

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

(Versió 3, Maig 2017)

(Doc. CO17/02) (Ciència Oberta\4 Gestio Dades Recerca\Repositori de dades\  
1705\_RecomanacionsSeleccionarRepositoriDades.docx, 15.05.17)

Aquest document té per objectiu recolzar els investigadors en el moment de seleccionar un repositori per al dipòsit de les seves dades de recerca atès que a partir de l'any 2017 l'*Open Research Data Pilot* (ORD Pilot) de l'Horizon 2020 requereix a tots els projectes que reben finançament la publicació de les dades de recerca en obert.

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.

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/2072/284974>

## 1. Selecció d'un repositori de dades de recerca

---

Avui en dia, la decisió d'escollir un repositori de dades pot arribar a ser tot un repte. Algunes disciplines acadèmiques han establert i consolidat certs repositoris de dades de recerca. En canvi, d'altres àmbits d'investigació no tenen aquestes possibilitats.

La selecció d'un o altre repositori comporta certa planificació a l'hora de generar les dades, ja que cal complir amb els requisits que aquests demanen (estàndards de metadades, documentació, drets d'autor, preservació...).

Els **repositoris de dades temàtics o específics per una disciplina** ofereixen el coneixement i l'experiència en la gestió de les dades en el seu context.

Els **repositoris multidisciplinars o de dades generals** permeten el dipòsit de dades de qualsevol àrea temàtica. Normalment ofereixen funcionalitats de navegació i visualització més amigables.

### 1.1 Procés per seleccionar un repositori de dades

En el procés de selecció d'un repositori de dades, cal seguir els següents passos (vegeu figura 1):

1. Comprovar si es disposa d'un repositori temàtic consolidat a l'àmbit [vegeu punt 2].
2. Si és possible, sol·licitar el dipòsit de dades al repositori institucional de la universitat.
3. Utilitzar Zenodo (repositori proposat per la Comissió Europea) o un altre repositori de dades multidisciplinars [vegeu punt 3].

En el cas que les dades es dipositin en un repositori temàtic o multidisciplinar, informar la universitat (unitats o serveis de suport a la recerca) sobre les dades dipositades i on s'han dipositat.

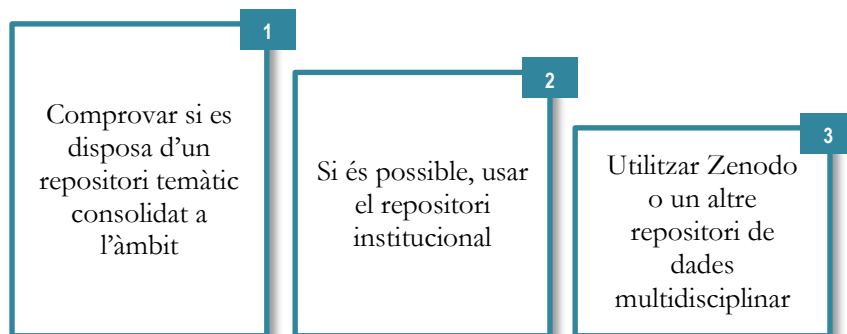


Figura 1. Proposta basada en les recomanacions d'OpenAIRE (vegeu bibliografia)

## 1.2 Criteris a tenir en compte a l'hora d'escollir un repositori de dades

A continuació, es mostren una sèrie de consideracions basades en la proposta de la John Hopkins University (vegeu bibliografia) que cal tenir en compte a l'hora de seleccionar un repositori de dades.

1. Existeix algun repositori representatiu i consolidat en el teu àmbit?
2. El repositori permet dipositar les dades en el format utilitzat?
3. La mida dels fitxers que pots dipositar s'ajusta a les teves necessitats?
4. Es permet la restricció d'accés a les dades (tancades, restringides o embargades)?
5. Els drets i llicències que t'ofereix el repositori s'ajusten amb els teus interessos?
6. El repositori atorga un identificador permanent?
7. Les condicions per dipositar o retirar contingut del repositori s'ajusten a les teves necessitats?
8. Hi ha costos associats a l'ús del repositori?
9. Es poden enllaçar les dades a les publicacions associades?
10. El repositori té una política de preservació (còpies, terminis de conservació)?
11. Permet dipositar diferents versions d'un mateix arxiu?
12. Compleix els requisits dels ajuts europeus?

## 2. Fonts per seleccionar un repositori de dades temàtic

---

A continuació, es mostren directoris de repositoris de dades de recerca i recomanacions d'editorials a l'hora de seleccionar un repositori de dades de recerca.

### 2.1 Directoris

Els següents directoris de repositoris de dades de recerca cobreixen diferents àrees de coneixement. Es presenten els resultats ordenats alfabèticament per disciplina.

