



¿QUÉ HACEN LOS INVESTIGADORES CON LOS DATOS DE LOS PROYECTOS?

FERNANDA PESET
[HTTP://WWW.DATASEA.ES](http://www.datasea.es)

datos-ua.es

**ENCUENTRO
DE DATOS
ABIERTOS DE LA
UNIVERSIDAD
DE ALICANTE**
#EDAUA2015

28 de abril de 2015
Sala Multimedia del
edificio de Rectorado de la
Universidad de Alicante
De 09:00h a 20:00h

Universitat d'Alacant
Universidad de Alicante

Consulta nuestro programa en:
<http://datos.ua.es/encuentro-de-datos-abiertos.html>
El programa se celebrará el día 28 de abril en el aula
de informática de la Universidad de Alicante.

Encuentro de datos abiertos de la
Universidad de Alicante
#EDAUA2015
Alicante 28 abril 2015

La intervención trata de mostrar la relación que existe entre los datos generados por los investigadores (RD researchdata) y los datos de la universidad (OD opendata). En estos momentos, observamos un amplio movimiento internacional trabajando sobre datos (OD y RD) que se ha trasladado al ámbito nacional en parte. El desarrollo de opendata está resolviéndose en los portales de datos generales y de las organizaciones. Sin embargo, los datos de investigación tienen una complejidad asociada a la internacionalización de los equipos y la sectorialización por disciplinas. En el ecosistema de la gestión de los datos de investigación apreciamos una infraestructura social de la ciencia formada por múltiples factores: las personas, la tecnología, los estándares, las culturas de trabajo...

En esta charla mostraremos en primer lugar qué dicen los investigadores que hacen con sus datos mediante los resultados de varias encuestas. En segundo lugar, abordaremos cómo los están compartiendo y los modelos que existen. Y por último, recomendaremos las pautas que consideramos correctas para gestionar los datos de los proyectos en una institución.

Voy a hablar de...

1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. Beneficios openscience
4. Entender qué hace un científico en su carrera
5. Infraestructura social
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo deberían proceder?
8. A modo de conclusión

1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. Beneficios openscience
4. Entender qué hace un científico en su carrera
5. Infraestructura social
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo deberían proceder?
8. A modo de conclusión

Realidad actual



Comisión Europea. **Hacia una economía de los datos próspera** (COM(2014) 442 final, 2.7.2014)



NSF financiación On March 29, 2012, six federal grant-making departments and agencies announced more than \$200 million in grant opportunities for the so-called “**Big Data research and Development Initiative**” (Office of Science and Technology Policy [OSTP] 2012).



