

Paving the way for adopting FAIR Data in a university

Ljubljana, 8th November 2019

Ignasi Labastida i Juan
Office for the Dissemination of Knowledge
CRAI Universitat de Barcelona



The ultimate goal: FAIR Data

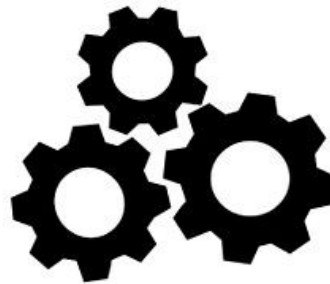
F
Findable



A
Accessible



I
Interoperable



R
Reusable

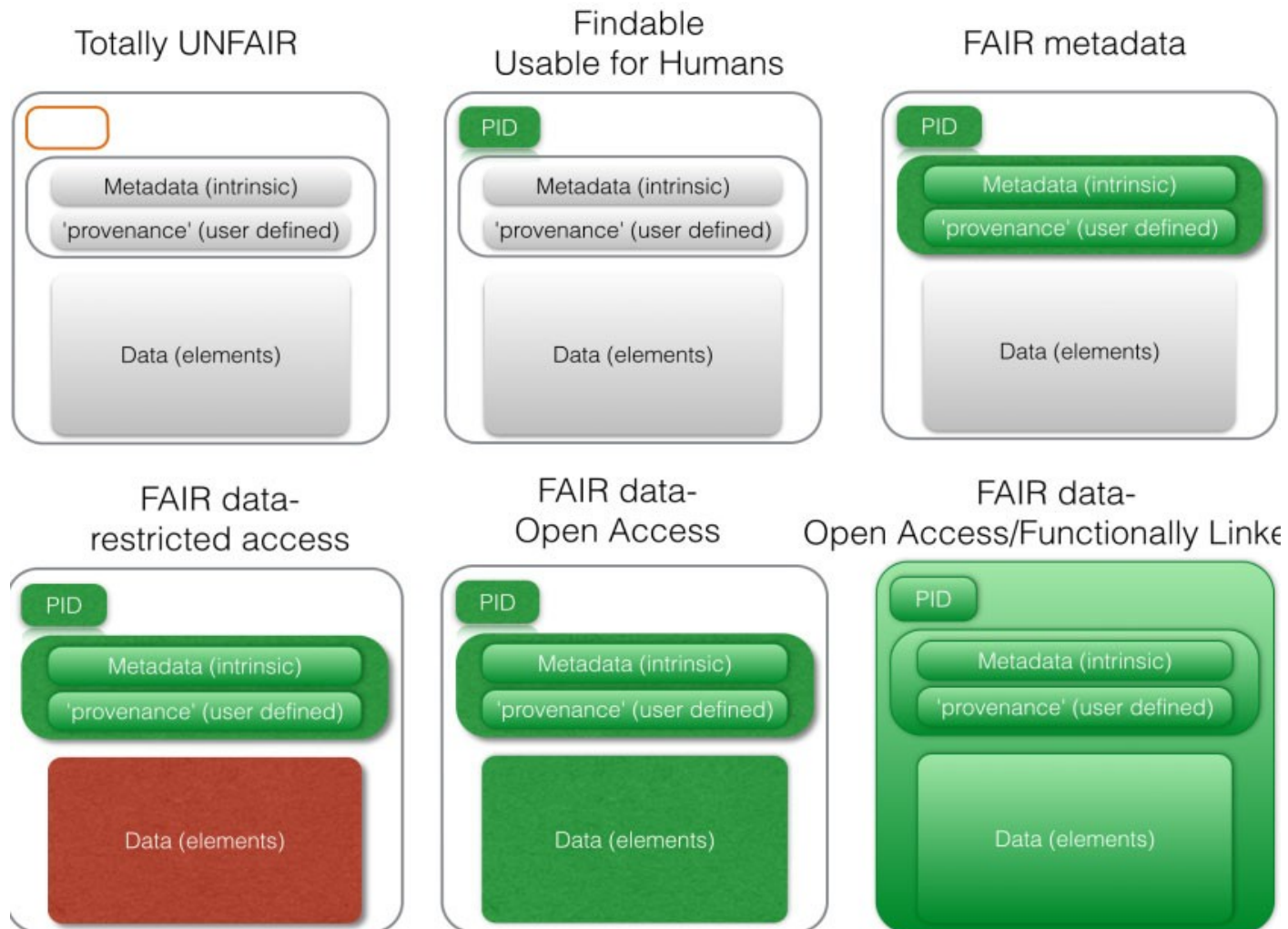


CC BY-SA SangyaPundir

https://commons.wikimedia.org/wiki/File:FAIR_data_principles.jpg

FAIR Data

Data as increasingly FAIR Digital Objects



<https://www.force11.org/fairprinciples>

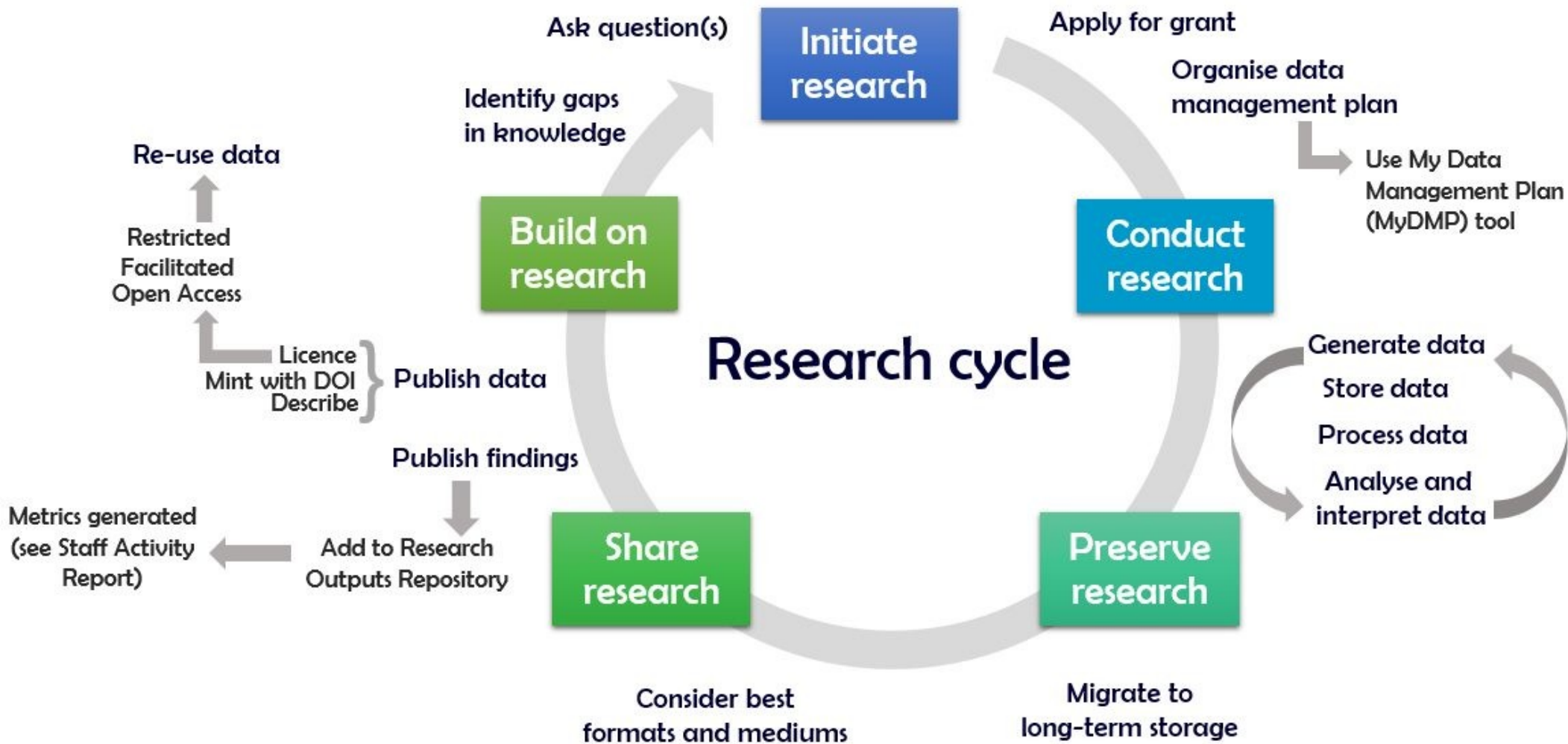
Overcome resistance

Problem 1: Data Excuse Bingo

