Què fer amb les dades que genera una investigació? Una visió pràctica.

Fernanda Peset. Universitat Politècnica de València

Palma de Mallorca, 9/11/18 10.15 h.

Evolución

2000	2003	2010	2011	2015	2017	2018
Comunicación científica	Acceso abierto	Datos enlazados	Datos de Investigación	Open Data	Measuring data	Blockchain Science/data

e-LiS e-prints in library & information science

VESTIGIUM BIBLIOTECA DIGITAL DE CIENTÍFICOS Y HUMANISTAS VALENCIANOS

FOTOGRAFÍA SOBRE ESPAÑA EN EL SIGLO XIX del Fondo Fotográfico de la Universidad de Navarra

IRALIS® International Registry for Authors: Links to Identify Scientists

DATASEA DATOS ABIERTOS DE INVESTIGACIÓN

DiSEA International registry on Research Data

MareData

UNIVERSITAT POLITÈCNICA DE VALÈNCIA · IUMPA · MADΦ

Scie-Chain

TABOOW ICOSYSTEM GUARDIAN

MADΦ MATHEMATICS & APPLICATIONS TO DATA & PHYSICS

UNIVERSITAT POLITÈCNICA DE VALÈNCIA · IUMPA · MADΦ · TABOOW ICOSYSTEM GUARDIAN · Scie-Chain

DATASEA DATOS ABIERTOS DE INVESTIGACIÓN

GOBIERNO DE ESPAÑA · MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

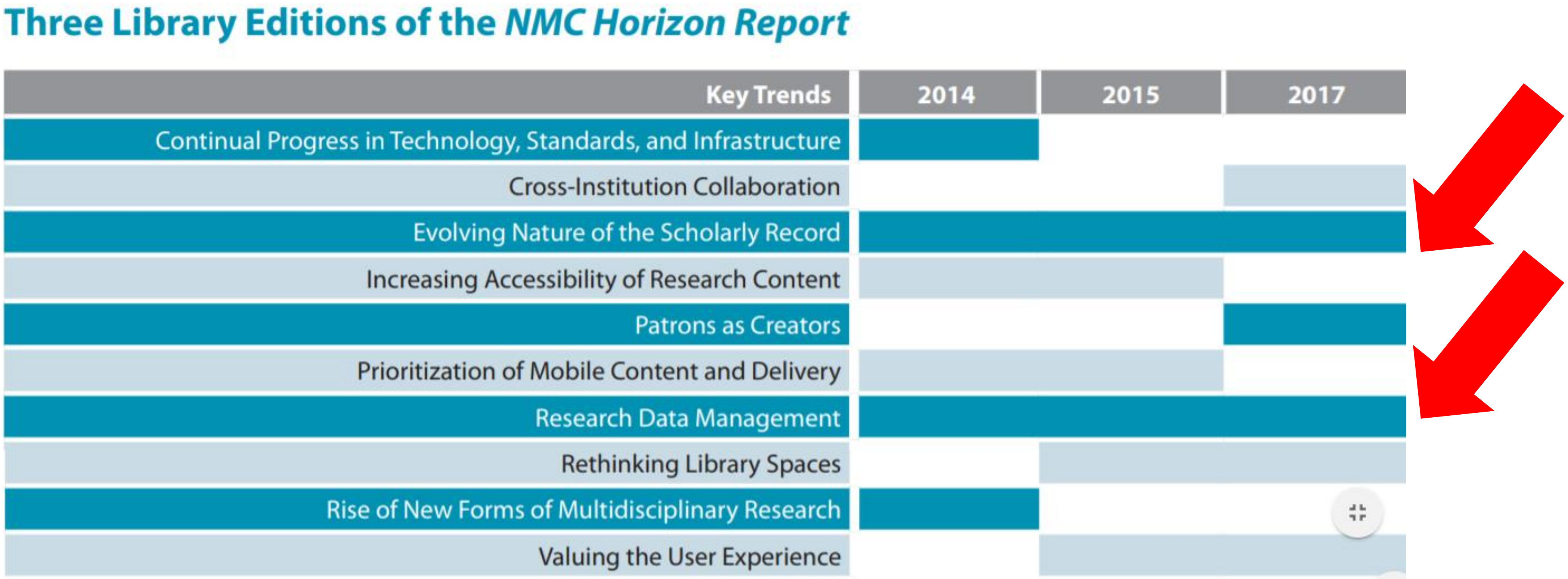
Hablamos de un ecosistema

- Los datos Importan
- Ecosistema científico innovador

Datos de investigación

- De qué hablamos: definiciones prácticas
- Selección
- Gestionar datos de investigación
  - Solicitud
  - Proyecto activo
  - Publicación

Conclusión



Horizon Report > 2017 Library Edition http://cdn.nmc.org/media/2017-nmc-horizon-report-library-EN.pdf

Torn, Pep (2016). Las tendencias de las bibliotecas académicas y de investigación del próximo lustro. Horizon Report: 2015 Library Edition. http://www.ub.edu/blokdebid/es/content/las-tendencias-de-las-bibliotecas-academicas-y-de-investigacion-del-proximo-lustro-horizon

https://ec.europa.eu/research/openscience/pdf/integrated_advice_opspp_recommendations.pdf#view=fit&pagemode=none

Para no ser yo quien lo diga…

The systemic change in the modus operandi of science and research…

- Exponential growth of data.
- Availability of **digital** technologies.
- Increase of the **global** scientific population.
- Public demand for better and more **efficient** science.

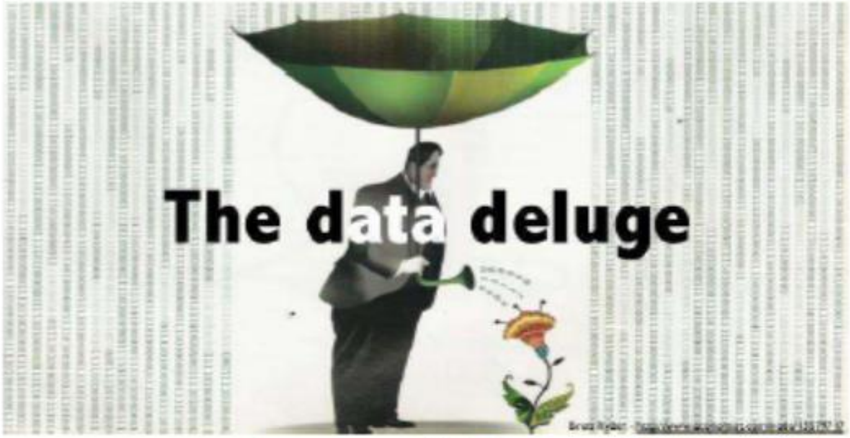


Image appeared at *The Economist* (March, 2010)

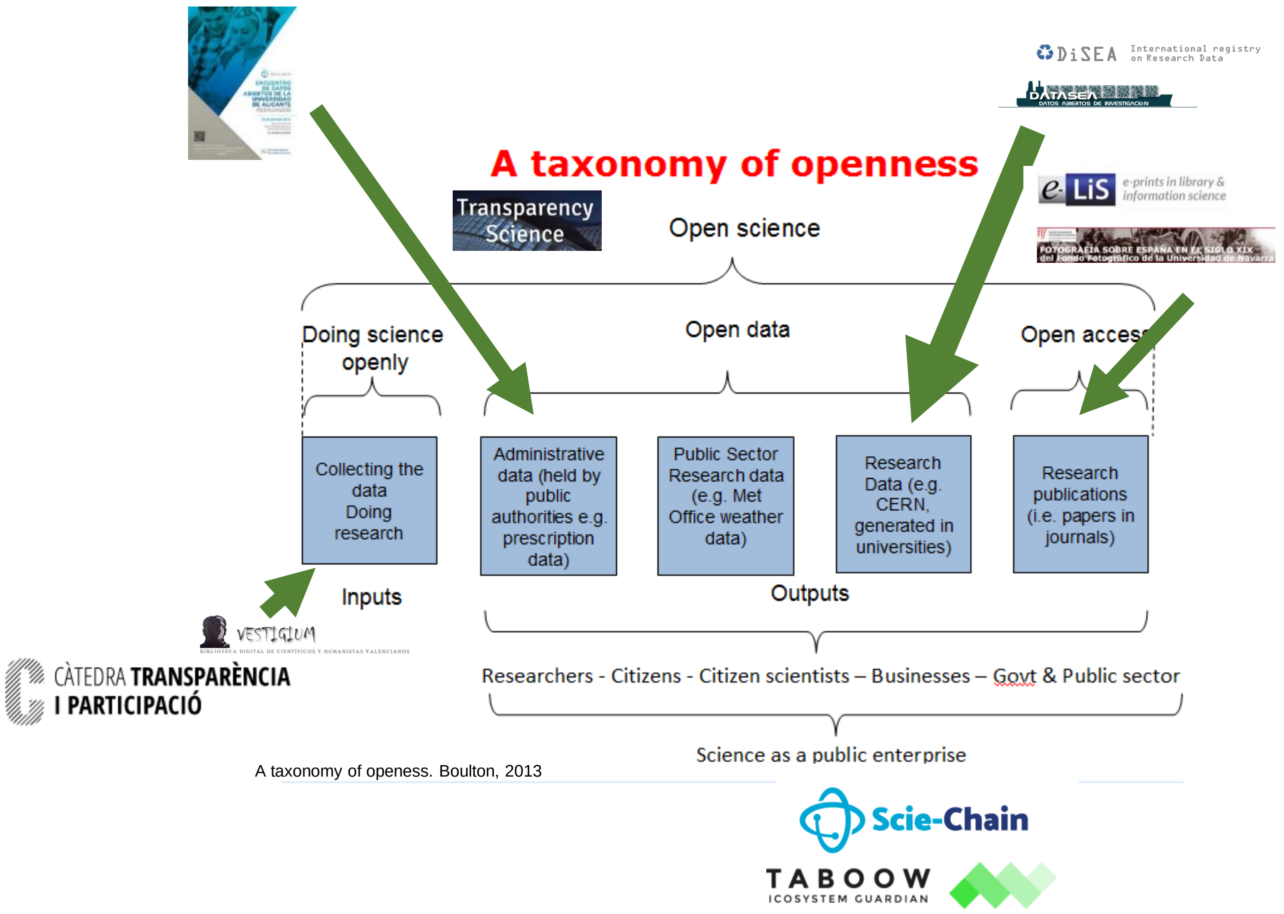
Eva Méndez, chair OSPolicyPlatform (EU)

- Demand for **accountable**, responsive and **transparent** science (PSI).
- Need to address faster **societal challenges**, and to contribute to economic growth.

