

Soporte a la gestión de datos de investigación: aproximación cooperativa de las bibliotecas de las universidades de Catalunya

Mireia Alcalá (CBUC - CSUC)

Con la ayuda del grupo de trabajo de soporte a la investigación del CSUC:

Ignasi Labastida (UB), Carme Besson (UAB), Rosa Fàbregas (UAB), Anna Rovira (UPC), Anna Casaldàliga (UPF),
Marina Losada (UPF), Brigit Nonó (UdG), Eva Estupinyà (UdL), Xosé Rius (URV), Mercè Montanyà (UVic-UCC),
Rosa Padrós (UOC), Laia Ros (URL), Gisela Coromines (URL), Lidón París (UJI), Lluís Anglada (CSUC),
Sandra Reoyo (CSUC) i Ricard de la Vega (CSUC)

Esquema

- 1. Las bibliotecas universitarias y la investigación**
- 2. Grupo de Trabajo de Soporte a la Investigación del CSUC**
- 3. Acciones realizadas en la gestión de los datos de investigación**
 - a) Soporte a los Planes de Gestión de Datos
 - b) Recomendación de repositorios
 - c) Política de gestión de datos de investigación
- 4. Trabajos futuros**

1. Las bibliotecas universitarias y la investigación (I)

- Nuevas necesidades de los investigadores
 - Cambios organizativos en las bibliotecas y las áreas de investigación para acercarse a estas nuevas necesidades
- En los últimos 10-15 años, las bibliotecas de las universidades de Catalunya proporcionan nuevos servicios:
 - bibliotecarios temáticos
 - herramientas para la gestión de referencias (Mendeley)
 - soporte a la publicación (repositorios) y asesoramiento sobre el cumplimiento de mandatos de acceso abierto
 - validación de datos de publicaciones de los CRIS (y relacionarlos con los repositorios institucionales)
 - implementación de identificadores (ORCID)

1. Las bibliotecas universitarias y la investigación (II)

- Nuevas oportunidades:
 - Ayudar en la aplicación de mandatos de OA
 - Portal de la recerca de les universitats de Catalunya
 - **Gestión de datos de investigación**
 - Elaboración de un catálogo de infraestructuras para la investigación
 - Adquisición de una herramienta para ayudar en la toma de decisiones estratégicas sobre la investigación
 - ...



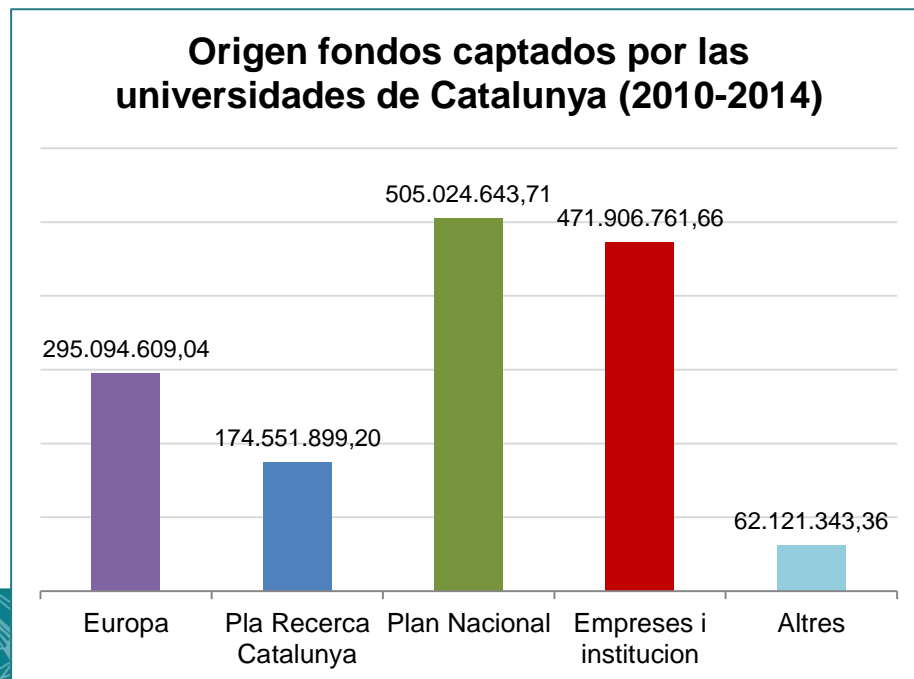
2. Grupo de Trabajo de Soporte a la Investigación del CSUC (I)