My data contains personal/sensitive information	My data is too complicated	People may misinterpret my data	My data is not very interesting
Commercial funder doesn't want to share it	We might want to use it in another paper	People will contact me to ask about stuff	Data Protection/ National Security
It's too big	People will see that my data is bad	I want to patent my discovery	It's not a priority and I'm busy
I don't know how	I'm not sure I own the data	Someone might steal/ plagiarise it	My funder doesn't require it

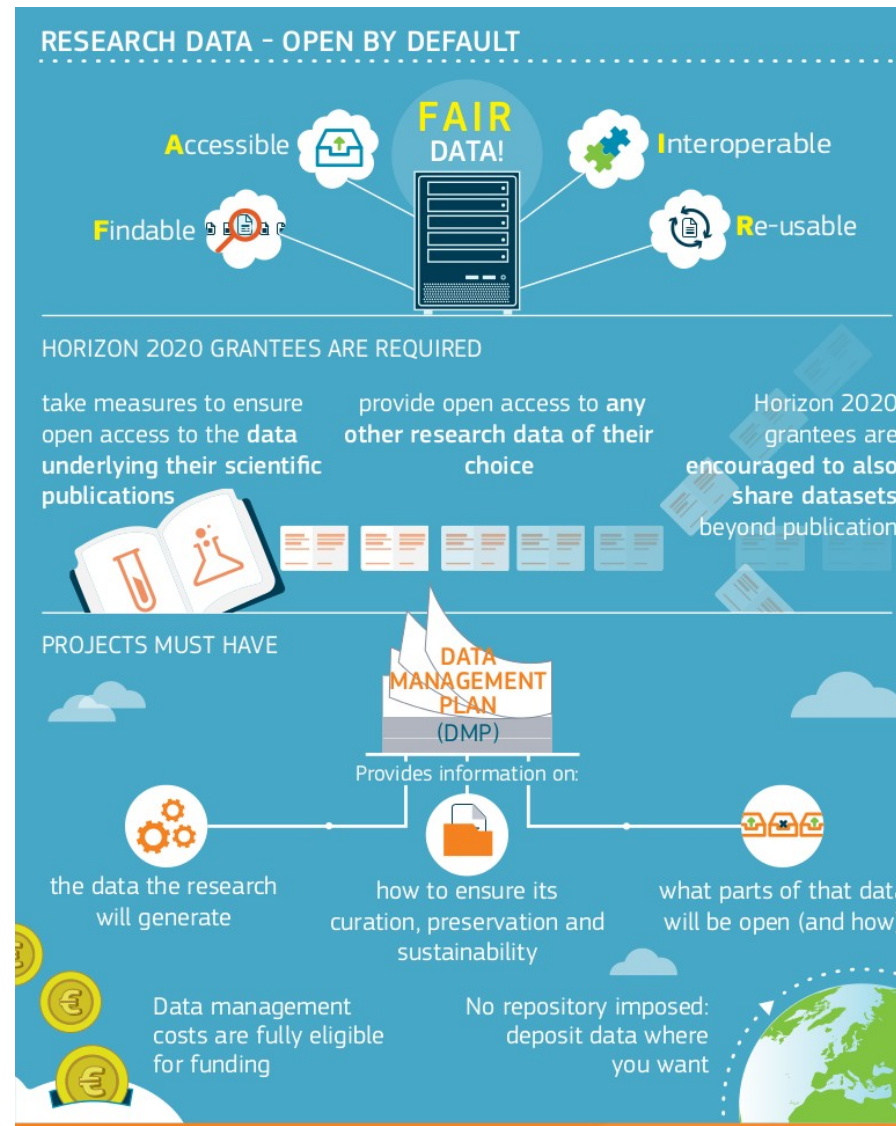
Data Excuse Bingo created by @jenny_molloy

Acting along the research cycle



<http://guides.library.unisa.edu.au/ResearchDataManagement>

Driven by funders' policies



We started by supporting DMP

Service aimed at helping researchers creating DMP

Choice of examples based on real DMP

Explanation on how to fill a DMP

Face to face meetings with researchers

Adaptation of the DMP Online: <http://dmp.csuc.cat>

Making infographies for DMP



10 passos per elaborar un Pla de Gestió de Dades

Un **Pla de Gestió de Dades** (PGD) o Data Management Plan (DMP) és un **document formal, que es presenta a l'inici de la investigació, en el qual es**

descriu què es farà amb les dades al llarg i després de finalitzar la investigació i que es pot modificar si es produeixen canvis en la mateixa.