DCC 10th international conference.. London, 9 - 12 February 2015



A taxonomy of openness

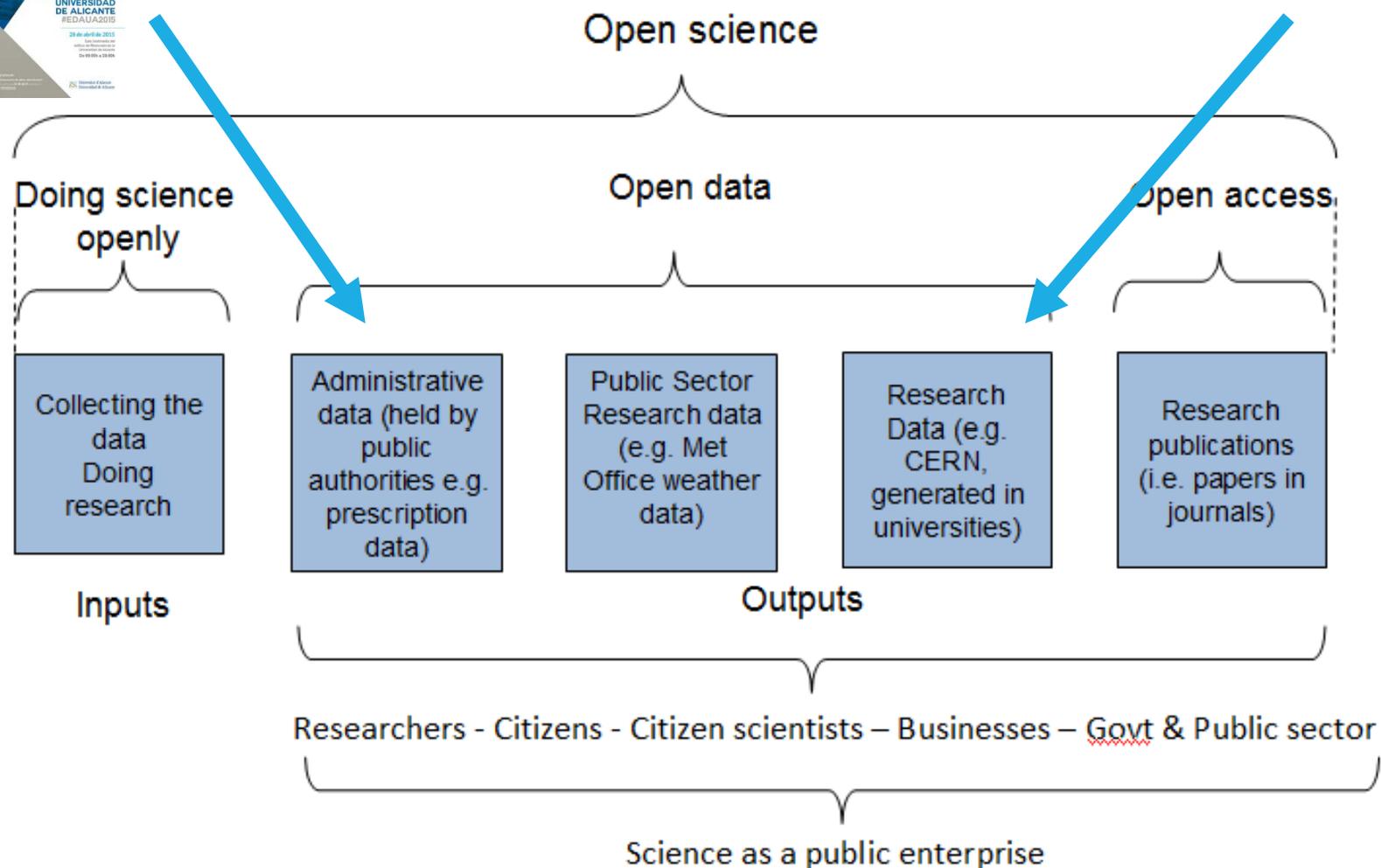


Figura 4. A taxonomy of openness. Boulton, 2013



www.datasea.es



OPENDATASCIENCE, centro de recursos para la preservación y gestión de datos abiertos de investigación", CS02012-39632-C02-02

<https://media.upv.es/player/?autoplay=true&id=a48abcd-21bc-ab4a-ac6a-af987b3e7f1a>

- Entender la situación española: encuesta
- Promover buenas prácticas: guías
- Diseñar un buscador de conjuntos de datos: [nonamed] cosechador contra repositorios y revistas
- Agenda / Noticias / ...

1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. Beneficios openscience
4. Entender qué hace un científico en su carrera
5. Infraestructura social
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo deberían proceder?
8. A modo de conclusión

What is Open Data?

Automático? ¿?

Data is open if it satisfies both conditions below:

Technically open: available in a machine-readable standard format, which means it can be retrieved and meaningfully processed by a computer application

Legally open: explicitly licensed in a way that permits commercial and non-commercial use and re-use without restrictions.

Word Bank <http://data.worldbank.org/about/open-government-data-toolkit/knowledge-repository>

Lo libre, lo open... recordemos

Gratis. Sólo implica el acceso sin coste, pero no otras dimensiones de reutilización y difusión. “Free” tiene esa acepción.

Libre. A menudo se mantiene en su forma española para diferenciarlo de lo simplemente gratis. En realidad este es el término que acoge otras connotaciones con la reutilización del contenido sin restricciones.

Abierto. Desde su semántica inicial, de código accesible y manipulable, incluye actualmente la posibilidad técnica y legal de transformarse.

Villamón, Miguel; Antonia Ferrer-Sapena; Rafael Aleixandre-Benavent; Tomás Saorín; Fernanda Peset, Influencia de los movimientos open en la transferencia de conocimiento *Universidad y Propiedad Intelectual: el presente y los Nuevos Retos* Universidad de Valencia, 24 y 25 de noviembre de 2011

Datos de investigación

“datos de investigación [es] todo aquel material que ha sido **registrado durante la investigación**, reconocido por la comunidad científica y que sirve para **certificar** los resultados de la investigación que se realiza. [...] debe provenir de una **f fuente única** y deben ser **difíciles o imposibles de obtener de nuevo**”

Torres-Salinas; Robinson-García; Cabezas-Clavijo, 2012

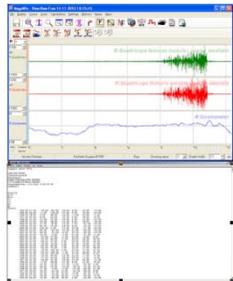


Figura 5. Datos recogidos de una electrografía de superficie (parte superior) y el correspondiente fichero en un formato plain text (parte inferior).

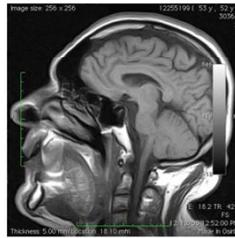


Figura 13. Imagen de tomografía axial con

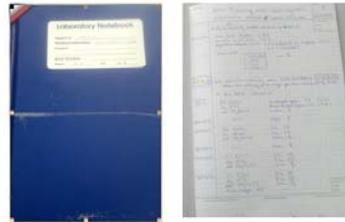


Figura 6. Ejemplo de cuaderno de laboratorio con anotaciones diarias. Laboratorio del prof. Dr. J.L. Toca (BOKU)

Figura 11. Datos tratados y publicados (García-Massó et al. en prensa).

Otros incluyen publicaciones (OECD) ¿?¿

Los datos incluyen: “cuadernos de laboratorio, cuadernos de campo, datos de investigación primaria (incluidos los datos en papel o en soporte informático), cuestionarios, cintas de audio, videos, desarrollo de modelos, fotografías, películas, y las comprobaciones y las respuestas de la prueba. Las colecciones de datos para la investigación pueden incluir diapositivas; diseños y muestras. [...] El código de software “

Universidad de Melbourne, del Informe Recolecta



¿Qué significa comunicar los datos (de investigación) de forma abierta?