- En 2014, **ampliación de las líneas estratégicas** de actuación del Área de Bibliotecas, Información y Documentación del CSUC con una dedicada al soporte a la investigación.
- A principios de 2015, el Grupo de Trabajo de Soporte a la Investigación del CSUC se centró en la **gestión de datos de investigación**.
- El grupo está formado por representantes de todas las universidades miembros del consorcio y también de las áreas BID y TIC del CSUC.



2. Grupo de Trabajo de Soporte a la Investigación del CSUC (II)

- El objetivo inicial respecto a la Gestión de Datos de Investigación era:
 - “Establecer un marco de referencia que permitiese a las universidades y a los centros de investigación definir una política de gestión de datos generados por las actividades de investigación.”
- Trabajos centrados en el programa Horizonte 2020 debido a:
 - Open Research Data Pilot
 - Captación significativa de fondos europeos



3. Acciones realizadas en la gestión de datos de investigación

Gestió de les dades de recerca

Benvolguts investigadors,

En els darrers anys, les dades recollides, generades o utilitzades al llarg de les activitats de la recerca han rebut l'atenció de la comunitat científica i dels òrgans gestors de la recerca. La difusió d'aquestes dades és un dels elements constituents del que s'anomena ciència oberta. Per aquesta raó, a Europa, dins del marc del programa Horizon 2020, s'ha engegat una acció pilot, adreçada específicament a algunes àrees del programa, on es demana tenir un pla de gestió de dades per als projectes atorgats.

Mentre que el valor de les dades de recerca augmenta, es constata la dificultat de la seva gestió. Arran d'aquesta constatació, les institucions de recerca més actives han començat a detectar quines necessitats tenen els investigadors i quins serveis els poden ser útils per tal de facilitar-los aquesta gestió. És per això que les biblioteques i les oficines de recerca de les universitats catalanes, que treballen coordinadament en el si de l'àrea de Biblioteques, Informació i Recerca, han iniciat un projecte pilot que, a partir de la informació provinent dels projectes del programa Horizon 2020, analitzi les necessitats dels investigadors i estableixi mètodes i processos estandaritzats per a la gestió de les dades de recerca. Us presentem aquesta enquesta que podreu completar en el termini de 15 dies.

9% completat

Prospección:
encuesta a los
investigadores



Cursos de
formación

CSUC Consorci de
Serveis Universitaris
de Catalunya

Plans de Gestió de Dades

(Versió 1, Juliol 2016)

(Doc. 16/31) (BESRGT/Support Recerca/PlansdeGestioDeDades_versio1/Publica_julio16.pdf, 13.07.16)

Plan de Gestión
de Datos
de Investigación

CSUC Consorci de
Serveis Universitaris
de Catalunya

Recomanacions per seleccionar un repositori per al dipòsit de dades de recerca

(Versió 1, Juliol 2016)

Recerca/PlansdeGestioDeDades.docx, 27.07.16)

Recomendaciones
de repositorios
temáticos y
multidisciplinares

de Catalunya

Proposta per establir una política d'accés obert a les dades de recerca a les Universitats de Catalunya

(Doc. 16/30) (BESR/Gestió Dades Recerca/1603/Polítiques rdm-v8.docx, 11.07.16)

Acuerdo marco
para aprobar
una política
de gestión de datos

Recomanacions científiques i a les dades de

condicionals pels quals es comunicaven els resultats de la recerca a l'Accés Obert (OA). Per accés obert s'entén la llibertat de l'accés obert s'ha desenvolupat sobretot a (2002) i de Berlin (2003) i s'ha refet fins fa poc als working papers i a comunicacions a congressos.