Per què?

És una **bona pràctica**, és un **element clau de l'Open Science** i és **obligatori** als nous projectes H2020.

Eines gratuïtes per elaborar un PGD



PGDOnline
(Consorcio Madroño)
<http://dmp.consorciomadrono.es/>



DMPOnline (Digital Curation
Centre, UK)
<https://dmponline.dcc.ac.uk/>



01

Reviseu els **requeriments** de l'ens financer (H2020).



02

Identifiqueu les dades: tipologia, procedència, volum, formats i fitxers.



03

Definiu com s'organitzaran i gestionaran les dades: nom dels fitxers, control de versions, programari necessari...



04

Expliqueu com es documentaran les dades: identifiqueu la informació a processar, consulteu si hi ha estàndards o esquemes de metadades, identifiqueu eines per a gestionar-los.



05

Descriviu els processos que asseguruen una **bona qualitat de les dades**.



06

Prepareu una **estratègia d'emmagatzemament** (durant el procés) i de preservació de dades (repositori).



07

Definiu les polítiques de dades del projecte: qüestions sobre propietat intel·lectual i com es tractaran les dades sensibles i personals.



08

Descriviu com es difondran les dades: on, quines, quan es difondran. Si publicareu les dades en un repositori, com a informació suplementària de l'article o com un "data paper".



09

Assigneu rols i responsabilitats a les persones i organitzacions participants en el projecte.



10

Prepareu un **pressupost realista:** la gestió de dades costa temps i diners en termes de programari, maquinari, serveis i personal



Universidades Españolas Red de Bibliotecas REBIUN

Adaptat de Michener, WK. Ten simple rules for creating a good data management plan. PLOS Comp Biol. 2015; 11(10)