Open data de Royal Society

Crterios a cumplir cuando son liberados:

- Accesibles (accessible): fáciles de encontrar y en una forma en que puedan ser usados [y preservados].
- Evaluados/certificados (asessable): credibilidad para diferentes grupos de interés.
- Inteligibles (intelligible): ser entendidos
- Reutilizable (useable): en un formato para usar y con licencias adecuadas.

include the JPEG 2000, PNG and SVG standard image formats; ASCII, PDF, Open Document Format and Office Open XML format (the native format for recent versions of Microsoft Word) for text; HTML, XHTML, RSS and CSS for the web and NetCDF for some scientific data.

Creative Commons' (CCZero) y Licencia al Dominio Público (Public Domain Dedication & Licence PDDL-OKF). Principios Panton (OKF)

Documentar : time consuming

Para recordar, ayudar a otros, verificar, replicar, archivar, reclamar autorías...

Ejemplos: cuadernos de laboratorio, notas de campo, metodologías cualitativas

- **Nivel de Proyecto:** documentar la base del estudio, métodos, instrumentos, hipótesis de trabajo
- **Nivel de archivo o dataset:** formatos, relaciones entre archivos
- **Nivel de variable o item:** como fue generada la variable y descripción de los campos

Experimento ESO-8 15-10-2003
ACETILACION DE SALICILALDEHIDO
Coryth, J., Du, M.H. J. Chem. Soc. Perkin Trans I, 2183 (199)

	1	2	3	4	5
Peso molecular	122.12	102.09	79.10	84.93	164.16
Peso (g)	2.29	2.17	2.94	—	2.9
Pureza (%)	99	98	99	98	—
milimols	18.56	20.79	36.76	—	17.67
Equivalentes	1.15	1.12	1.98	—	—
Densidad (g/ml)	1.146	1.082	0.978	1.325	—
Volumen (ml)	2.0	2.0	3.0	3.5	—

Materia: equipo para calentamiento a reflujó
Procedimiento: Reactivos disueltos en CH₂Cl₂ y reflujados
Proceso: Adición de hielo (2g), agua (20 ml). la
suspensión se lava con HCl 2N (3 x 15 ml), con
NaHCO₃ sol. sat. (3 x 15 ml) y NaCl sol. sat. (1
Se seca con Na₂SO₄, filtra y evapora.
Peso obtenido: 3g

Destilación a presión reducida (trampa agua, 18 mm Hg
140-142°C.

Rendimiento: $\frac{17.67 \text{ mmol producto}}{18.56 \text{ mmol reactivo limitante}} = 95\%$

Infra-rojo y ¹H RMN (Clave ESO-8P1)

1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. **Beneficios openscience**
4. Entender qué hace un científico en su carrera
5. Infraestructura social
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo *deberían* proceder?
8. A modo de conclusión

Compartir datos tiene muchos beneficios, algunos nuevos/proprios

- ✓ responder de forma rápida y mucho más eficiente ante las emergencias (e-coli 2011)
- ✓ capacidad potencial de señalar los fraudes (datos inventados: reciente polémica del artículo de Science) o las malas prácticas (tensionar la integridad del sistema o de las personas).
- ✓ estimula nuevas y altamente creativas formas de colaboración científica
- ✓ estimula un movimiento social hacia la ciencia, quizá fundamental para el cambio de las dinámicas científicas: Galaxy Zoo, Fold-it, Ash-Tag, etc.

Pero además...

los **ciudadanos**, quienes pagan con sus impuestos la ciencia, demandan cada vez más las evidencias que dan lugar a las políticas que se les acaban imponiendo en todos los órdenes sociales

única forma de afrontar **retos globales** que implican a toda la sociedad: cambio climático, crisis energética, las pandemias o residuos mundiales.

1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. Beneficios openscience
4. Entender qué hace un científico en su carrera
5. Infraestructura social
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo deberían proceder?
8. A modo de conclusión