a) Plan de Gestión de Datos (I)

- Guía que sirve de apoyo a los investigadores en el momento de crear su Plan de Gestión de Datos; concretamente, para los proyectos financiados en el marco de Horizonte 2020 de la Unión Europea
- Se muestran los diferentes campos requeridos en la convocatoria, como también diferentes aspectos concretos a tener en cuenta y ejemplos de PGD reales
- Se encuentra disponible en catalán (hdl.handle.net/2072/266502) y en inglés (hdl.handle.net/2072/266523) en RECERCAT
- Instrumentalización mediante la herramienta en línea DMPOnline (dmp.csuc.cat)

a) Plan de Gestió de Dades (II)

Plans de Gestió de Dades

(Versió 1, Juliol 2016)

(Doc. 16/31) (B6SRIGT Suport Recerca/PlansdeGestioDeDades_versio1Publica_julio16.pdf, 13.07.16)

Aquest document vol servir de recolzament als investigadors a l'hora de crear els seus Plans de Gestió de Dades; concretament, per als projectes finançats en el marc de l'Horitzó 2020 de la Unió Europea.

A continuació es mostra:

- Amb una lletra majúscula, els diferents camps que es requereixen a Horitzó 2020
- Amb un número, els elements que caldrà tenir en compte a l'hora d'omplir cada camp
- Amb una lletra minúscula, les descripcions de cada element i una mostra d'exemples reals

Aquest document ha estat elaborat pel Grup de Treball de Suport a la Recerca del CSUC que està format per representants de les universitats següents: Universitat de Barcelona, Universitat Autònoma de Barcelona, Universitat Politècnica de Catalunya, Universitat Pompeu Fabra, Universitat de Girona, Universitat de Lleida, Universitat Rovira i Virgili, Universitat Oberta de Catalunya, Universitat de Vic-Universitat Central de Catalunya, Universitat Ramon Llull i Universitat Jaume I.

Els exemples citats són una mostra¹ de diferents Plans de Gestió de Dades disponibles a la xarxa.

Aquest document està subjecte a la llicència de Reconeixement de Creative Commons (<http://creativecommons.org/licenses/by/4.0/>).

Versió digital: <http://hdl.handle.net/2072/266522>

¹ Actris (Grant 654109), Citilab (Grant 635896), ConnectingGEO (Grant 641536), EGI-Engage (Grant 654142), FREAM (Grant 644771), iCizms (Grant 644526), MMT (Grant 645487), RAMCIF (Grant 643433), SatisFactory (Grant 636302), Step (Grant 649493), Tsadem (Grant 654206), U-Mobile (Grant 645124), U-Turn (Grant 635773)

C. Descripció del conjunt de dades (dataset)

C.1 Descripció del dataset

C.1.a) Descripció

Text curt introductori que explica el contingut del dataset

C.1.b) Exemple real

Ex. 1 This dataset contains the posts and the contact details of all subscribers to the project collaborative platform.

Ex. 2 MMT is a translation memory obtained by finding and aligning at the sentence level parallel documents from the web.

Ex. 3 Dataset for incident detection, along with high-level activities and business processes monitoring (e.g. activities occurring at the shop-floor, etc.), obtained with thermal and depth cameras mounted at specific locations in the shop-floor. [...]

C.2 Col·lectiu a qui pot ser d'interès el conjunt de dades (dataset)

C.2.a) Descripció

Indicar per a quin col·lectiu poden ser d'interès les dades.

C.2.b) Exemple real

Ex. 1 The dataset will be valuable for benchmarking algorithms for object recognition, robotics navigation and grasping

C.3 Origen de les dades produïdes

C.3.a) Descripció

En el cas que les dades es generin dins el projecte, indicar l'origen de les dades.