TABOOW ICOSYSTEM GUARDIAN

Scie-Chain

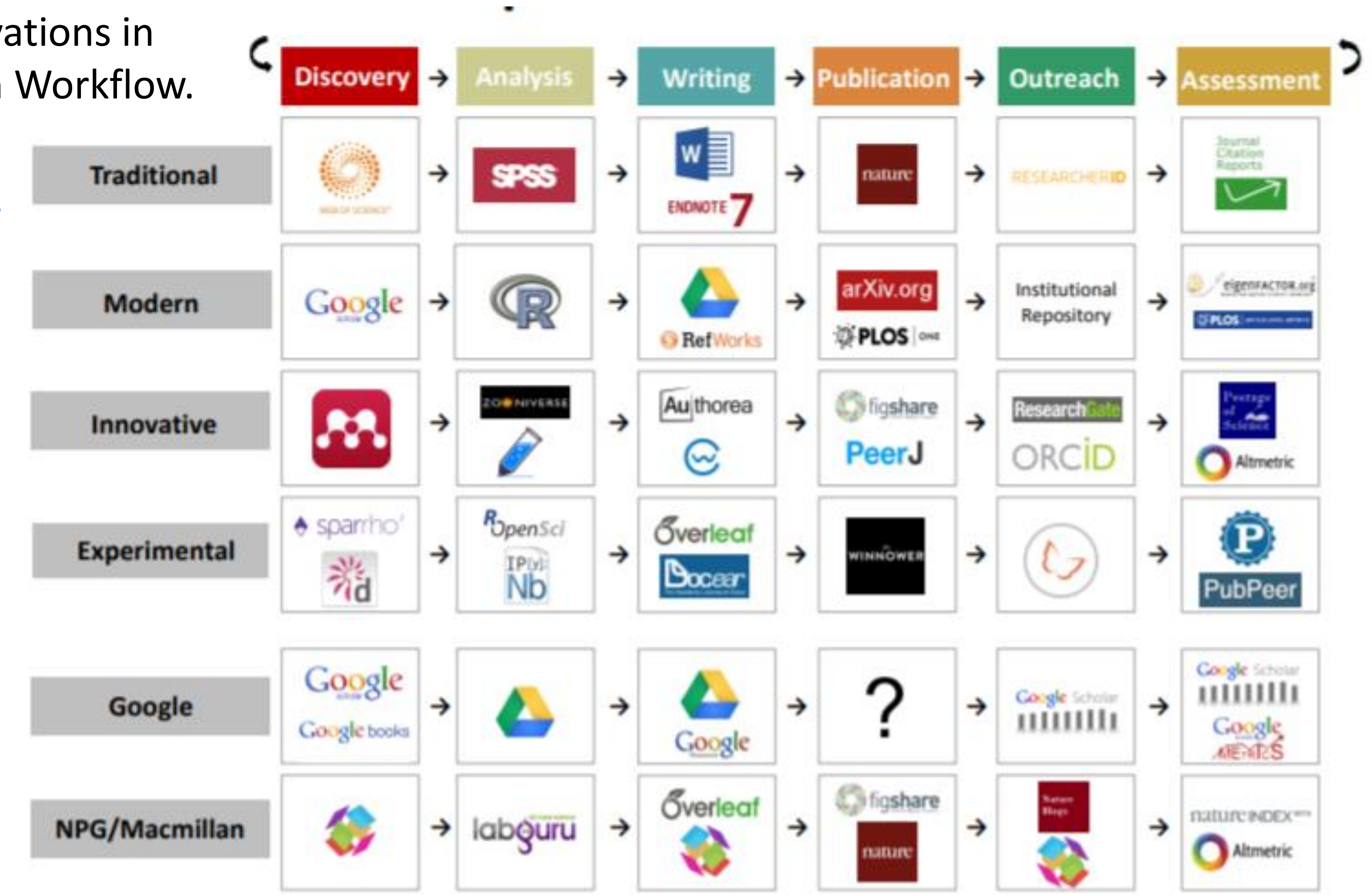
Ecosistema

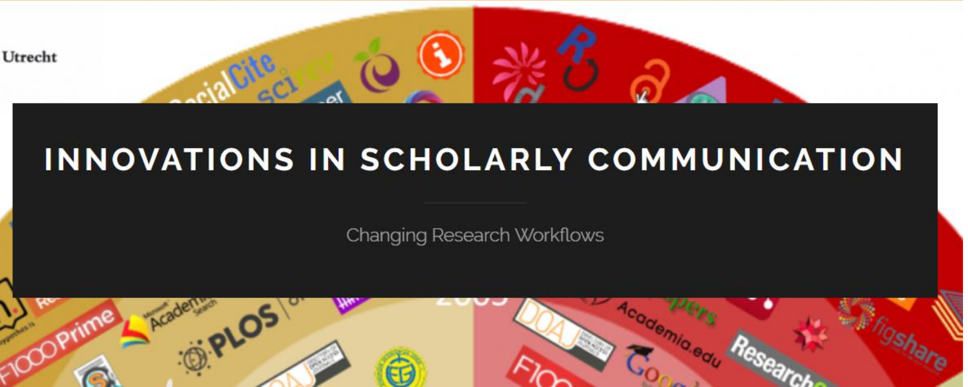


A taxonomy of openess. Boulton, 2013

https://docs.google.com/spreadsheets/d/1KUMSeq_Pzp4KveZ7pb5rddcssk1XBTiLHniD0d3nDqo/edit#gid=0

Kramer, Bianca; Bosman, Jeroen (2015): 101 Innovations in Scholarly Communication - the Changing Research Workflow. figshare. Poster.
https://doi.org/10.6084/m9.figshare.1286826.v1





Hablamos de un ecosistema

- Los datos Importan
- Ecosistema científico innovador

Datos de investigación

- De qué hablamos: definiciones prácticas
- Selección
- Gestionar datos de investigación
  - Solicitud
  - Proyecto activo
  - Publicación

Conclusión

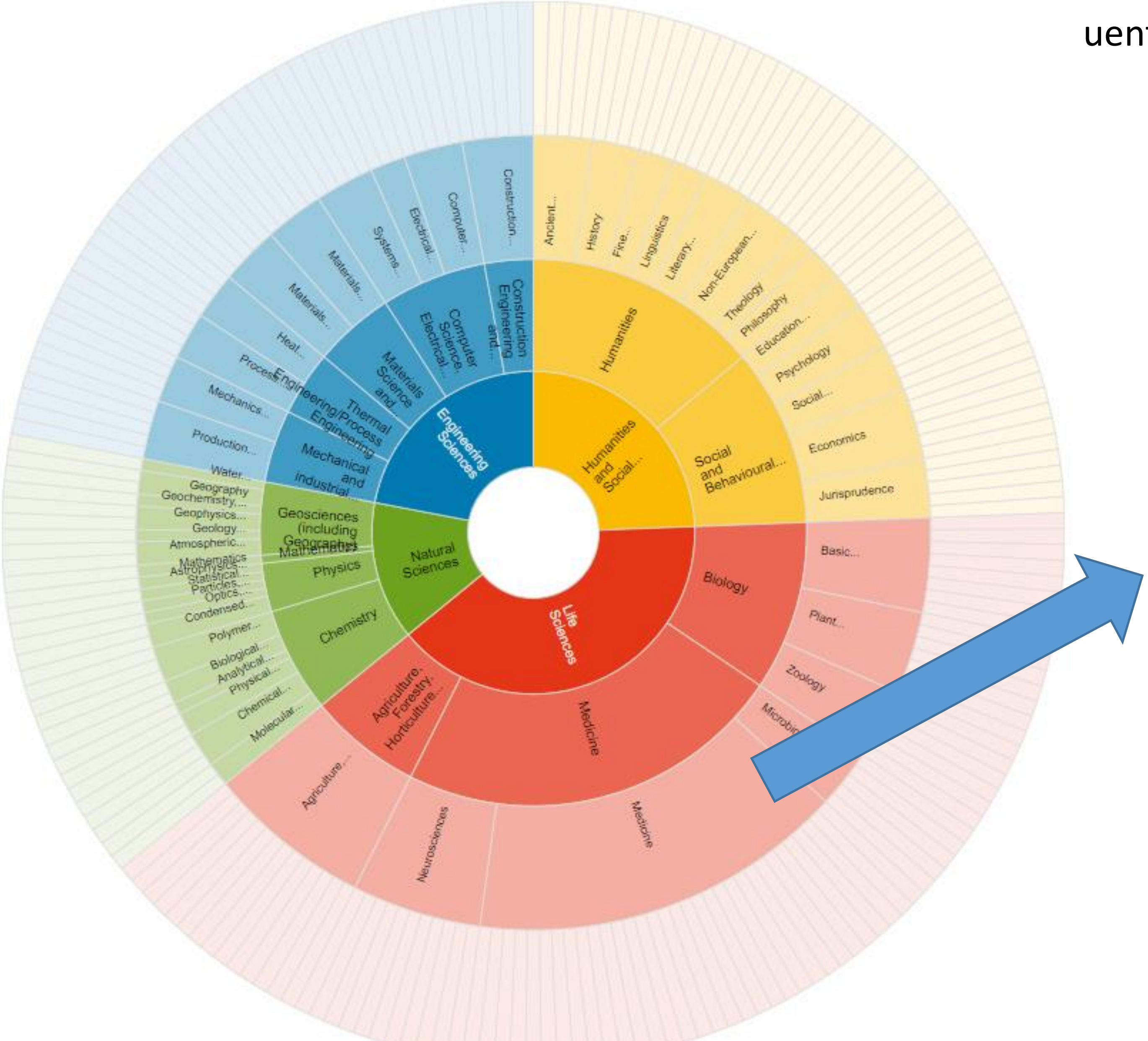
Definiendo…

“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 **fuente única** y deben ser **difíciles o imposibles de obtener de nuevo**” (Torres-Salinas; Robinson-García; Cabezas-Clavijo, 2012)

Necesarios para **validar** los resultados de una investigación: “…factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated.” OECD (2007)

Digital Repositories, INVEMAR,, Nov 2017

Bancos de datos en la categoría Medicine por submaterias.
uente: re3data

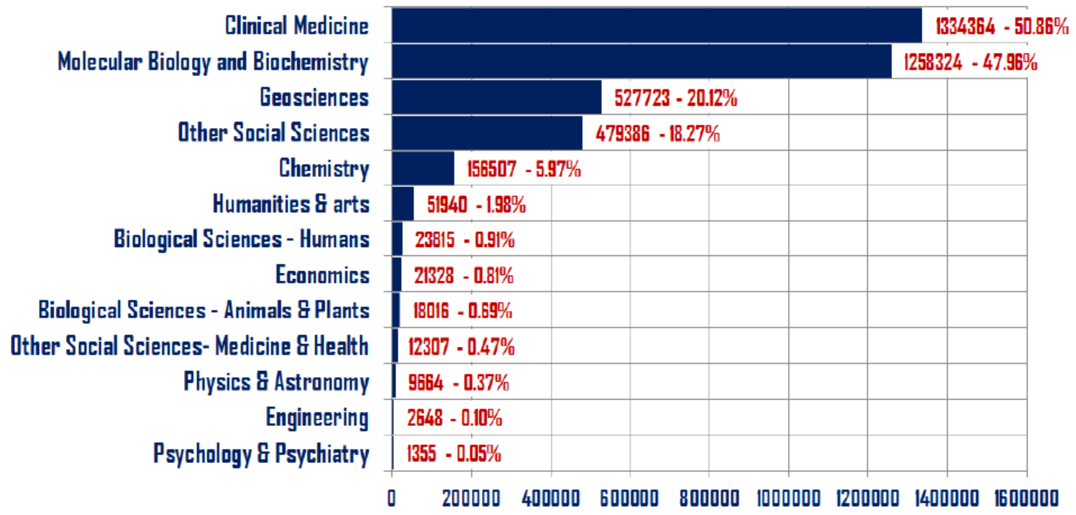


Categoría	Nº total
Public Health, Health Services Research, Social Medicine	99
Human Genetics	96
Epidemiology, Medical Biometry, Medical Informatics	40
Pharmacology	16
Biomedical Technology and Medical Physics	15
Physiology	12
Hematology, Oncology, Transfusion Medicine	11
Pharmacy	10
Toxicology and Occupational Medicine	10
Gastroenterology, Metabolism	7
Pediatric and Adolescent Medicine	7
Pathology and Forensic Medicine	6
Radiology and Nuclear Medicine	6
Gynaecology and Obstetrics	4
Clinical Chemistry and Pathobiochemistry	3
Pneumology, Clinical Infectiology Intensive Care Medicine	3
Gerontology and Geriatric Medicine	3
Radiation Oncology and Radiobiology	3
Nutritional Sciences	2
Endocrinology, Diabetology	2
Rheumatology, Clinical Immunology, Allergology	2
Cardiology, Angiology	1
Dermatology	1
Urology	1
Total Medicina	301