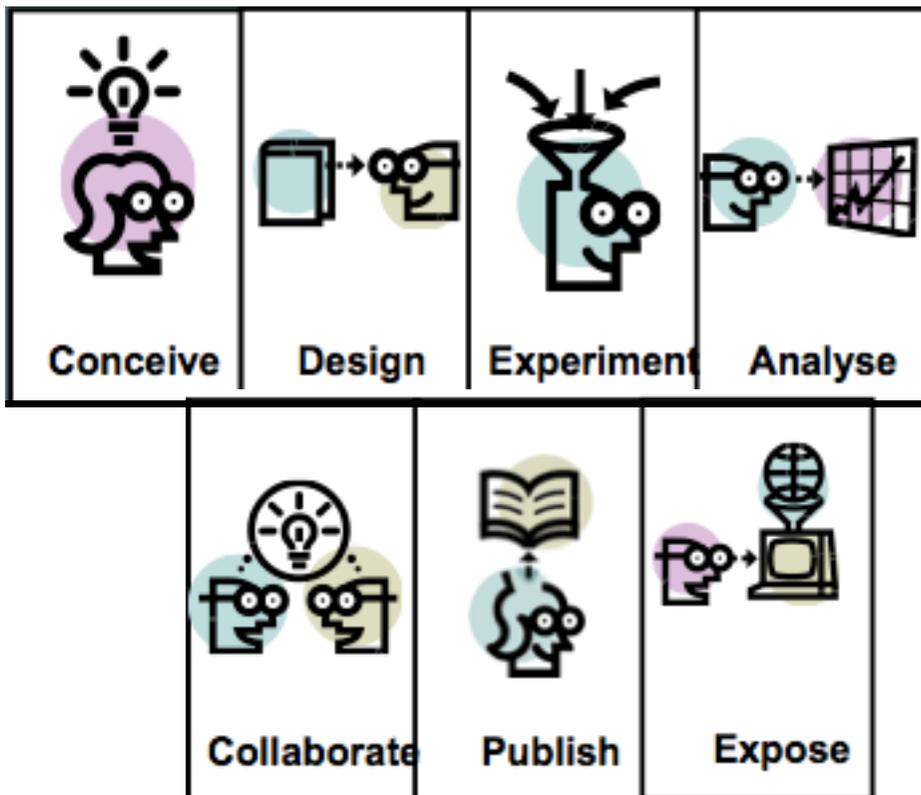


Imagen de Anthony Beitz

CREATING DATA

- design research
- plan data management (formats, storage etc.)
- plan consent for sharing
- locate existing data
- collect data (experiment, observe, measure, simulate)
- capture and create metadata

PROCESSING DATA

- enter data, digitise, transcribe, translate
- check, validate, clean data
- anonymise data where necessary
- describe data
- manage and store data

ANALYSING DATA

- interpret data
- derive data
- produce research outputs
- author publications
- prepare data for preservation

PRESERVING DATA

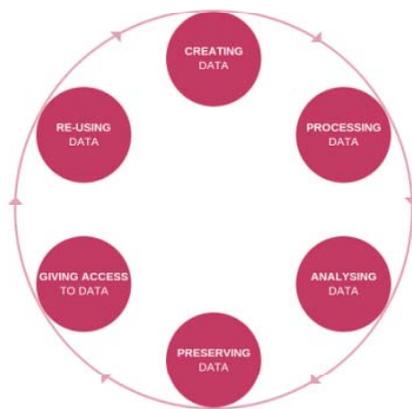
- migrate data to best format
- migrate data to suitable medium
- back-up and store data
- create metadata and documentation
- archive data

GIVING ACCESS TO DATA

- distribute data
- share data
- control access
- establish copyright
- promote data

RE-USING DATA

- follow-up research
- new research
- undertake research reviews
- scrutinise findings
- teach and learn



^[1] <http://datalib.edina.ac.uk/mantra/>

¿qué interesa a un investigador?

Producción
Explotación
Networkng/Datasharing

Su trabajo

Reconocimiento
Financiación
Rapidez

Sus intereses

Preservación
Compartir

**Nuevas
obligaciones**

1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. Beneficios openscience
4. Entender qué hace un científico en su carrera
5. **Infraestructura social**
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo *deberían* proceder?
8. A modo de conclusión

Mundo complejo

La **infraestructura social de la ciencia** está formada por múltiples factores:

- las personas,
- la tecnología,
- los estándares,
- las culturas de trabajo

Los objetos **digitales-datasets** tienen sus propios

- contextos disciplinares
- regulaciones legales
- requisitos de los financiadores

Culturas de trabajo de los protagonistas

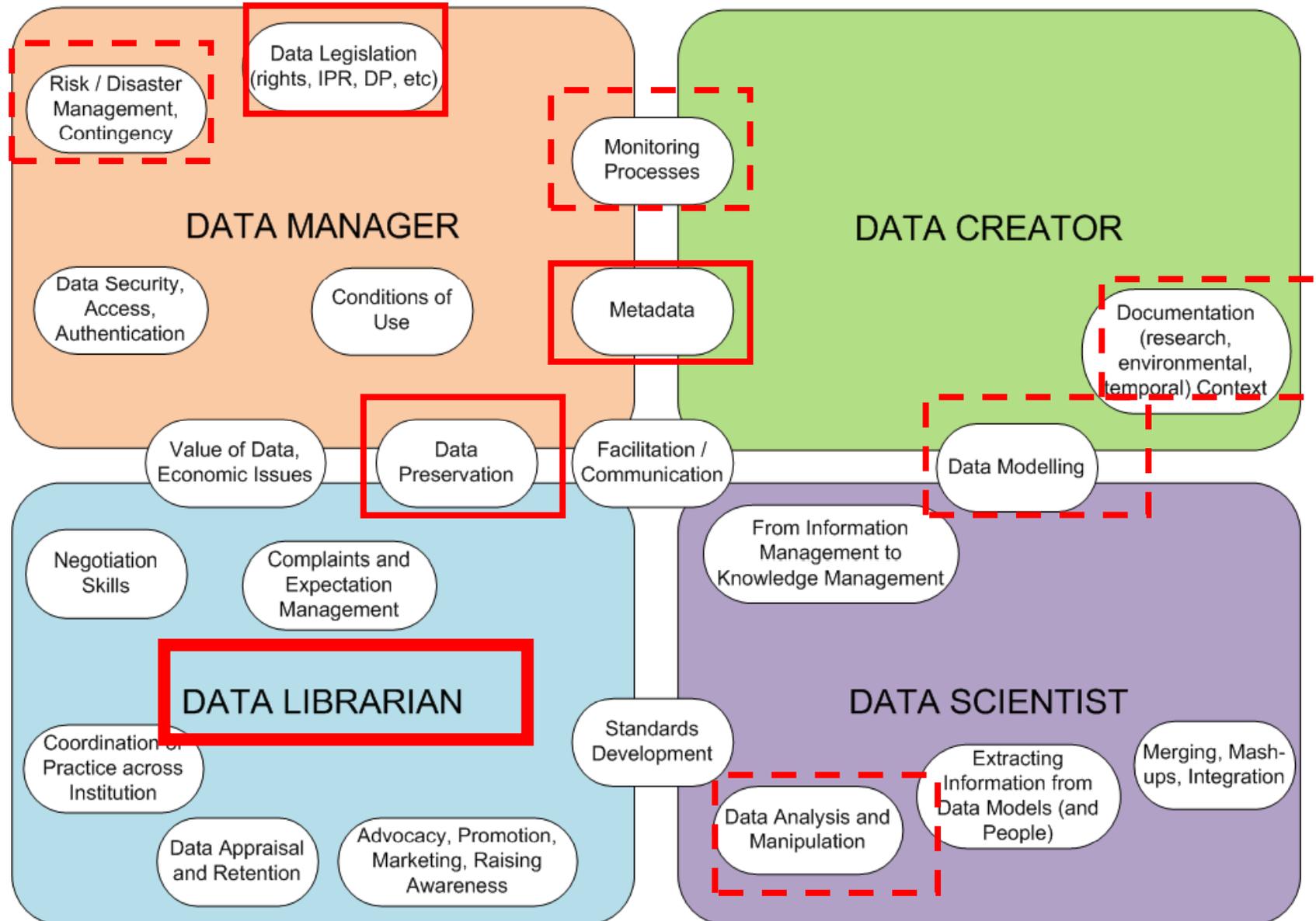
Investigadores : buscan el refuerzo de la comunidad científica