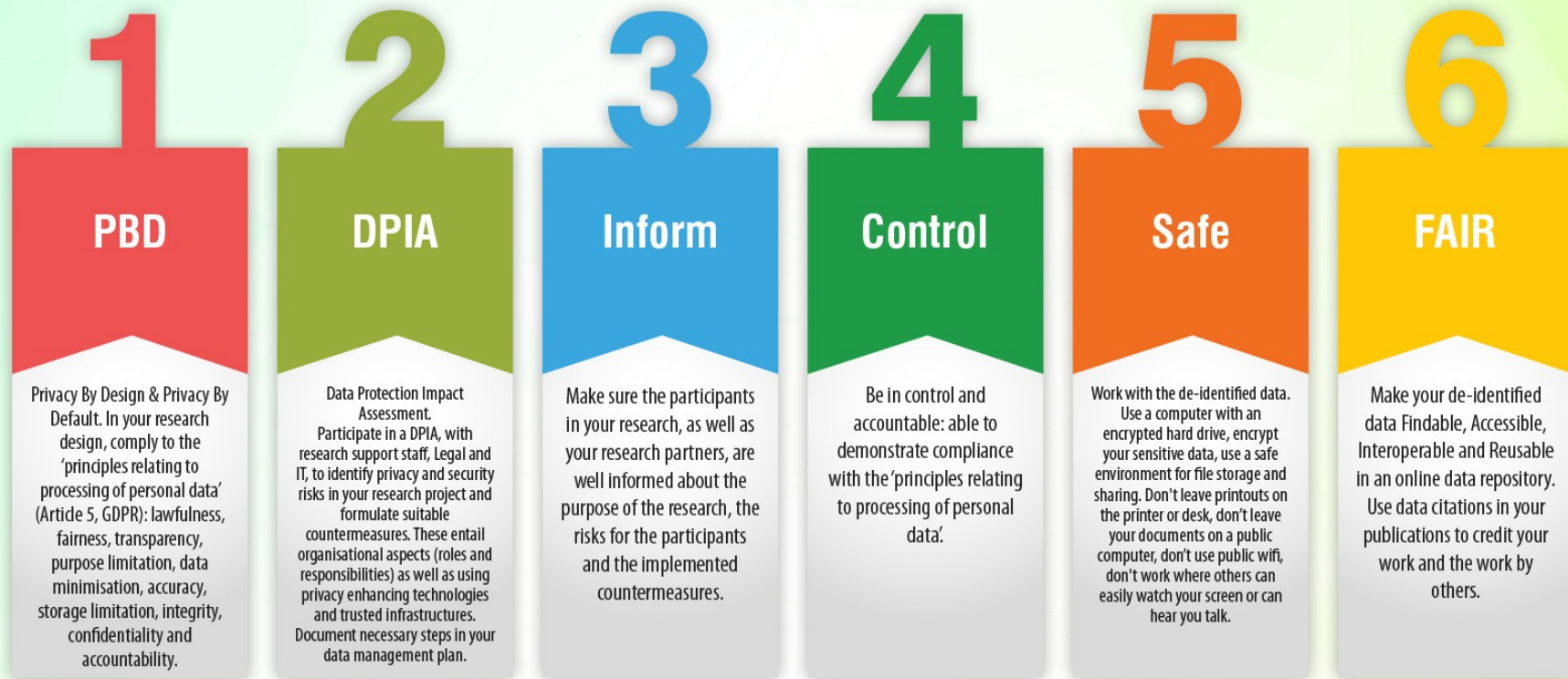


El personal de la teva Biblioteca pot assessorar



Reusing materials to alert researchers on personal data

HOW TO TREAT PERSONAL DATA IN RESEARCH. RESPONSIBLE USE OF PERSONAL DATA BEFORE, DURING AND AFTER RESEARCH.



For More Information visit
www.gdprcoalition.ie

Twitter
[@GDPR_Coalition](https://twitter.com/GDPR_Coalition)

Linkedin
[gdpr Coalition](https://www.linkedin.com/company/gdpr-coalition)

Brought to you by:



Reusing materials to alert researchers on personal data

HOW TO TREAT PERSONAL DATA IN RESEARCH?

Responsible use of personal data before, during and after research.

PRIVACY BY DESIGN AND BY DEFAULT



In your research design, address these six security and privacy goals, as identified by: www.datenschutzzentrum.de

Participate in a *data protection impact assessment* to identify risks and formulate countermeasures.

Communicate the security and privacy measures for your research in your *data management plan* and with all participants and data subjects.

BEFORE RESEARCH

Make sure your data subjects are well informed about the purpose of the research and their risks, before they sign the *informed consent form*.

Only generate and use data that are relevant for the purpose of your research: *data minimisation*

Use a computer with an encrypted hard drive, encrypt your sensitive data, use *SURFdrive* for safe and secure file storage and sharing.

DURING RESEARCH

Anonymise and / or pseudonymise the data and work with the de-identified data.

Work safe: don't leave printouts on the printer or desk, don't use public wifi, don't work where others can easily watch your screen or can hear you talk.

During research feel free to consult us in case of practical issues or just to reflect on aspects.

AFTER RESEARCH

Make your de-identified data *Findable and Accessible* in the DANSeasy national data repository. If possible, additionally provide metadata for making your data *Interoperable and Reusable*. #FAIR data

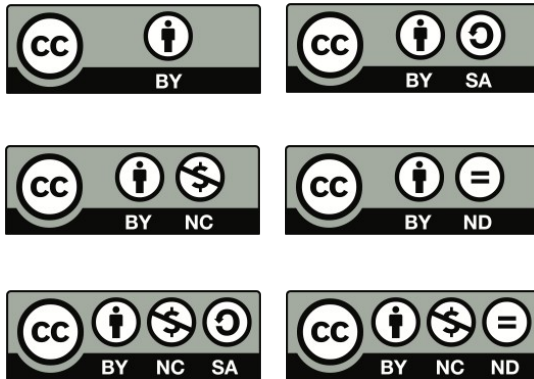
Share your experiences with us, contact us for support before, during and after research:

DATASUPPORT@UBIB.EUR.NL
+31 10 4088006

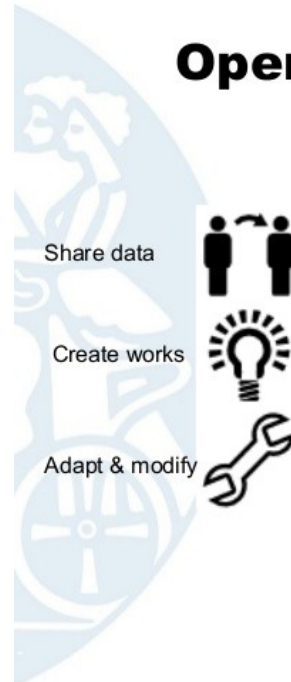


SEE ALSO:
[HTTPS://WWW.EUR.NL/RESEARCHMATTERS/RDM/RDM_SERVICES/](https://www.eur.nl/researchmatters/RDM/RDM_SERVICES/)

Advising on licensing options



Open Data Commons



ODC Open Database License (ODbL)

<http://opendatacommons.org/licenses/odbl/>



Attribution, Share alike & keep open

ODC Attribution

<http://opendatacommons.org/licenses/by/>



Attribution







ODC Public Domain Dedication

<http://opendatacommons.org/licenses/pddl/summary/>

None

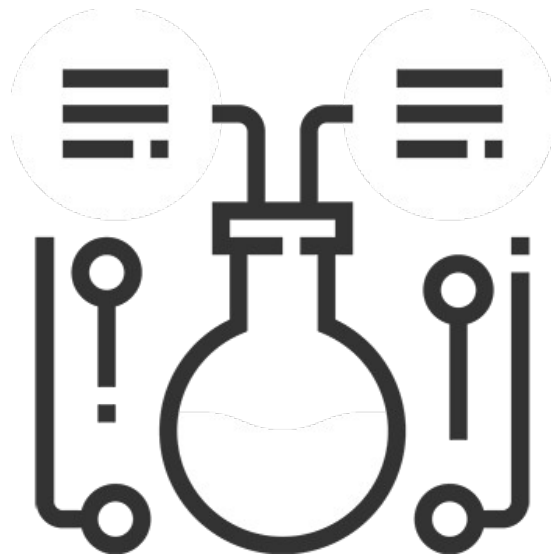
Advising on where to share data



	 CERN	 EUDAT	 Digital Science (Macmillan)	 NCSU	 IOSS (Harvard University)	 Elsevier
1. Àmbit temàtic i geogràfic	Multidisciplinar Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional	Amb èmfasi en ciència i medicina Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional
2. Tipologia de dades (programari, imatges, raw data, etc.)	Tot tipus de formats i també documents publicats: publicacions, literatura grisa, imatges i audiovisuals, text pla, etc.	Tot tipus de formats	Tot tipus de formats : xifres, conjunts de dades, imatges i audiovisuals, literatura grisa, conjunts d'arxius, etc.	Tot tipus de formats (amb exclusió de software) i amb preferència per formats de codi obert	Tot tipus de formats amb funcions específiques per alguns (Stata, SPSS, R, Excel, CSV, etc.)	Tot tipus de formats : text sense format, imatges i audiovisuals, PDF, HTML, etc.
3. Mida aproximada dels fitxers a incloure	Fins a 50 GB Per a mides superiors es pot contactar amb ells	Fins a 2 GB. Per a mides superiors, contactar amb ells	Fins a 5 GB	20 GB per un conjunt de dades (cost addicional per sobre de 20 GB)	Fins a 2 GB. Per fitxers més grans contactar amb ells	10 GB
4. Dades obertes, embargades, restringides o tancades	Obertes Tancades Restringides Embargades (Polícies)	Obertes Restringides Embargades. La descripció del conjunt de dades és pública.	Obertes Tancades (per defecte fins que no es publiquen) Restringides Embargades	Obertes Embargades (per defecte fins el moment de la publicació)	Obertes Restringides La descripció (títol, etc.) d'un conjunt de dades és pública però els arxius poden ser públics o restringits. Sense informació sobre embargaments	Obertes Tancades (per defecte fins que no es publiquen) Restringides Embargades
5. Llicències amb les que es volen difondre les dades	Permet varietat de llicències encara que recomanen llicències obertes. La descripció de les dades està sota CC0	Permet varietat de llicències per a dades i per a software	- CC0 per a dades (<i>metadata & datasets</i>) - CC-BY per a no dades (audiovisuals, <i>papers</i> , etc.) - MIT per a codi i software	CC0 (FAQ)	Per defecte CC0 però es pot canviar	Permet varietat de llicències: CC-BY 4.0, CC-BY-NC, CC0 1.0, MIT, Apache 2.0, BSD 3 clause, BSD 2 clause, gpl3
6. Identificador permanent	DOI	Handle	DOI	DOI	DOI	DOI
7. Condicions per dipositar o retirar contingut	Es permet retirar les dades justificant-ne el perquè. El DOI i la URL quedaran retinguts.	Per retirar contingut contactar amb: EUDAT	Sense informació	Cessió de dades irrevocable	Sense informació	Es permet retirar les dades abans de publicar-les. Una vegada publicades cal contactar amb Mendeley.