Datos en salud. Fuente: DCI

Robinson, Jiménez-Contreras y Torres-Salinas, 2015

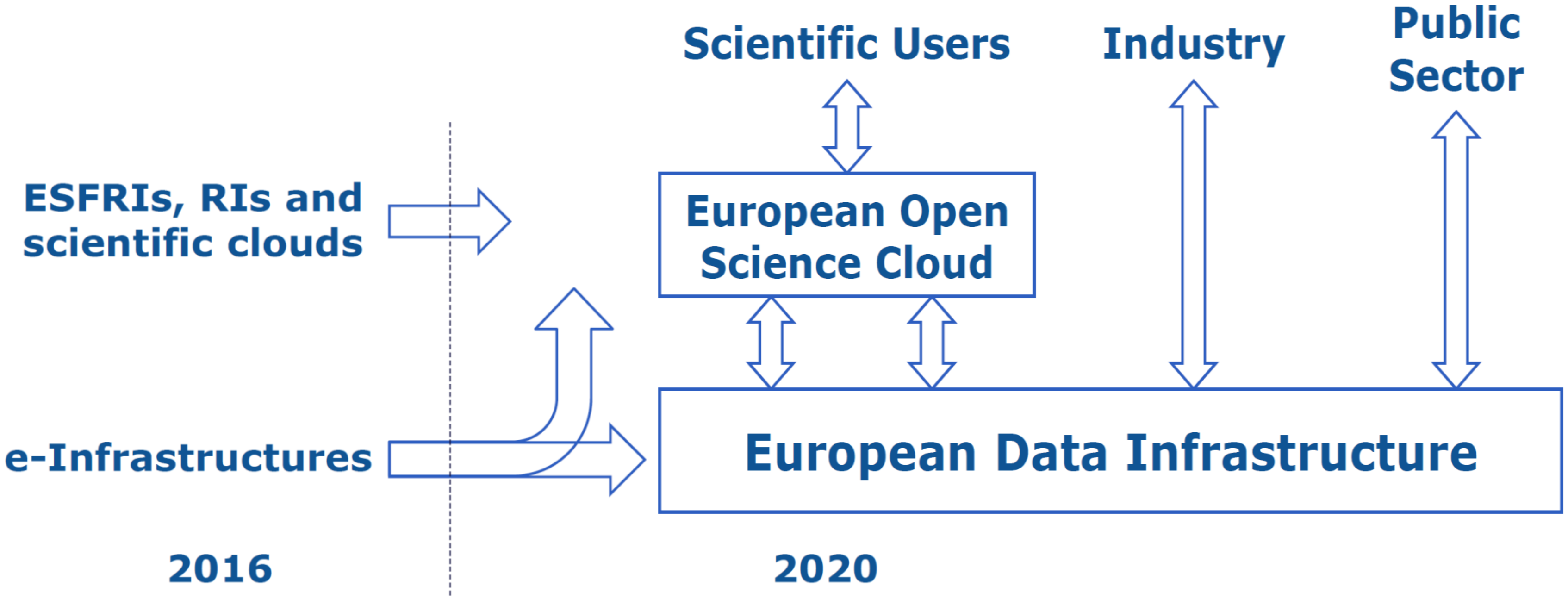
Figure 3
Record distribution per discipline in the Data Citation Index



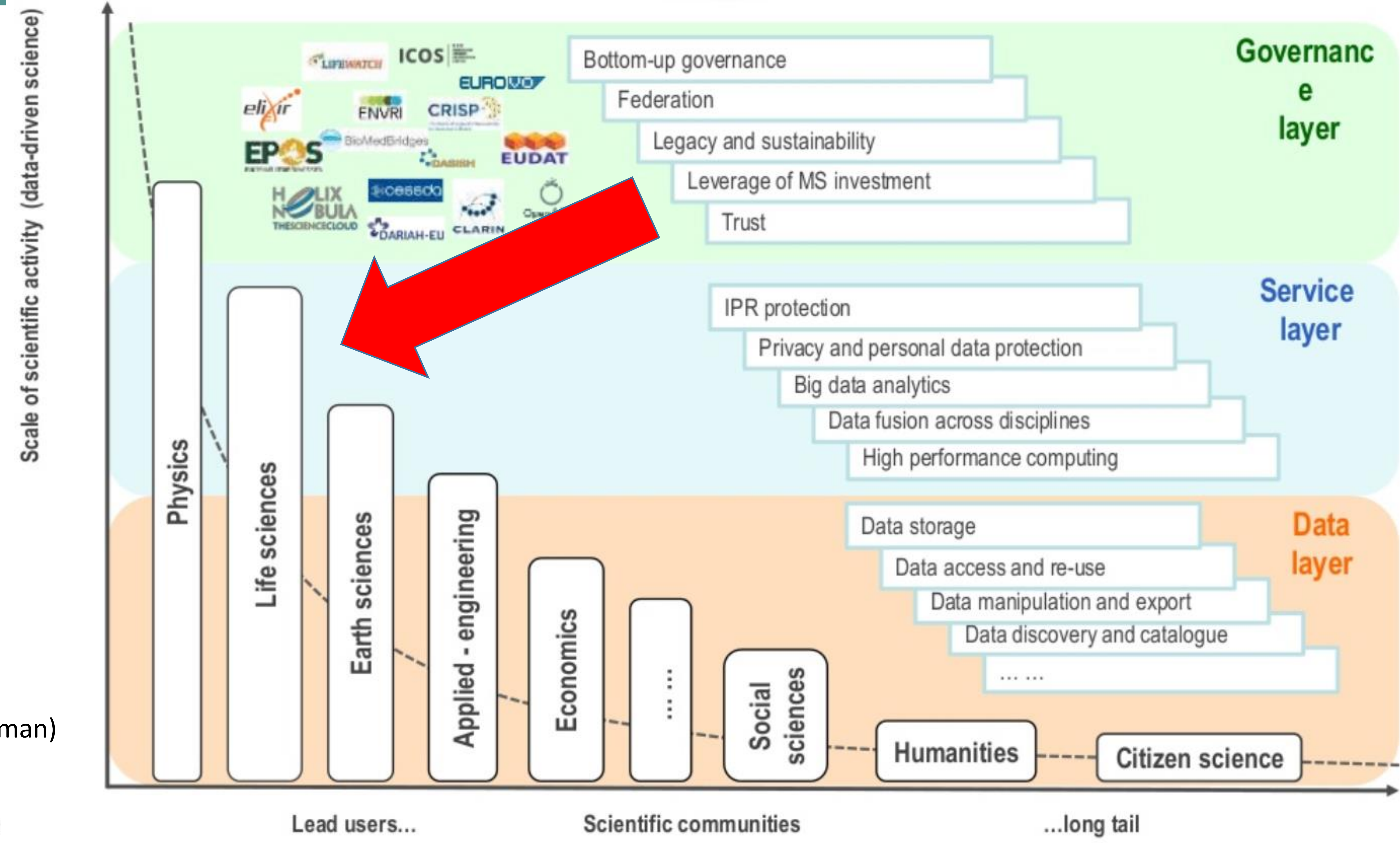
DISCIPLINA	Nº de bancos DCI 10/10/2018
Life Sciences	144
Social Sciences	82
Physical Sciences	74
Multi-discipline	42
Arts & Humanities	9
Technology	7
Total general	362

Hay muchos!!!

Evolución de la infraestructura



https://eoscpilot.eu/about/eoscpilot-brief



Scale of scientific activity (data-driven science)
Governance layer
Bottom-up governance
Federation
Legacy and sustainability
Leverage of MS investment
Trust
Service layer
IPR protection
Privacy and personal data protection
Big data analytics
Data fusion across disciplines
High performance computing
Data layer
Data storage
Data access and re-use
Data manipulation and export
Data discovery and catalogue
... ...
Physics
Life sciences
Earth sciences
Applied - engineering
Economics
... ...
Social sciences
Humanities
Citizen science
Lead users...
Scientific communities
...long tail

(Burgelman)

FAIR PRINCIPLES

Crédito: https://www.dtls.nl/fair-data/fair-data/

Interoperabilidad

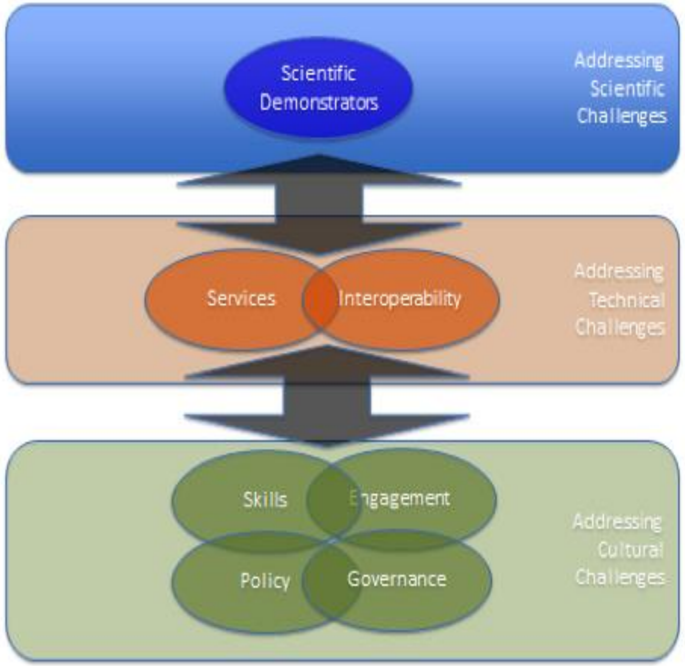
Enfoque interdisciplinar

Social y cultural

El cambio cultural hará más abierta la ciencia

Infraestructura

El acceso a recursos compartidos cumple la promesa de hacer una ciencia abierta intensiva en datos