Financiadores y autoridades académicas : persiguen una imagen de marca

Infomediarios : buscan las formas más eficientes

CORE SKILLS FOR DATA MANAGEMENT

A follow-up from the second DCC Research Data Management Forum (November 2008)



1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. Beneficios openscience
4. Entender qué hace un científico en su carrera
5. Infraestructura social
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo deberían proceder?
8. A modo de conclusión

¿Qué están haciendo con los datos de sus proyectos?

¿Conocéis algún MANDATO (=Ley de la ciencia) en España?



Es que no hay



¿Conocéis algún MANDATO en el mundo?

- Recomendaciones de revistas (extranjeras)
- Recomendaciones de financiadores en la solicitud: Data management plan
 - En vigor NSF
 - En pruebas Opendatapilot del H2020



Técnica de encuesta

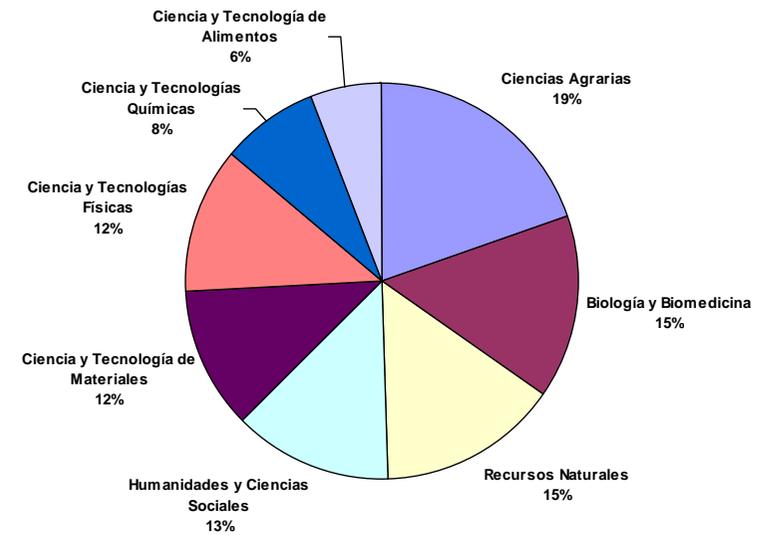
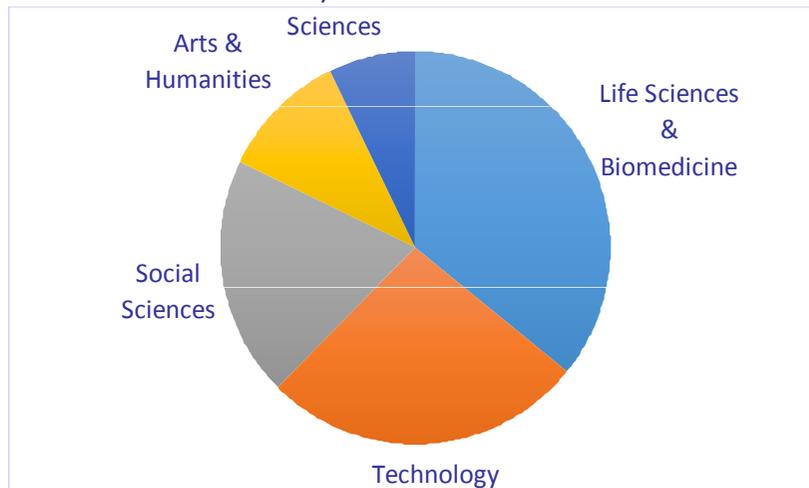
Realismo

Encuesta DATASEA con aspectos relacionados con el investigador y con la institución:

- Respecto a investigadores se trata de:
 - Averiguar los datos activos que se mantienen dentro de las instituciones
 - Estudiar cómo se almacenan estos datos, se administran, comparten y reutilizan
 - Identificar los riesgos, como por ejemplo el mal uso, la pérdida de datos, o la irrecuperabilidad de los mismos
 - Conocer cómo se han creado y las posibilidades de intercambio (si usan páginas web, repositorios...)
- Respecto a la institución se trata de identificar:
 - Acciones de *curation* que son aplicables a los datos de investigación
 - Diseñar los agentes y los flujos principales en una organización en España

Datos nacionales: DATASEA y CSIC

Disciplinas



DATASEA <https://docs.google.com/forms/d/1NNS5tuUVOYOZLtmYUGSb7YWbT7oXEFgm30e7nCDDhiU/viewform>

Bernal, Isabel; Román Molina, Juan (2014). *Prácticas en la gestión, difusión y preservación de datos de investigación en el CSIC* <https://digital.csic.es/handle/10261/92404>

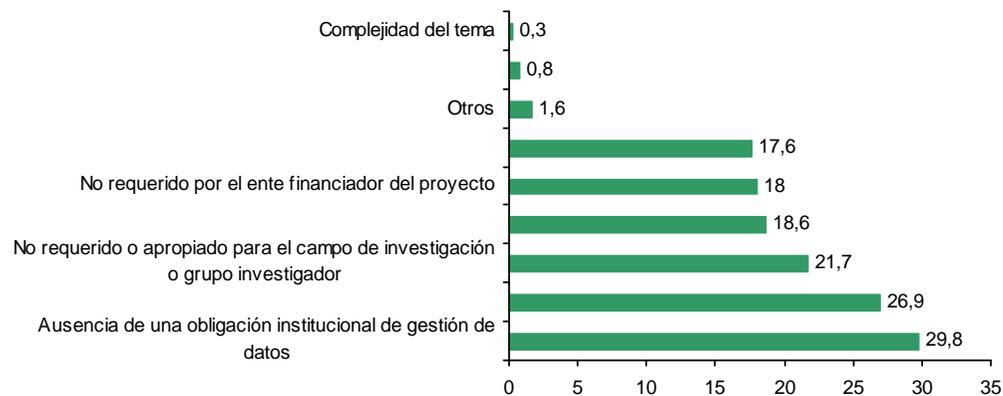
No suele haber normas para datos



En los proyectos de investigación en los que ha participado, ¿existe alguna política o plan de gestión de los datos que incluya algunos de estos aspectos?



Razones de no tenerla: no requerida



En caso de no tenerla, ¿es debido a alguna de estas razones?



Encuestas

España

UPV

UV

UM

Hospital Severo Ochoa

Clínico Valencia

CSIC con variaciones

Internacional

Parse.Insight

DATARES

...

METHODS

40 questions were emailed to 3 public universities and 2 hospitals

RESULTS

Main issues in Datasea

607 respondents about data,
mainly male researchers

Parse.Insight for data and publications

1.389 respondents from
researchers, mainly european
from UK

Datares

296 respondents
about data,
mainly librarians

Coincident=  Not coincident=  Not appear= 

	Parse.Insight for data and publications	Datares
Major disciplines: Life Sciences (39%)		
Does your project have a data plan?: Not really, only for participant's anonymization and how to storage the data		
Storage: 80% keep their current data either on local computers or on machines that collect data		
Access methods for reusing: personal contacts, Google and institutional database		
Barriers for sharing: legal issues (46%) and possible misuse (42%)		
Infrastructures for sharing: REAL, institutional repository (53%); IDEAL, infrastructure should be by discipline 60%)		
Motivation for sharing: - re-analysis of existing data - advancement of science		
Services of their organization: they do not know		

1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. Beneficios openscience
4. Entender qué hace un científico en su carrera
5. Infraestructura social
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo deberían proceder?
8. A modo de conclusión

Gestiona y Comparte Datos de Investigación



Son los datos que se generan durante la investigación y validan los resultados de la ciencia

Durante LA SOLICITUD DE FINANCIACIÓN

INFORMATE SOBRE:

- Formas de almacenamiento
- Costes

+ info



INFORMATE SOBRE:

Organización:

- Estructura los directorios
- Nombra los ficheros
- Documenta los datos

Acceso:

- Propiedad de los datos
- Almacenamiento
- Seguridad al compartir
- Integridad de los datos

Durante EL PROYECTO

+ info



PUBLICACIÓN EN ABIERTO DE LOS DATOS

INFORMATE SOBRE:

Selección de los datos:



Prepara:

- Fichero
- Metadatos
- Documentación
- Licencia



Dónde publicarlos:

- Junto al artículo
 - DiSEA
- En un banco o repositorio
 - zenodo
- En una revista de datos
 - SCIENTIFIC DATA

Cómo citarlos:



http://www.datasea.es/dt/index.php?option=com_faqbook&view=category&id=23&Itemid=134

Solicitudes Horizon2020



En el **artículo 29.3** del *Agreement* de los socios aparece *Open access to research data*, con las condiciones

Open Research Data Pilot orientada a mejorar el acceso y reutilización. Las áreas que participan recibirán (2014-2015) sobre 3.000 millones de euros (Vlaeminck, 2013).