- [Re3data](http://bit.ly/Re3Data) (bit.ly/Re3Data)
- [Data repositories OAD](http://bit.ly/OADDataRepositories) (bit.ly/OADDataRepositories)

### 2.2 Recomanacions d'editorials

Algunes editorials recomanen repositoris temàtics per dipositar les dades suplementàries a les publicacions. Vegeu com a exemple:

- [Elsevier](http://bit.ly/ElsevierDataRepositories) (bit.ly/ElsevierDataRepositories)
- [Nature](http://bit.ly/NatureDataRepositories) (bit.ly/NatureDataRepositories)
- [Plos](http://bit.ly/PlosDataRepositories) (bit.ly/PlosDataRepositories)
- [Springer](http://bit.ly/SpringerDataRepositories) (bit.ly/SpringerDataRepositories)

## 3. Fonts per seleccionar un repositori de dades multidisciplinar

---

A continuació es mostra una taula comparativa dels diferents repositoris multidisciplinars realitzada per la Universitat Autònoma de Barcelona i ampliada i revisada pel Grup de Suport a la Recerca del CSUC.

	 <a href="#">CERN</a>	 <a href="#">EUDAT</a>	 <a href="#">Digital Science (Macmillan)</a>	 <a href="#">NCSU</a>	 <a href="#">IOSS (Harvard University)</a>	 <a href="#">Elsevier</a>
<b>1. Àmbit temàtic i geogràfic</b>	Multidisciplinar Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional	Amb èmfasi en ciència i medicina Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional
<b>2. Tipologia de dades (programari, imatges, raw data, etc.)</b>	<a href="#">Tot tipus de formats</a> i també documents publicats	<a href="#">Tot tipus de formats</a>	<a href="#">Tot tipus de formats</a>	<a href="#">Tot tipus de formats</a> (amb exclusió de software) i amb preferència per formats de codi obert	<a href="#">Tot tipus de formats</a> amb funcionalitats addicionals per alguns (RData, SPSS, STATA, CSV, xlsx, FITS)	<a href="#">Tot tipus de formats</a>
<b>3. Mida aproximada dels fitxers a incloure</b>	Fins a 50 GB Per a mides superiors es pot <a href="#">contactar</a> amb ells	Fins a 2 GB. Per a mides superiors, <a href="#">contactar</a> amb ells	Fins a 5 GB	20 GB per un conjunt de dades. Per a mides superiors revisar les <a href="#">condicions</a>	Fins a 2 GB.	10 GB
<b>4. Dades obertes, embargades, restringides o tancades</b>	Obertes Tancades Restringides Embargades ( <a href="#">Polícies</a> )	Obertes Restringides Embargades	Obertes Tancades (per defecte fins que no es publiquen) Restringides Embargades	Obertes <a href="#">Embargades</a> (per defecte fins el moment de la publicació)	Obertes Restringides	Obertes Tancades (per defecte fins que no es publiquen) Restringides Embargades
<b>5. Llicències amb les que es volen difondre les dades</b>	Permet varietat de llicències encara que recomanen llicències obertes. La descripció de les dades està sota <a href="#">CC0</a>	Permet varietat de llicències per a dades i per a software	- <a href="#">CC0</a> per a dades ( <i>metadata &amp; datasets</i> ) - <a href="#">CC-BY</a> per a no dades (audiovisuals, <i>papers, etc.</i> ) - <a href="#">MIT</a> per a codi i software - Altres llicències: GNU, Apache 2.0, BSD 3 clause	<a href="#">CC0</a> ( <a href="#">FAQ</a> )	Per defecte <a href="#">CC0</a> 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, gplv3
<b>6. Identificador permanent</b>	DOI o URL	Handle	DOI	DOI	DOI	DOI
<b>7. Condicions per retirar contingut</b>	Es permet <a href="#">retirar</a> les dades justificant-ne el perquè. El DOI i la URL quedaran retinguts.	Per retirar contingut contactar amb <a href="#">EUDAT</a>	Es pot sol·licitar la <a href="#">retirada</a> de contingut públic justificant-ne el motiu	Cessió de dades <a href="#">irrevocable</a>	Es permet <a href="#">retirar</a> les dades, tot i que les metadades quedaran sempre visibles	Per retirar el contingut contactar amb <a href="#">Mendeley</a> .