UNIVERSITAT POLITÈCNICA DE VALÈNCIA IUMPA TABOOW ICOSYSTEM GUARDIAN Scie-Chain

Hablamos de un ecosistema

- Los datos Importan
- Ecosistema científico innovador

Datos de investigación

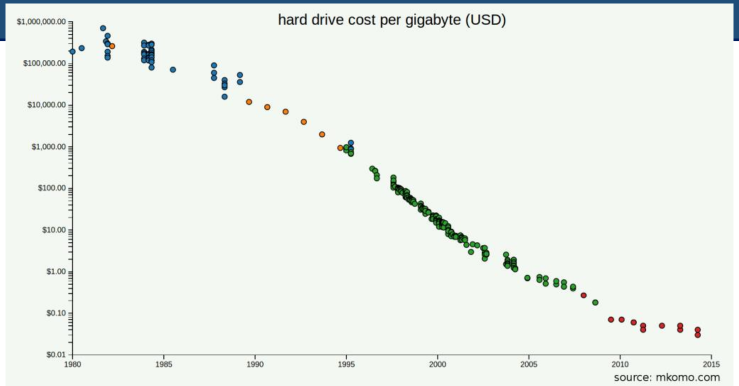
- De qué hablamos: definiciones prácticas
- Selección
- Gestionar datos de investigación
  - Solicitud
  - Proyecto activo
  - Publicación

Conclusión

Seleccionar

Preservar

Previamente, expurgar



[Matt Komorowski](http://www.mkomo.com/cost-per-gigabyte-update) http://www.mkomo.com/cost-per-gigabyte-update

- La selección toma tiempo pero ahorra costes de preservación y limita los riesgos a la institución (obligaciones legales)
- La bajada de precios de almacenamiento local no corre pareja con los precios de la nube
- El servicio debe ayudar junto al equipo de investigación a identificar qué datasets van a tener valor a largo plazo (no demasiados) y cómo se van a preservar

Criterios DCC. *How to Select and Appraise Research Data for Curation*

1. Relevancia para la misión de la institución o financiador, incluyendo aquí si existen obligaciones legales (underlying data)
2. Valor científico o histórico, lo que indica que puede usarse en el futuro para la investigación o la formación
3. Unicidad y riesgo de pérdida
4. Potencial para la redistribución
5. Replicabilidad casi imposible
6. Costos de preservación razonables o justificables
7. documentación completa

http://www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services#Selection-criteria

- Hablamos de un ecosistema
  - Los datos Importan
  - Ecosistema científico innovador

- Datos de investigación
  - De qué hablamos: definiciones prácticas
  - Selección
  - Gestionar datos de investigación
    - Solicitud
    - Proyecto activo
    - Publicación

- Conclusión

LEARN Toolkit of Best Practice for Research Data Management (2017). http://learn-rdm.eu/wp-content/uploads/RDMToolkit.pdf DOI: https://doi.org/10.14324/000.learn.01

El ciclo de vida de los datos

http://www.datasea.es/dt/index.php/resultados/gestiona-y-comparte-datos-de-investigacion-guia-datasea-para-investigado

En la práctica

Escenarios posibles

- Queremos solicitar una ayuda financiera y nos piden un plan de gestión de datos ¿Qué es? ¿Cómo hacerlo? ¿Ejemplos?

- Estamos trabajando en un proyecto internacional y no se cuál es la mejor forma de permitir al equipo que accedan a los datos ¿Infraestructuras? ¿Organización? ¿Precauciones?

- He escrito un artículo financiado con un proyecto y nos obligan a depositar los datos ¿Todos? ¿Dónde? ¿Formatos?

http://www.datasea.es/dt/index.php/resultados/gestiona-y-comparte-datos-de-investigacion-guia-datasea-para-investigadores/presentacion

H2020

- If you will take part in the pilot on **Open Research Data**[2] , include information on how the participants will manage the research data generated and/or collected during the project, in particular addressing the following issues:[3]
    - What types of data will the project generate/collect?
    - What standards will be used?
    - How will this data be exploited and/or shared/made accessible for verification and re-use? If data cannot be made available, explain why.
    - How will this data be curated and preserved?

⚠ *You will need an appropriate consortium agreement to manage (amongst other things) the ownership and access to key knowledge (IPR, data etc.). These will allow you, collectively and individually, to pursue market opportunities arising from the project's results.*

- Outline the strategy for **knowledge management and protection**. Include measures to provide **open access** (free on-line access such as the 'green' or 'gold' model) to peer-reviewed scientific publications which might result from the project[4].

⚠ *Open access publishing (also called 'gold' open access) means that an article is immediately provided in open access mode by the scientific publisher. The associated costs are usually shifted away from readers, and instead (for example) to the university or research institute to which the researcher is affiliated, or to the funding agency supporting the research.*

⚠ *Self-archiving (also called 'green' open access) means that the published article or the final peer-reviewed manuscript is archived by the researcher - or a representative - in an online repository before, after or alongside its publication. Access to this article is often - but not necessarily - delayed ('embargo period') as some scientific publishers may wish to recoup their investment by selling subscriptions and charging pay-per-download/view fees during an exclusivity period.*

Formularios, 2. Impact. pag. 18. Se recomienda máx. 2 pags.

http://ec.europa.eu/research/participants/data/ref/h2020/call_ptef/pt/2016-2017/h2020-call-pt-eranet-2016-17_en.pdf

Ejemplos del DCC http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples

RIO (Research Ideas and Outcomes) is new online journal and its purpose is to share all documents related to the research lifecycle, including proposals. http://rio.pensoft.net/browse_user_collection_documents.php?collection_id=3&journal_id=17

DMPTool : planes públicos
https://dmptool.org/public_dmps?e=z&public%3Apage=1&s=a

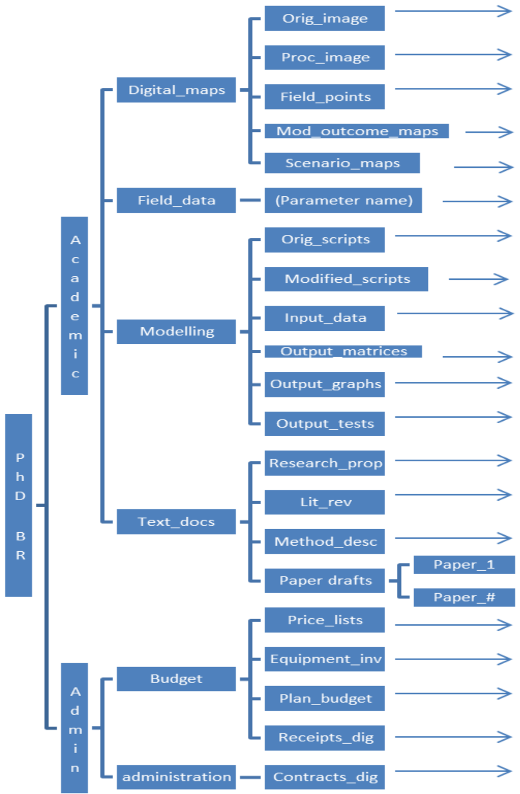
PGDonline PAGODA

http://www.datasea.es/dt/index.php/resultados/gestiona-y-comparte-datos-de-investigacion-guia-datasea-para-investigadores/presentacion

Organizar

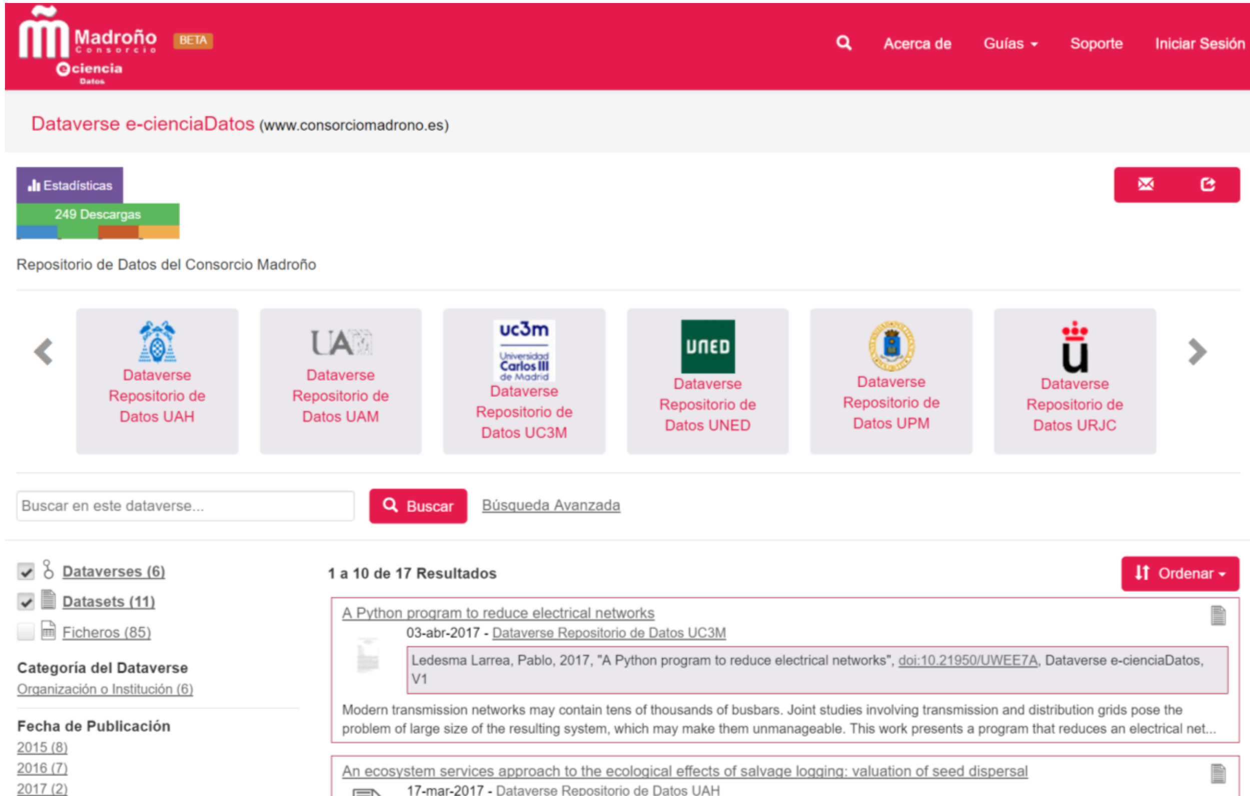
Gestionar

Wigham (2014)



Name	Example
Month_year_type	01_1989_landsat
Month_year_type_process	01_1989_landsat_orthorect
Field_points_date of actualization	Field_points_11_09_13
Month_year_outcome type	01_1989_wateryield
Year_parameter	2050_wateryield
Parameter_date of actualization	Rainfall_11_09_2013
Model_script_version	Shetran_script_v2.5
Model_script_modification_trial#	Shetran_script_cloudwater_t1
Parameter_col#_row#_trial#	Rainfall_4_250_t1
Est. parameter_col#_row#_trial#	ET_4_250_t1
Graph_x_y_trial#	Scatterplot_rainf_ET_t1
Test_parameter_trial#	Senstest_rainfall_t1
RP_trial# & _revised by	RP_t1_revLG
Topic_subtopic	Modelling_wateryields
Topic_method	Soil inf_infiltrometer
Paper code_date & revised by	TMCFvsPas_11_09_2014_LG
Company_date	Eijkelkamp_11_09_2013
Equipment list	Equipment_list
Budget_date	Budget_11_09_2013
Item_rec_date (month_year)	Mercado_rec__01_2014
Item_con_date	Houserent_con_01_2014

E-ciencia Datos. Fuente: https://edatos.consorciomadrono.es/dataverse/Madrono



http://www.datasea.es/dt/index.php/resultados/gestiona-y-comparte-datos-de-investigacion-guia-datasea-para-investigadores/presentacion

Aspectos legales

Retener la propiedad sobre los datos es una cuestión crítica en investigación

- Intellectual Property Rights (IPR) son inealineables al trabajo de los autores excepto que renuncien expresamente. Están limitados en el tiempo. Permiten publicar, licenciar la distribución y defenderlos

- Los datos gozan de una protección peculiar: sólo se aplica a los que hayan sido procesados de alguna manera. Esto significa que los *raw data* puros, los generados por los aparatos, tienen el tratamiento de texto plano y no pueden ser amparados en la ley de propiedad intelectual ni industrial españolas. Son intrínsecamente libres, abiertos.