Los beneficiarios deben:

- depositar en un repositorio abierto de datos de investigación: los datos necesarios para validar las publicaciones de investigación, con sus metadatos, tan pronto como sea posible; y otros datos, con sus metadatos, dentro del plazo previsto en sus planes de gestión de datos así como el momento de liberación que indicaron;
- ofrecer información en el repositorio sobre las herramientas para validar la investigación e incluso, si es posible, proporcionarlas

DataSharing



“... the practice of making data used for scholarly research available to others (sin ayuda del productor original).” [Wikipedia]

Involucra a: productor de los datos, el repositorio y el personal de apoyo, y el usuario que los consume

Consideraciones: pensar en el futuro (formatos, software, anonimización/ética/confidencialidad, documentación); descripción y documentación; futura gestión de derechos y licencias; infraestructura

(Anne Donnelly)

Privado o Público ¿?¿?

¿Qué está sucediendo?

Opciones habituales: dropbox, drive, pero también es posible en Figshare o Dataverse

Otras opciones: muchos proyectos crean plataformas a medida

dispendio O investigación ¿?¿?

Data curation



“any services related to the organization, management, or long-term preservation of data developed through scholarly research.

These services encompass a range of activities, including **consultations on creating data management plans and strategies, physical or electronic archiving of datasets, and workshops**”

(Deards, 2013).

Liberar datos: licencias (OKF)

OpenDataCommons Attribution License (ODC-BY), solo obliga a atribuir/reconocer el propietario original. Permite la copia, distribución y uso para producir obras derivadas sobre ella y modificar, transformar y construir sobre ella para cualquier propósito.

PDDL, Public Domain Dedication and License, similar a la CC más permisiva (CC0) sin atribución ni limitación geográfica de uso.

Open Database License, ODbL, es más restrictiva, pero permite compartir –copiar, distribuir y usar la base de datos—, producir trabajos sobre ella y modificar, transformar y construir otros productos sobre la base de datos. Mantiene la (Attribution y obliga a ceder el nuevo producto con esta misma licencia (Share-Alike). Añade condiciones a las bases de datos derivadas y puede aplicar restricciones tecnológicas (DRM) a los datos originales o los derivados. Por tanto, reduce la interoperabilidad y la reutilización

Formas de compartir/liberar

Repositorio de los financiadores

[ESRC Data Store-Reshare](#)

En España ...

Repositorio Institucional

[Edinburgh DataShare](#)

En España!!!!

Repositorio por Disciplina/Formato

Directorios: Re3data

[Figshare](#)

Y... en las revistas [ODiSEA](#)

(Donnelly) Edina MANTRA

Repositorios de datos

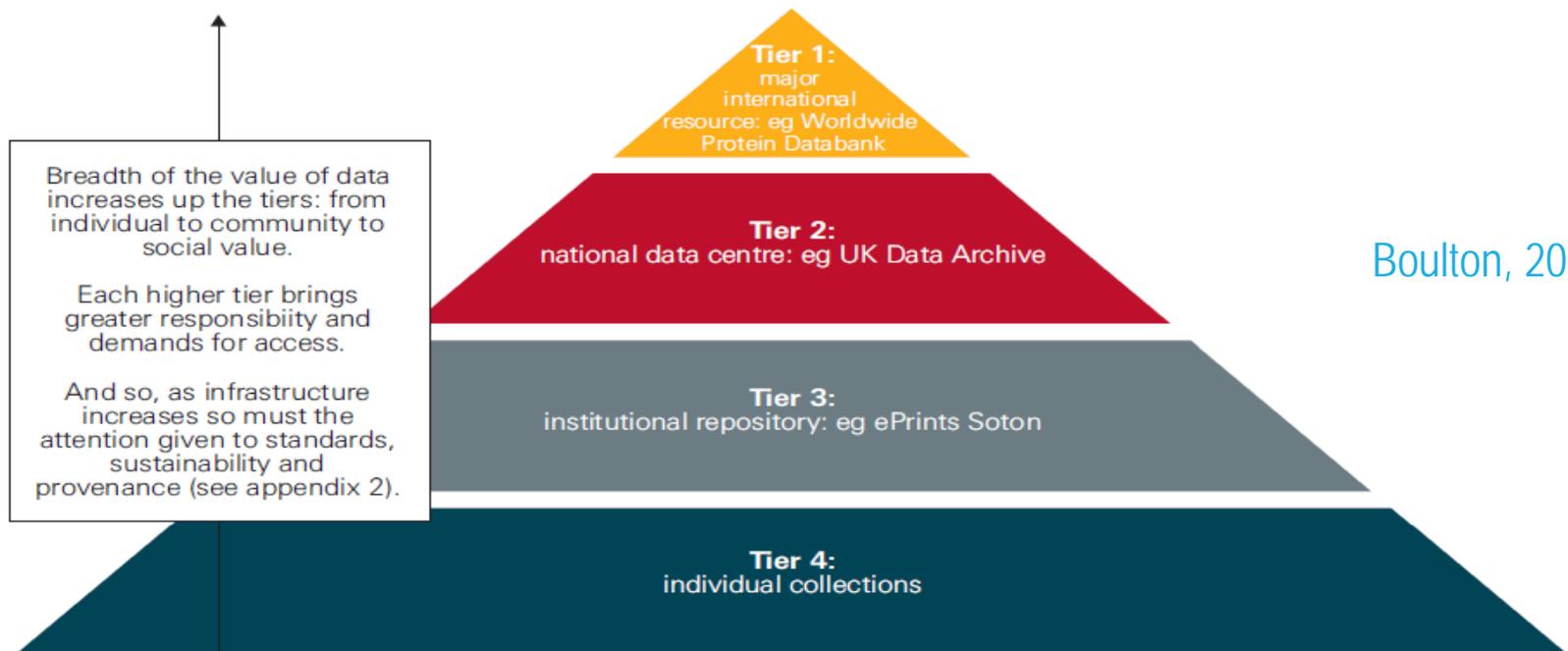
No existe consenso en cómo preservar y compartir los datos, y parece que las soluciones mixtas entre los servicios universitarios y los externos