<b>8. Cost</b>	Gratuït. Però hi ha la possibilitat de més espai ( <a href="#">FAQS</a> )	Gratuït	Gratuït 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. <a href="#">Contacte</a>	110 € per un conjunt de dades de fins a 20GB. <a href="#">Cost</a> addicional per sobre de 20 GB ( <a href="#">Payment plans</a> )	Gratuït	Gratuït
<b>9. Dades enllaçades a les publicacions</b>	Permet afegir enllaços a documents relacionats	Permet afegir enllaços a documents relacionats	Permet afegir enllaços al text complet o a documents relacionats	-Les dades han d'estar associades amb un document de recerca <a href="#">publicat</a> o en procés de publicació - Integració amb les publicacions ( <a href="#">Submission integration</a> )	Permet afegir enllaços a documents relacionats	Permet afegir enllaços a documents i programaris relacionats
<b>10. Preservació</b>	<a href="#">Dades emmagatzemades</a> durant mínim 20 anys al <a href="#">CERN Data Center</a>	Com a mínim 2 anys i amb garanties de migració a altres repositoris (el centre de dades d'emmagatzematge està certificat per la ISO)	<a href="#">Preservació (Preservation policies)</a>	<a href="#">Dades emmagatzemades</a> a <a href="#">DataONE</a>	<a href="#">Harvard Dataverse Preservation Policy</a> . Conveni amb Data Preservation Alliance per preservar les dades a llarg termini	<a href="#">Dades emmagatzemades</a> a <a href="#">Amazon S3</a> , a Alemanya. Conveni amb DANS per preservar les dades a llarg termini
<b>11. Versions dels documents</b>	Permet tenir <a href="#">diferents versions</a> d'un arxiu	Sense informació	Permet tenir <a href="#">diferents versions</a> d'un arxiu ja públic	Permet tenir <a href="#">diferents versions</a> d'un arxiu	Permet tenir <a href="#">diferents versions</a> d'un arxiu	Permet tenir diferents versions d'un arxiu
<b>12. Compleix els requisits dels ajuts europeus</b>	Recomanat per projectes europeus. Desenvolupat pel CERN (projecte OpenAIRE 7PM de la UE)	Recomanat per als projectes europeus. Projecte finançat per la CE amb participació del <a href="#">BSC</a> , <a href="#">DANS</a> , <a href="#">LIBER</a> i la <a href="#">Univ. d'Edimburg</a> entre d'altres	Sense informació	Sense informació	Sense informació	Sense informació
<b>13. Més informació (URL)</b>	<a href="#">FAQS</a> <a href="#">Policies</a> <a href="#">Features</a> <a href="#">About</a>	<a href="#">FAQS</a> <a href="#">DARUP</a>	<a href="#">FAQS</a> <a href="#">Features</a> <a href="#">Blog</a> <a href="#">Tools</a> <a href="#">About</a>	<a href="#">FAQS</a> <a href="#">Terms of service</a> <a href="#">Wiki</a> <a href="#">Blog</a>	<a href="#">User guide</a> (Dataverse general) <a href="#">Demo</a> (proves Harvard) <a href="#">About</a> (general) <a href="#">Blog</a>	<a href="#">FAQS</a> <a href="#">Terms of use</a>
<b>14. Altres</b>	- Integració amb ORCID i GitHub - Espais col·laboratius ( <i>Community collections</i> ) - Permet exportar cites a gestors bibliogràfics i xarxes socials	- El cercador de dades <a href="#">B2FIND</a> permet cercar dades de <a href="#">B2SHARE</a> i d'altres repositoris - Es pot utilitzar de manera independent <a href="#">B2DROP</a> com una infraestructura per actualitzar i intercanviar dades	- <a href="#">Integració amb ORCID</a> - <a href="#">Espais col·laboratius</a> - Estadístiques de descàrregues - Les dades públiques es veuen a Google Scholar - Permet afegir al perfil personal enllaços a xarxes socials i a les nostres publicacions - Permet exportar cites a gestors bibliogràfics i xarxes socials	- <a href="#">Integració amb ORCID</a> - Estadístiques de descàrregues - Permet exportar cites a gestors bibliogràfics i xarxes socials	- Aplicació de codi obert (una institució o investigador pot crear el seu propi Dataverse) - <a href="#">Integració amb OJS</a> i DropBox - Permet obtenir la citació en format XML o RIS - Permet compartir un conjunt de dades a les xarxes socials	- Estadístiques de descàrregues

## 4. Bibliografia

---

OpenAIRE. *Research Data Management. Briefing paper : Understanding Research Data Management* [en línia]. European Commission, February 2016. Disponible a: <[bit.ly/2016OpenAIRE](http://bit.ly/2016OpenAIRE)> [Consulta: 15.05.2017]

*Repositoris multidisciplinars per a dades en Accés Obert* [en línia]. Barcelona: Universitat Autònoma de Barcelona. Servei de Biblioteques, 2016. Disponible a: <[bit.ly/2016UAB](http://bit.ly/2016UAB)> [Consulta: 15.05.2017]

*Selecting a Repository for Data Deposit* [en línia]. Maryland: Johns Hopkins University Data Management Services, 2016. Disponible a: <[bit.ly/2016Hopkins](http://bit.ly/2016Hopkins)> [Consulta: 15.05.2017]

Whyte, A. *Where to keep research data: DCC checklist for evaluating data repositories* [en línia]. Edinburgh: Digital Curation Centre, 2015. Disponible a: <[bit.ly/2015Whyte](http://bit.ly/2015Whyte)> [Consulta: 15.05.2017]