://doi.org/10.3145/thinkepi.2019.e13e09>
- EDLC 2020** Emilio Delgado López-Cózar (2020), Guía de buenas prácticas en la publicación científica,
<https://dialnet.unirioja.es/descarga/articulo/7610015.pdf>
- EDLC** Emilio Delgado López-Cózar <https://scholar.google.fr/citations?hl=fr&user=kyTH0h0AAAAJ>
- ECCTI 2021-27** Estrategia Española de Ciencia, Tecnología e Innovación 2021-2027,
<https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/EECTI-2021-2027.pdf>
- TM 2020** Tambiama Madiaga (2020), Digital sovereignty for Europe,
[https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI\(2020\)651992_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf)
- CNRS 2020** CNRS (2020), La Science Ouverte : une révolution nécessaire. 1re journée pour la science ouverte au CNRS, https://www.science-ouverte.cnrs.fr/wp-content/uploads/2021/01/Recueil_Journee-ScienceOuverte_Cnrs_sept2020_VF.pdf
- TGDTR 2020/21** Teresa Gomez-Diaz, Tomas Recio (POLIS 2020, Version 3 en 2021), Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs,
http://uet.edu.al/polis/images/Teresa_Gomez-Diaz.pdf, <https://zenodo.org/record/4577066>
- JB et al 2021** Jeroen Bosman et al. (2021), The OA Diamond Journals Study, Part 1: Findings,
<https://zenodo.org/record/4558704>
- AB et al 2021** Arianna Becerril et al. (2021), The OA Diamond Journals Study, Part 2: Recommendations,
<https://zenodo.org/record/4562790>
- JMGB, 2021** JM González-Barahona (2021), A Brief History of Free, Open Source Software and Its Communities,
<https://dirkriehle.com/wp-content/uploads/2021/03/09353517.pdf>