C.3.b) Exemple real

Ex. 1 Dataset produced by simulation tools and/or by real life trials will be used as means to quantify the performance advantages the project architecture offers compared to current practices

Ex. 2

(1) Dataset1: Data collected through literature review, and questionnaire survey among project pilot partners (Countries).

(2) Dataset2: This data originates through qualitative and standardized open-ended interviews with 28 Young European Citizens and 12 Public Authorities/Policy Makers. Datasets generated for WP2 comply with the University Research Data Policy (<https://intranet...>)

a) Plan de Gestión de Datos (III)

Gestió de les Dades de Recerca
de les Universitats de Catalunya

Home About Contact Help

Sign in +

Sign up +

New to DMPonline? Sign up today.

The Pla de Gestió de Dades de les Universitats de Catalunya tool can help you write data management plans. The tool has been developed by **Consorti de Serveis Universitaris de Catalunya** based on the model created by the Digital Curation Centre.

Powered by: Coordination:

DMP
ONLINE

CSUC

Legal Disclaimer

UNIVERSITAT DE BARCELONA

UB

UNIVERSITAT POLITÈCNICA DE CATALUNYA

upf

Universitat Pompeu Fabra

Universitat de Girona



Universitat de Barcelona

UNIVERSITAT ROVIRA I VIRGILI

UOC

UVIC

UNIVERSITAT DE VIC

UNIVERSITAT CENTRAL DE CATALUNYA

UNIVERSITAT RAMON LLULL

UNIVERSITAT DE VALÈNCIA

Standards and metadata

B *I* -

← →

Save

Answered 3 months ago by Mireia Alcalá **Unsaved changes**

Guidance **Share note**

Quines normes o metodologies s'utilitzen?
guidance

Referència a estàndards de metadades de la disciplina. Si no existeixen, descripció de quines metadades es crearan i com.

Exemples reals:

1. Generic metadata standard: Dublin core <http://dublincore.org> (used by the X repository).
2. The documents are based on XML according to a DTD. The vocabulary is represented in SKOS. The RDF data is based on an OWL ontology.

Com s'anomenen i s'estructuren les carpetes i els arxius? guidance

Com s'identifiquen fàcilment les diferents versions d'un conjunt de dades? guidance

Com es capturen / creen les metadades? guidance






Quins estàndards de metadades utilitzarà i per què? guidance

b) Recomendaciones de repositorios temáticos y multidisciplinares (I)

- Documento para dar soporte a los investigadores en el momento de seleccionar un depósito para sus datos de investigación.
- Incluye el proceso a seguir para seleccionar un repositorio de datos y diferentes criterios a tener en cuenta
- Facilita fuentes para seleccionar repositorios temáticos y multidisciplinares (directorios, recomendaciones de editoriales, etc.)
- Se adjunta una tabla comparativa entre los principales repositorios multidisciplinares
- Se encuentra disponible en RECERCAT (<http://hdl.handle.net/2072/266502>)



b) Recomanacions de repositoris temàtics y multidisciplinaries (II)