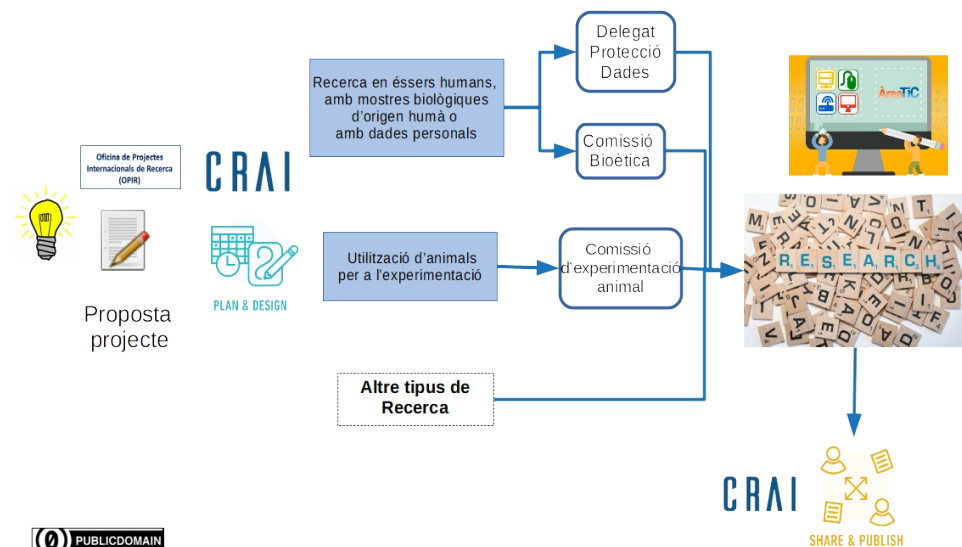
<https://ddd.uab.cat/record/150829>

Organising workshops with researchers



<http://www.ub.edu/dadesreerca/#>

Com gestionar les dades des de la idea a la publicació
Qui us pot ajudar i per on cal passar



PUBLICDOMAIN

Defining the requirements for a FAIR Data Repository

FAIR x FAIR

Feasible, Affordable and Implementable
Requirements for a FAIR research data
repository

<http://hdl.handle.net/2072/357521>

Defining the requirements for a FAIR Data Repository

Document based on analysis from other countries, technical reports and 25 interviews with experts

Output: 25 minimal requirements, 4 best practices and 3 recommendations

Defining the requirements for a FAIR Data Repository

To immediately create a repository in Catalonia for publication of research data according to FAIR requirements

Follow the 25 minimal requirements, and use available open-source software.

Publicise the infrastructure, provide a support service, and organise training sessions at all levels

And finally a Research Data Policy

Based on the LEARN model

An initial framework and a later implementation

Approved by the University Council in July, now working in its implementation

First action: asking PhD Students to make a DMP

Second action: choosing an infrastructure

Third action: Training

Next steps

New workshop with researchers on storage and preservation

Developing a DMP layout for PhD students

Adopting an institutional infrastructure

Learning from best practices: Course on Data Stewards

Training, training and training

Questions ?

ilabastida@ub.edu

@ignasi

Oficina de Difusió del Coneixement

CRAI, Universitat de Barcelona

<http://www.ub.edu/odc>