CONSENTIMIENTO INFORMADO

Título del estudio de investigación

XXXXX

Investigadores principales

XXXXX

Objeto y antecedentes

XXXXX

Procedimientos

Si consiente participar en el estudio sucederá lo siguiente:

1. Tendrá que responder a una serie cuestionarios
2. Será entrevistado a lo largo de aproximadamente 1 h, tiempo durante el cual podrá ser grabada la conversación.
3. Las respuesta que usted proporcione a los investigadores serán empleadas de forma anónima para la confección de informes y documentos científicos.

Beneficios

XXXXX

Riesgos

No existen riesgos reseñables que se puedan derivar de su participación en este estudio.

Confidencialidad

Los resultados obtenidos de todas las pruebas sólo podrán utilizarse en el ámbito científico y académico. Con la excepción de esta revelación, toda la información se considerará confidencial.

Sus datos personales y las grabaciones que se realicen serán tratados de acuerdo con las leyes vigentes sobre protección de datos.

Preguntas

El personal investigador ha discutido esta información conmigo y se ha ofrecido a responder a todas mis preguntas.

Derecho a rehusar o abandonar

La participación en el estudio es voluntaria y soy consciente de que puedo abandonar en cualquier momento la investigación.

Consentimiento

Consiento en participar en este estudio. He recibido copia de este impreso y he tenido la oportunidad de leerlo.

Firma del participante y/o representante legal:

Fecha:

Firma del investigador:

Vías reconocidas

- Vía verde de los repositorios interoperables
- Vía de los bancos de datos, a veces en LOD, a veces OAI-PMH compliant y otras veces nuevos silos invisibles a las máquinas

- Vía dorada del material suplementario, en la que las bibliotecas pueden hacer algo aún

International registry on research data http://odisea.ciepi.org

- Cada una tiene ventajas e inconvenientes si se analiza desde la óptica de un gestor de la información, quien ha de tomar en consideración cuestiones de preservación, de visibilidad, de calidad de los datos, de interoperabilidad... Por ejemplo, la liberación de los datos de un grupo de investigación en su propio web es una opción numerosa y cómoda, pero deja esos datos casi invisibles a la comunidad internacional.

Vía dorada

- Ciertas revistas ofrecen su plataforma para subir el material con una mínima descripción y otras recomiendan repositorios que luego enlazan. La más reciente forma de publicar los datos son las *datajournals*, que incluyen *data papers* o artículos que solo describen los *datasets* y cómo han sido obtenidos, y no los resultados o las conclusiones que se derivan de ellos. O al contrario, convertir un artículo en un datapaper (Data in brief, Elsevier)

- Adjuntar los datos a la publicación facilita su descubrimiento y la correcta interpretación de los datos pero en la plataforma editorial existen límites de almacenamiento. Riesgo de su preservación y acceso: ahora parece que proporcionan acceso al material suplementario aún sin tener suscrita la revista

Ad hoc:

EVISE

Upload Files

Provide Additional Information

Review & Submit

Guide for Authors

Click Upload Files to select and upload submission files.

- Select a File Type for each submission file.
- Mandatory File Types are indicated in the drop-down list.
- The total size of your submission files may not exceed 700MB.
- The Manuscript File size may not exceed 150MB.
- Update the File Order if necessary, then click Save to preserve the new order before proceeding.
- **Please note that in order to proceed with your submission (Only for Review Articles and Research Papers) you are required to to upload the author contributions.**
- Click Save & Continue to proceed to the next step.
- **Data in Brief:** Authors are invited to submit a Data in Brief article alongside their research artic it, which is normally buried in supplementary material, is actively reviewed, curated, formatted, lexed, given a DOI and publicly available new, Open Access journal, Data in Brief. Please use the Data in Brief template available through e below link: Data in Brief Template.
- **MethodsX (optional)**

  We invite you to submit a method article alongside your research article. This is an opportunity to get full your method article will be automatically transferred over to the open access journal, MethodsX, where it will be editorially reviewed and published as a sep available here when preparing your article: Open access fees apply. For any questions, please send your inquiries to mexjm@elsevier.com

CRediT Author Statement

We encourage you to submit an author statement file outlining all authors' individual contributions, using the relevant CRediT roles: Conceptualizatio Resources; Software; Supervision; Validation; Visualization; Roles/Writing – original draft; Writing – review & editing. Please format with author name

***Data Visualization**: Include interactive data visualizations in your publication and let your readers interact and engage more closely with your resea with your article.*

Enrich your article to present your research with maximum impact. This journal supports the following Content Innovations:

- Interactive Plots: Interactive plot viewer providing easy access to the data behind plots. Please prepare a .CSV file with your plot data and test

Upload Files | Order Files

* File Order	* File Name	* File Type
No data to display.		

Share your research data (Optional)

In this section, you can make your research data available with your article. Sharing research data helps other researchers evaluate your findings, and increases t protocols, methods and materials.

Please cite your research data within your article; consult your journal□s Guide for Authors for instructions.

The following options are available:

Link research data

If your research data is already hosted in a repository, you can link it to your article here. Learn more

PLOS

PLOS | ONE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

An author keyword analysis for mapp

Luis-Millán González, Xavier García-Massó, Alberto Pardo-Ibañez, Fernanda

Published: August 1, 2018 • https://doi.org/10.1371/journal.pone.0201435

Article | Authors | Metrics

Abstract

Introduction

Materials and methods

Results

Discussion

Conclusions

Supporting information

Abstract

Scientific production has increased expone methodological strategies for understandi trends developed in specific fields. Data m particular, our study presents a global ana in the Sport Sciences Category (SSC) incl the frequency of appearance and the dyna

NISO-NFAIS (Nat Fed of advanced inf systems)

Odisea

http://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor_en

NISO
How the information world CONNECTS

NFAIS
SERVING THE GLOBAL INFORMATION COMMUNITY

NISO RP-15-2013

Recommended Practices for Online Supplemental Journal Article Materials

Agricultural Science (22)
Biology and Chemistry (407)
Chemistry (80)
Clinical Medicine (1513)
Computer Science (91)
Economics and Bussines (155)
Engineering (113)
Environment Ecology (140)
Geoscience (77)
Immunology (86)
Material Science (29)
Mathematics (168)
Microbiology (138)
Molecular Biology and Genetics (451)
Multidisciplinary (60)
Neuroscience and Behaviour (207)
Pharmacology and Toxicology (147)
Physics (63)
Plant and Animal Science (278)
Psychiatry / Psychology (113)
Social Science General (41)
Space Science (13)

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

Proyecto financiado por Plan Nacional de I+D+i del Ministerio de Economía y Competitividad:

A.2.1 Primary Publishing

Publisher	Editor	Peer Reviewer	Author(s)
Educate other parties about requirements for posting and curating content.	Set editorial policy.	Follow journal guidelines for reviewing Supplemental Materials.	Be aware of Journal expectations and follow them to the best of their ability.
Provide appropriate resources for managing supplemental content.	Make final decisions on content.	Inform the editor in a timely fashion if unable to review any content.	Provide context and demonstrate that the Supplemental Materials add substance to scholarship in the field.
Provide systems and policies to facilitate the decision-making process.	Determine whether supplemental content is integral to the article.[3]	Alert the editor to instances in which integral data are not provided, but are needed to understand the manuscript.	Be responsible for providing Supplemental Materials at the same level of quality as the article.
Be clear about the level of delivery and preservation that can be provided.	Set expectations for acceptable content with an understanding of what is entailed in vetting, delivering, and preserving content.		Be aware of trusted repositories in the field and knowledgeable about their practices.
Encourage authors to post Additional Content in endorsed archives that ensure good preservation and provide bidirectional linking to the journal.[4]	Encourage authors to post Additional Content in endorsed archives that ensure good preservation and provide bidirectional linking to the journal.[4]		Be aware of and adhere to policies of your institution and funder for sharing of research data.

Roles de las bibliotecas en la vía dorada (NISO, 2013) https://www.niso.org/standards-committees/supplemental

Bibliotecas	Servicios de indización y resumen	Administradores de repositorios
Ser repositorio del material de investigación generado por su personal	Indicar la disponibilidad de Material suplementario si el editor de la revista da indicaciones claras sobre ello	Hacer accesible el contenido depositado asignando identificadores persistentes como DOI y otros
Incluir Material suplementario con el artículo que se suministra por préstamo bibliotecario si el editor de la revista da indicaciones claras sobre ello	Incluir los formatos y tipos de archivos de Material suplementario si el editor de la revista da indicaciones claras sobre ello	Incluir DOIs para cualquier artículo que enlace el contenido en su repositorio
	Incluir el DOI u otro identificador proporcionado por el editor	Gestionar enlaces bidireccionales entre el contenido depositado y el artículo

Vía verde de los Repositorios
(que no infraestructuras de investigación)

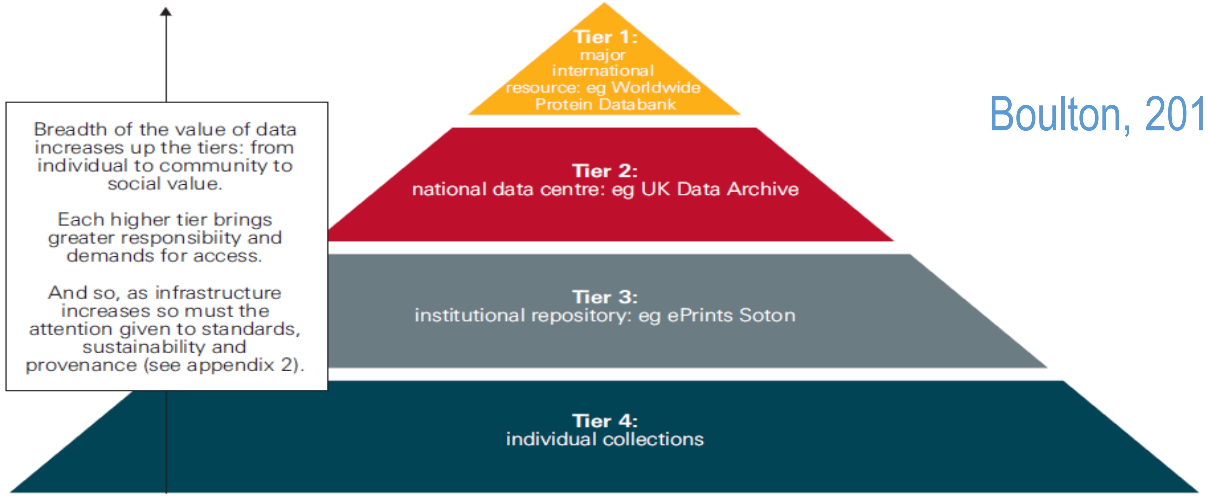
Funciones

- Cuando hablamos de repositorios de datos nos referimos a una muy amplia gama de infraestructuras donde se deposita la versión final de algo.
- Es decir, prima la preservación frente a la evolución (frente al ciclo de vida)
- Características:
  - Preservación supuestamente a largo plazo, por lo que suelen ser públicos
  - Interoperabilidad de alguna de las muchas formas (OAI-PMH), por lo que NO son modelos de negocio (RG o Academia)
  - Ofrecen un conjunto de servicios asociados, como métricas, CV…