	 zenodo CERN	 DRYAD NCSU	 Harvard Dataverse IOSS (Harvard University)	 figshare Digital Science (Macmillan)	 MENDELEY DATA Elsevier
1. Àmbit temàtic i geogràfic	Multidisciplinària Internacional	Amb èmfasi en ciència i medicina Internacional	Multidisciplinària Internacional	Multidisciplinària Internacional	Multidisciplinària Internacional
2. Tipologia de dades (programari, imatges, raw data, etc.)	<u>Tot tipus de formats</u> i també documents publicats: publicacions, literatura grisa, imatges i audiovisuals, text pla, etc.	<u>Tot tipus de formats</u> (amb exclusió de software) i amb preferència per formats de codi obert	<u>Tot tipus de formats</u> amb funcions específiques per alguns (Stata, SPSS, R, Excel, CSV, etc.)	<u>Tot tipus de formats</u> : xifres, conjunts de dades, imatges i audiovisuals, literatura grisa, conjunts d'arxius, etc.	<u>Tot tipus de formats</u> : text sense format, imatges i audiovisuals, PDF, HTML, etc.
3. Mida aproximada dels fitxers a incloure	Fins a 2 GB Per a mides superiors es pot <u>contactar</u> amb ells	Fins a 1 GB per fitxer i 20 GB per un conjunt de dades (<u>cost</u> addicional per sobre de 20 GB)	Fins a 2 GB. Per fitxers més grans contactar amb <u>support@dataverse.org</u>	Fins a 5 GB	10 GB
4. Dades obertes, embargades, restringides o tancades	Obertes Tancades Restringides Embargades (<u>Polítiques</u>)	Obertes <u>Embargades</u> (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	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 <u>CC0</u>	<u>CC0</u> (<u>FAQ</u>)	Per defecte <u>CC0</u> però es pot canviar	- <u>CC0</u> per a dades (<u>metadata & datasets</u>) - <u>CC-BY</u> per a no dades (audiovisuals, papers, etc.) - <u>MIT</u> per a codi i software	Permet varietat de llicències: CC-BY 4.0, CC-BY-NC, CC0 1.0, MIT, Apache 2.0, BSD 3 clause, BSD 2 clause, gplv3
6. Identificador permanent	DOI	DOI	DOI	DOI	DOI
7. Condicions per depositar o retirar contingut	Es permet <u>retirar</u> les dades justificant-ne el perquè	Cessió de dades <u>irrevocable</u>	Sense informació	Sense informació	Sense informació
8. Cost	Gratuit. Però hi ha la possibilitat de més espai pagant (<u>FAQS</u>)	110 € per un conjunt de dades (de fins a 20GB) (<u>Payment plans</u> <u>Data publishing charges</u>)	Gratuit	Gratuit fins a una certa mida: -5 GB pels arxius individuals -20 GB d'espai privat -Espai públic il·limitat -10 espais col·laboratius Per més espai i altres prestacions, ofereixen versió de pagament. <u>Contacte</u>	Gratuit

c) Política de gestión de datos de investigación (I)

- Antecedente y modelo: Mandatos de Acceso Abierto (2010)
- Acuerdo marco elaborado entre todas las universidades donde se definen los elementos mínimos a los que las universidades deben comprometerse en materia de gestión de datos de investigación
- Estructura y contenidos siguiendo las recomendaciones de *“Policy RECommendations for Open Access to Research Data in Europe (RECODE)”*
- Cada universidad deberá aprobar su propia política definiendo concretamente cada elemento

c) Política de gestión de datos de investigación (II)

Contenido de la política

AO por defecto

- AO como requerimiento estándar y de obligado cumplimiento (con la posibilidad de cerrar los datos en casos concretos)
- Licencias que permitan la reutilización

Competencias

- Definición de la responsabilidad de todos los agentes implicados
- Investigadores gestionaran sus datos siguiendo estándares y procedimientos específicos
- La institución asumirá la responsabilidad de ofrecer servicios necesarios para que lo puedan hacer

Lugar de depósito

- Determinación del lugar de depósito de los datos (repositorio institucional u otros repositorios de confianza)

Periodo de depósito

- Abierto tan pronto como sea posible
- Períodos razonables entre la obtención de datos y su publicación en abierto

Plan de Gestión de Datos

- PGD como requerimiento estándar (implementándose de manera progresiva)

Costes

- Mecanismos oportunos para repercutir los costes de infraestructura

Preservación y conservación

- Preservar y conservar los datos depositados durante un período de tiempo determinado (10-20 años, por ejemplo)

4. Trabajos futuros

Monitorizar el servicio de gestión de datos de investigación

Elaborar una propuesta de repositorio de datos

Aprobar las diferentes políticas de Acceso Abierto a los datos de investigación en cada universidad

Muchas gracias!

Mireia Alcalá Ponce de León
(CSUC)

mireia.alcala@csuc.cat

www.csuc.cat