Garantizar la procedencia, la persistencia de los identificadores, la utilización de estándares para los metadatos y formatos de citación y periodos de embargo, así como aspectos técnicos derivados de las posibilidades de recombinarlos, analizarlos automáticamente, migrarlos de un formato a otro, o los métodos de almacenamiento

Herramientas: especializadas y huérfanos

Figshare, Dryad, Dataverse o Zenodo

Pisos desiguales:



Consejo....

Monash University “adopta, adapta y en último término desarrolla”

“Developing a new product may be expensive, costly to support, and could split researchers from their community.”

La lealtad de los científicos con su comunidad es mayor que con su institución.

1. El por qué de mi presencia aquí: ¿Qué hace DATASEA?
2. ¿De qué estamos hablando? OD / RD y lo abierto
3. Beneficios openscience
4. Entender qué hace un científico en su carrera
5. Infraestructura social
6. ¿Qué están haciendo con los datos de sus proyectos? En España y más allá
7. Un poco de alfabetización de datos: ¿Cómo *deberían* proceder?
8. A modo de conclusión

- Desechar una cultura donde se contemplan los datos como un bien **privado**
- Aumentar los criterios usados en la **evaluación** de la carrera de los investigadores para reconocer el esfuerzo que supone la liberación de datos útiles y las nuevas formas de colaboración (Leiden Manifesto; *San Francisco Declaration on Research Assessment, Putting science into the assessment of research*)
- Desarrollar **estándares** comunes para comunicar esos datos, también incluyendo criterios de calidad de los datos y cuestiones relacionadas con las licencias y derechos

- Ordenar **mandatos** que promuevan la liberación de los datos relevantes de las publicaciones científicas
- Reforzar el número de **científicos de datos** que son necesarios para gestionar y apoyar el uso de los datos digitales (lo que es también crucial para el sector privado de análisis de los datos y las estrategias open data de los gobiernos), sobre todo en las tareas de evaluación, que garantizan la calidad de los datos
- Desarrollar y utilizar nuevas herramientas de **software** para automatizar y simplificar la creación y explotación de conjuntos de datos, proporcionando un acceso uniforme a la variedad de *outputs* de la investigación

Nada más

¿Cómo lo ven ustedes?

<http://www.datasea.es>

mpesetm@upv.es en nombre del

Grupo DataSEA

Eventos nacionales de mayo

CRECS Murcia Peset, F.; González, LM.; Villamón, M.; Ferrer-Sapena, A.; Aleixandre-Benavent, R.; Rodríguez-Gairín, JM.; García-García, A (2015). La vía dorada de la publicación de datos. <http://www.thinkepi.net/crecs2015>

FOSTER WINNING HORIZON2020 WITH OPENSOURCE
Madrid, Valencia, Gijón: integrar las herramientas openscience como parte del ciclo de vida de la investigación y armonizar los requisitos de los financiadores y de la Ley de la Ciencia española con los derechos de los autores y las editoriales científicas <https://docs.google.com/forms/d/1x3RSV6rkiwyVRirBLk4zqwc9zoju5bOp1rur0QiabrU/viewform?c=0&w=1>

FESABID Gijón. Cultura abierta: conocimiento compartido <http://www.fesabid.org>

Peset, F. (2014). Datasea.

<https://polimedia.upv.es/visor/?id=a48abcad-21bc-ab4a-ac6a-af987b3e7f1a>



A screenshot of a video player interface. The main content is a presentation slide for 'DATASEA - Research Open Data - Open Research Data'. The slide lists 'Best methods for sharing and presenting research data' and 'OPEN DATA' with sub-points like 'Health', 'High Cost', and 'Clean Access'. It also mentions 'Funded by research projects in grants' and 'Operational Science'. A woman is visible in the bottom right corner of the video frame, standing next to the slide. The video player has a 'poliMedia' logo in the bottom left and a university logo in the bottom right.

Algo de bibliografía

Parse insight project. www.parse-insight.eu.

DATARES project <http://digital.library.unt.edu/ark:/67531/metadc228265/m/>

Mossink, Wilma; Bijsterbosch, Magchiel & Nortier, Joeri (2013). *European Landscape Study of Research Data Management*. Utrecht: SURF, 2013

Council on Library and Information Resources (2013). *Research Data Management: Principles, Practices, and Prospects*. Washington, D.C.: CLIR. <http://www.clir.org/pubs/reports/pub160/pub160.pdf>

Beagrie, N. & Houghton, J. (2013). *The value and impact of the Archaeologydata service*. Charles Beagrie Ltd

Beagrie, N. & Houghton, J. (2014). *The Value and Impact of Data Sharing and Curation A synthesis of three recent studies of UK research data centres*