Vía de los repositorios
Herramientas: especializadas y huérfanos

- Figshare, Dryad, Dataverse o Zenodo
- Pangaea https://www.pangaea.de/
- DCI master list http://wokinfo.com/product

- Pisos desiguales



Boulton, 2013

Consejos prácticos

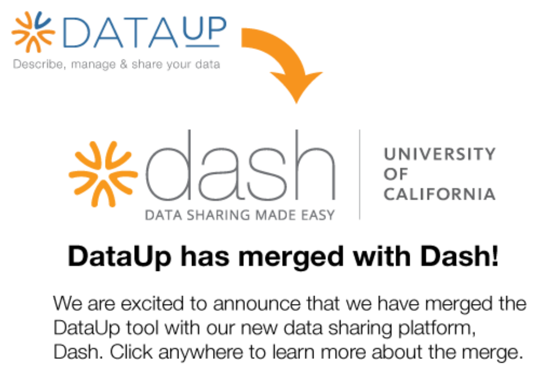
Consejo 1

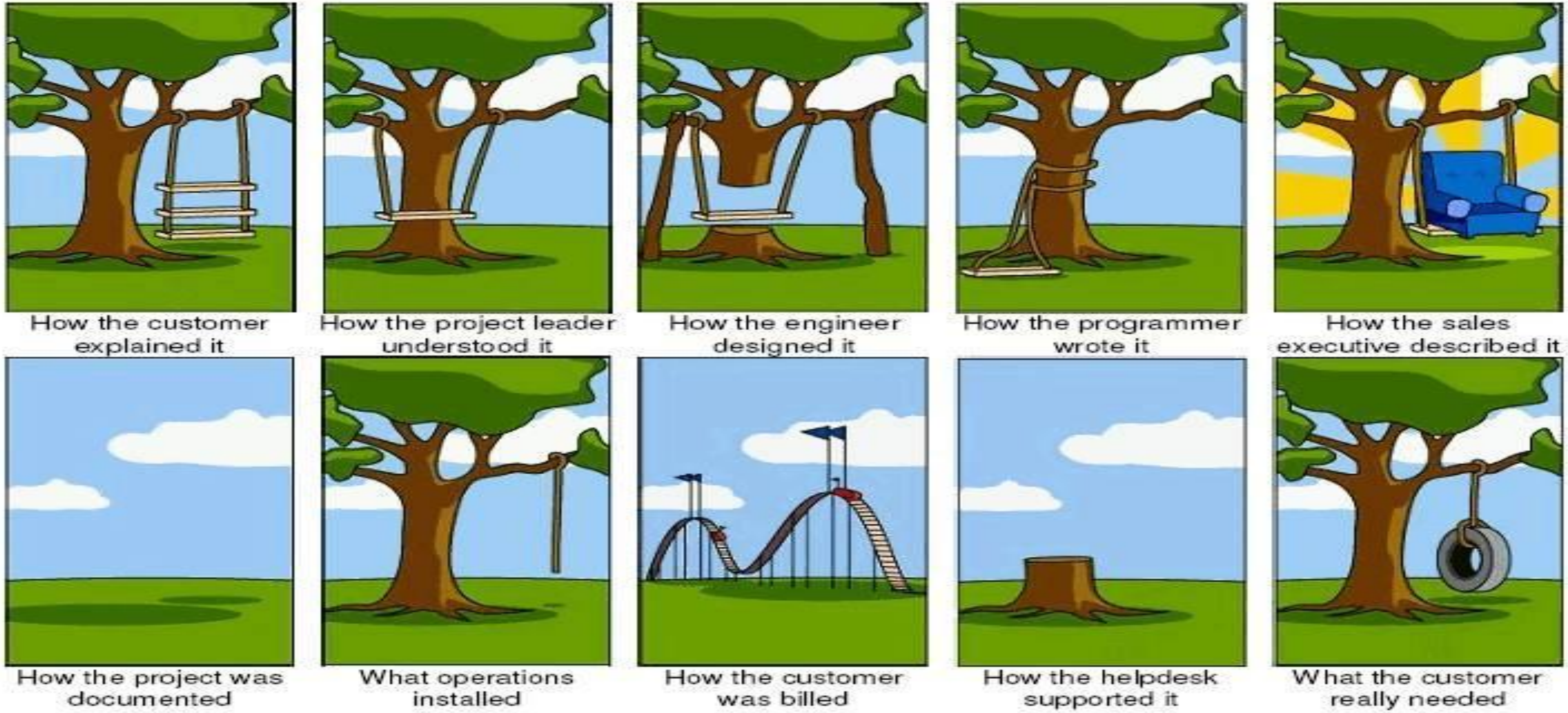
“adopta, adapta y en último término desarrolla” Monash University

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

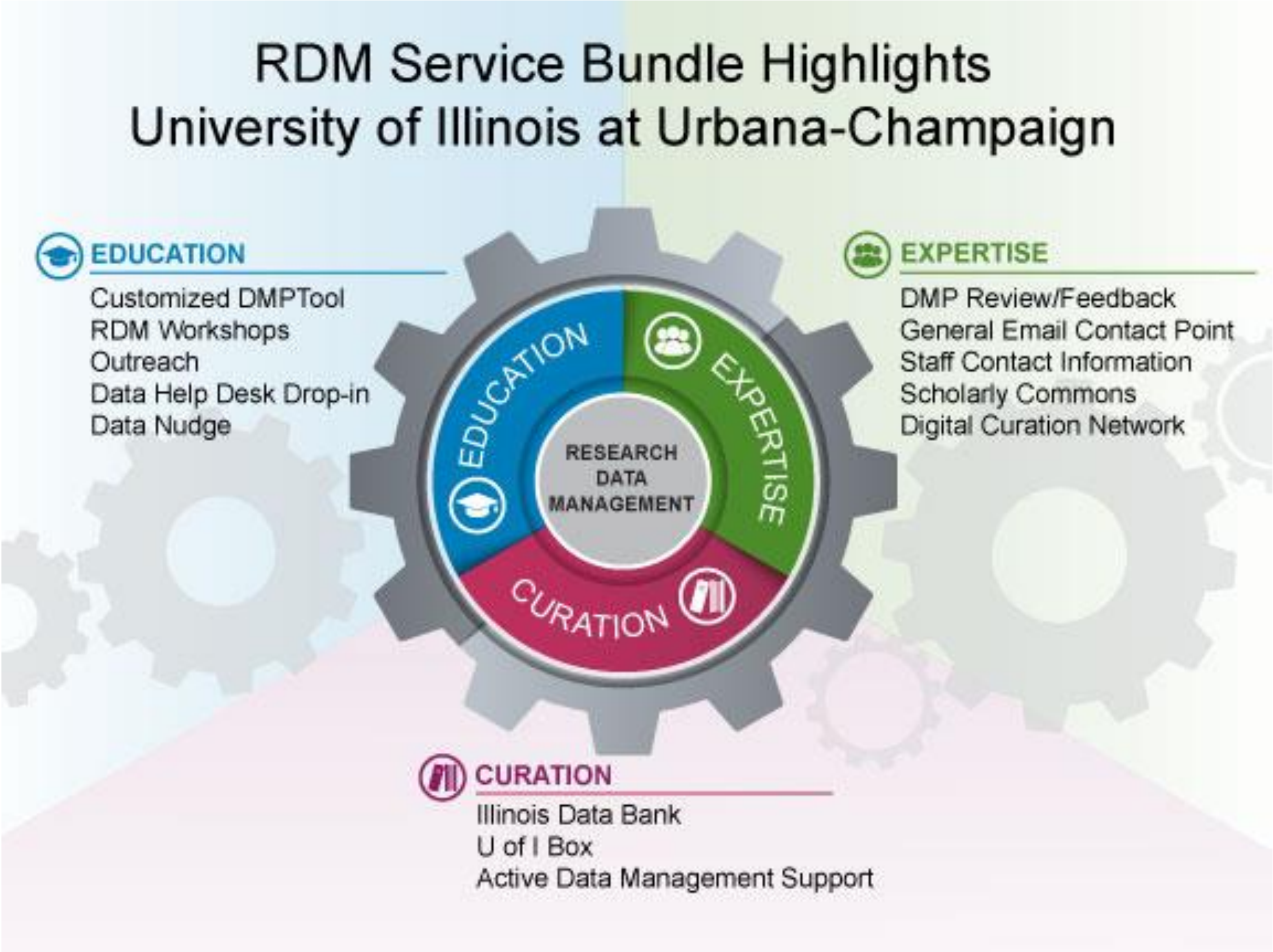
La lealtad de los científicos con su comunidad es mayor que con su institución.

Integra

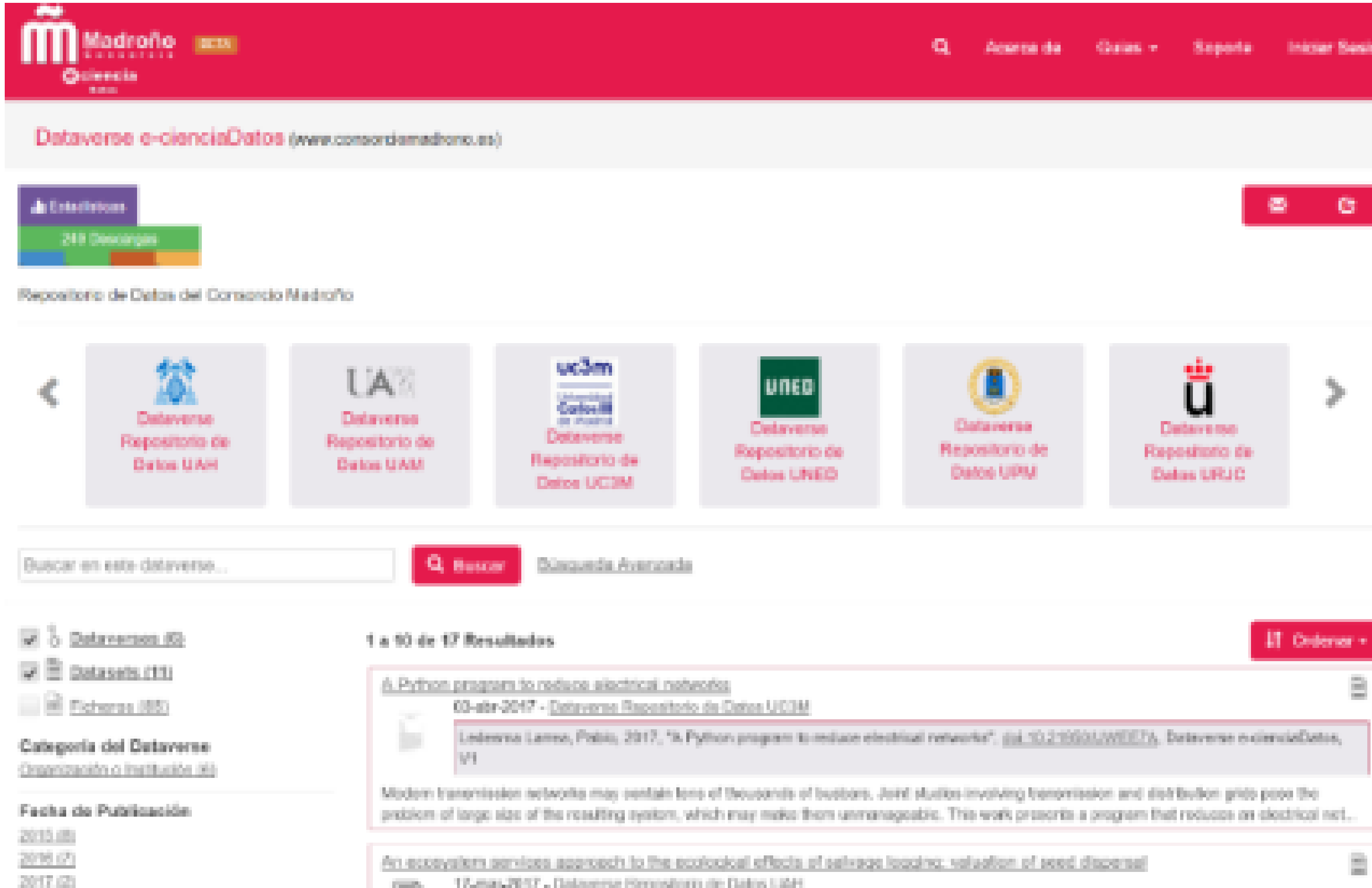



How the customer explained it
How the project leader understood it
How the engineer designed it
How the programmer wrote it
How the sales executive described it
How the project was documented
What operations installed
How the customer was billed
How the helpdesk supported it
What the customer really needed

Consejo 2 ¿Qué quiere mi institución?



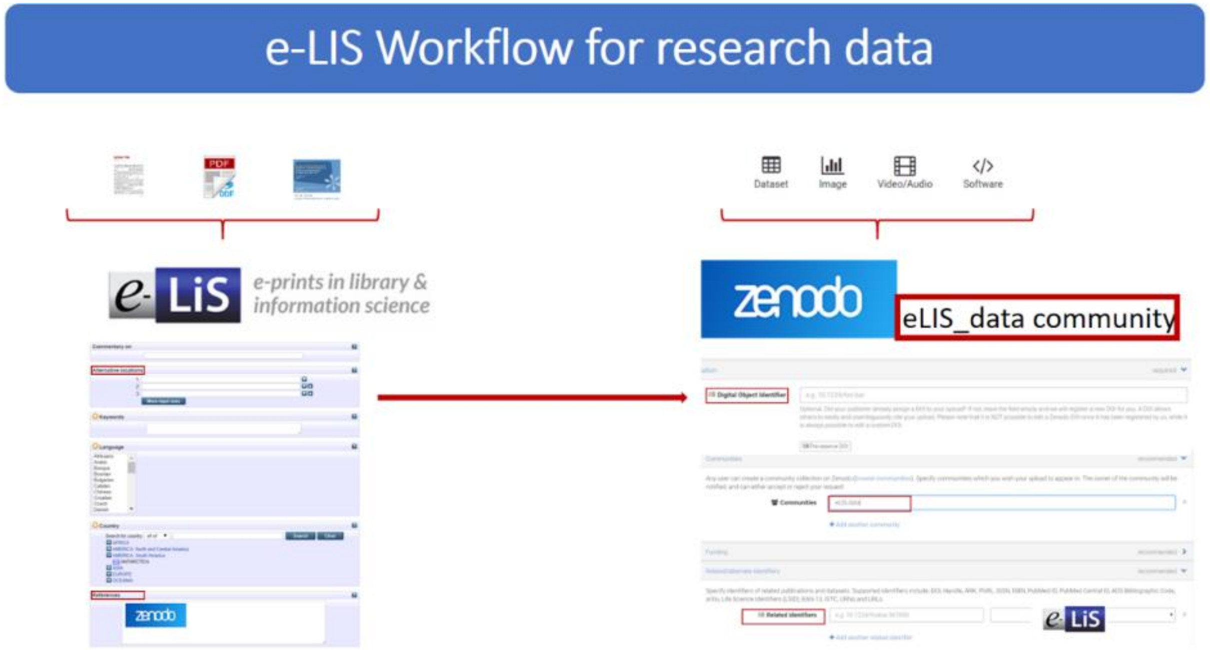
Si mi repositorio institucional/revista no acepta datos: opción DataVerse



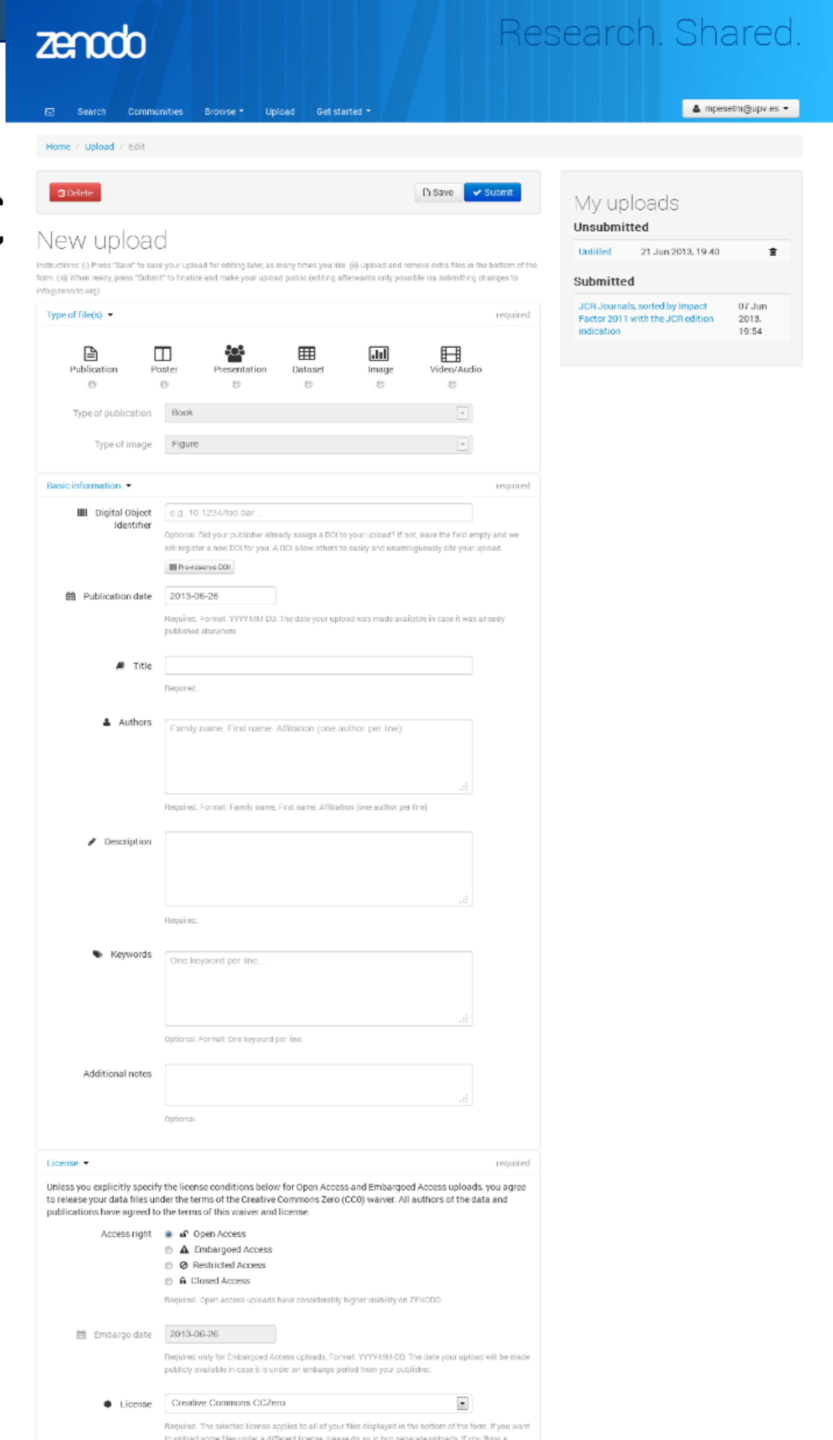
UNIVERSITAT POLITÈCNICA DE VALÈNCIA
IUMPA
TABOOW ICOSYSTEM GUARDIAN
Scie-Chain

Si tengo un repositorio que no acepta datos: caso de e-LIS, repositorio + colección

https://docs.google.com/document/d/1U1bvAG2MQy9fWNuBn6G3Q-eUcRmq2EhuvXsjTID8B2E/edit?ts=59f1a9d3



Zenodo https://zenodo.org/ y colecc



Desarrollado como repositorio huérfano por el CERN sobre Invenio. Recientemente ha sido relanzado para todo tipo de datos dentro el proyecto europeo OpenAIREplus y probado con datasets del *Wellcome Trust* (Príncipe, 2013).

UNIVERSITAT POLITÈCNICA DE VALÈNCIA IUMPA TABOOW ICOSYSTEM GUARDIAN Scie-Chain

Para cualquier dato: opción Figshare

figshare

Si los datos se han publicado: DRYAD / repositorio recomendado en la revista / temático

Dryad Membership: Members

- American Association for the Advancement of Science *
- American Society of Naturalists *
- The American Genetic Association *
- Association for Tropical Biology and Conservation
- Beulah Williams Library, Northern State University
- Botanical Society of America
- British Ecological Society *
- BMJ Publishing Group, Ltd. *
- Cambridge University Press
- European Society for Evolutionary Biology *
- The Genetics Society *
- German National Library of Medicine
- Oikos *
- Oxford University Press *
- The Royal Society
- Society for the Study of Evolution *
- Society of Systematic Biologists *
- United States Fish and Wildlife Service *
- University Library, University of Illinois at Urbana-Champaign
- Wiley *

Si no existe… **Construyamos el futuro con nuestras necesidades**

Servicios comunes

- **Almacenamiento**, recopilando datos de manera activa;
- Una plataforma de **publicación**, que ofrezca al público los datos de investigación y el software relacionado para ser compartidos y reutilizados;
- Servicios de **archivo**, respuesta a los requerimientos de los financiadores;
- Un servicio de **localización**, que permita a los investigadores y a la ciudadanía buscar

LEARN Toolkit of Best Practice for Research Data Management(2017). http://learn-rdm.eu/wp-content/uploads/RDMToolkit.pdf *DOI: https://doi.org/10.14324/000.learn.01*

FAIRización

- Rec.1. Ampliar la definición
- Rec.2. Explicitar en las políticas las limitaciones a as open as posible
- Rec.3. Modelos de objetos que apunten a identificadores persistentes

Definir

- Rec.4. Registros interconectados para los componentes mínimos del ecosistema: políticas, DMP, identificadores, normas y repositorios
- Rec.5. Financiación sostenible para mantener un nivel profesional
- Rec.6. Financiación basada en la evidencia del impacto y de la comunidad

Aplicar y desarrollar

- Rec.7. Marcos de interoperabilidad por disciplina
- Rec.8. FAIRización interdisciplinar con estándares globales (anti-silos)
- Rec.9. Métricas robustas y revisables
- Rec.10. Certificación de los repositorios
- Rec.11. Métricas para evaluar y certificar los servicios de datos

Asegurar

- Rec.12.DMP para gestionar
- Rec.13. Profesionalizacion de las capacidades
- Rec.14. Incentivos y recompensas

Alfabetizar

Hodson, Jones et al. (2018) Turning FAIR data into reality. Interim report of the European Commission Expert Group on FAIR data. https://doi.org/10.5281/zenodo.1285272

Repositorios de confianza

- Transparencia en sus políticas
- Certificación:
  - DRAMBORA-DCC (selfaudit);
  - Data Seal Approval-DANS (externo). Sustituido por [CoreTrustSeal Application Management Tool (AMT)](#) . Trabaja con Membership Management Tool (MMT) de [World Data System (ICSU-WDS)](#).
  - ISO 16363 Audit and Certification of Trustworthy Digital Repositories (evidencias)
  - FAIR

This course was made possible in part by the Institute of Museum and Library Services, under award #RE-06-13-0052-13

CSUC

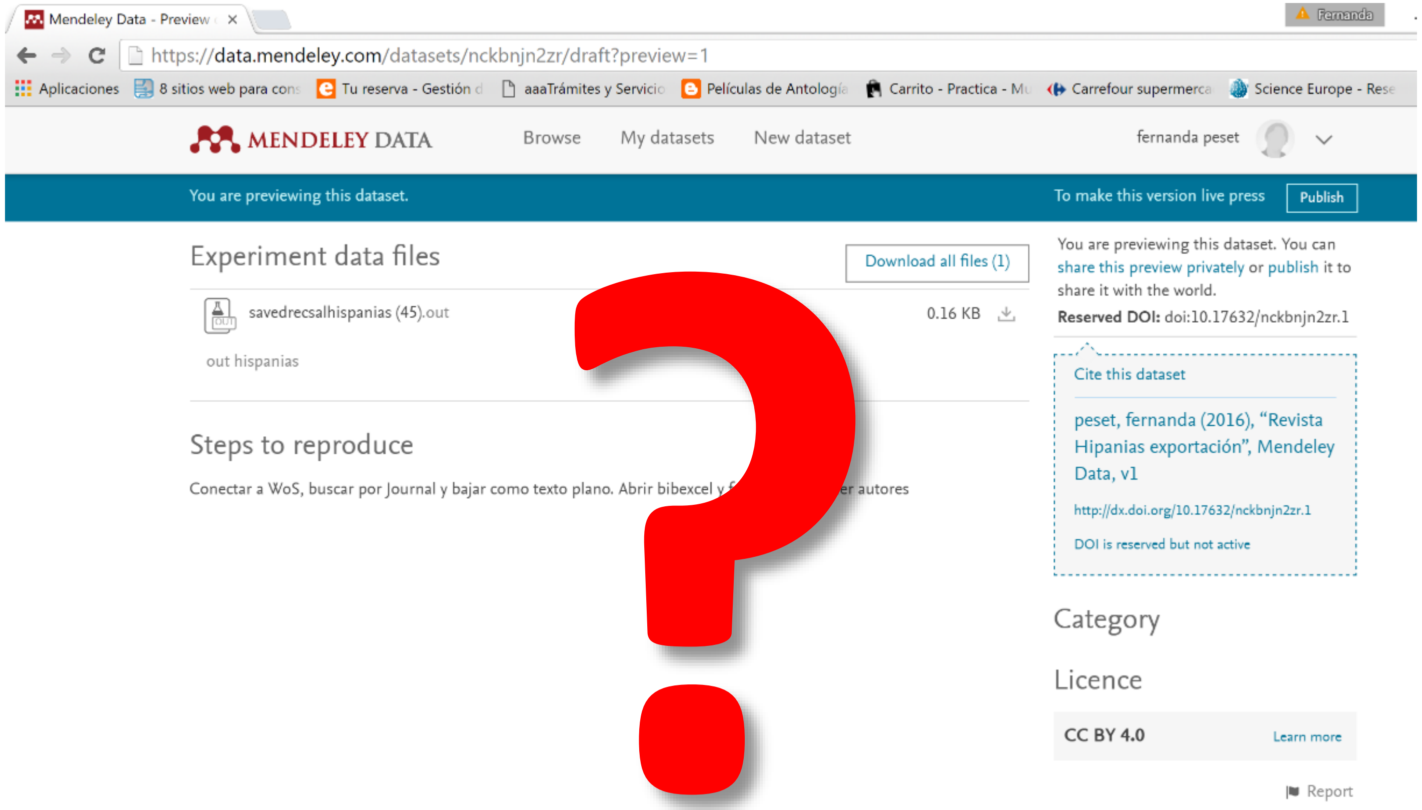
	zenodo CERN	EUDAT	figshare Digital Science (Macmillan)	DRYAD NCSU	Harvard Dataverse IQSS (Harvard University)	MENDELEY DATA Elsevier
1. Àmbit temàtic i geogràfic	Multidisciplinar Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional	Amb èmfasi en ciència i medicina Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional
2. Tipologia de dades (programari, imatges, raw data, etc.)	Tot tipus de formats i també documents publicats	Tot tipus de formats	Tot tipus de formats	Tot tipus de formats (amb exclusió de software) i amb preferència per formats de codi obert	Tot tipus de formats amb funcionalitats addicionals per alguns (RData, SPSS, STATA, CSV, xlsx, FITS)	Tot tipus de formats
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. Per a mides superiors revisar les condicions	Fins a 2 GB.	10 GB
4. Dades obertes, embargades, restringides o tancades	Obertes Tancades Restringides Embargades (Policies)	Obertes Restringides Embargades	Obertes Tancades (per defecte fins que no es publiquen) Restringides Embargades	Obertes Embargades (per defecte fins el moment de la publicació)	Obertes Restringides	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 (*metadata* & *datasets*) -CC-BY per a no dades (audiovisuals, *papers*, *etc.*) -MIT per a codi i software - Altres llicències: GNU, Apache 2.0, BSD 3 clause	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, gplv3
6. Identificador permanent	DOI o URL	Handle	DOI	DOI	DOI	DOI
7. Condicions per retirar contingut	Es permet retirar les dades justificant-ne el perquè. El DOI i la URL quedaran retinguts.	Per retirar contingut contactar amb EUDAT	Es pot sol·licitar la retirada de contingut públic justificant-ne el motiu	Cessió de dades irrevocable	Es permet retirar les dades, tot i que les metadades quedaran sempre visibles	Per retirar el contingut contactar amb Mendeley.

https://recercat.cat/bitstream/handle/2072/284974/1705_RecomanacionsSeleccionarRepositoriDades.pdf?sequence=5

Tabla comparativa de repositorios. Fuente: UAB http://ddd.uab.cat/pub/guibib/150829/repositorisdades_2016m9iSPA.pdf

https://data.mendeley.com/

Almacenado en servidores de Amazon S3 Germany. Preservado en DANS (Data Archiving and Network Services. Cita acorde con Force11. Datos modificables que se renombran. Públicos y no visibles. Facilidad de uso



Hablamos de un ecosistema
- Los datos Importan
- Ecosistema científico innovador

Datos de investigación
- De qué hablamos: definiciones prácticas
- Selección
- Gestionar datos de investigación
  - Solicitud
  - Proyecto activo
  - Publicación

Conclusión

Muchas formas de aproximarse al tema

- Institución/policymaker/financiador: conocer la producción de los empleados, cumplir con las políticas sobre lo abierto e incluso gestionar correctamente la información

- Gestor de información: preservar la información, procurar la interoperabilidad, controlar la producción científica…

- Investigador: cumplir con las exigencias del financiador o la revista en la que quiero publicar

- Desarrollador: implantar tecnologías y protocolos para ofrecer servicios avanzados en competencia

Las estrategias dependen de la perspectiva de quien esté siendo protagonista

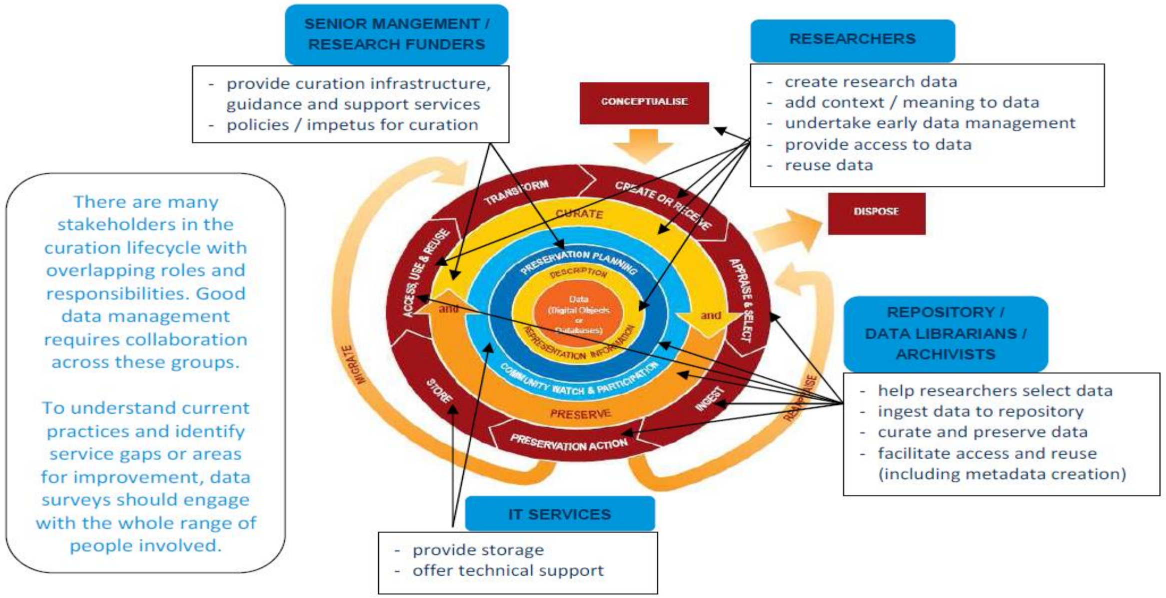
O

Instituciones vs. Financiadores

Integración vs. Dispersión

Orden vs. datadeluge

Modelos institución (ciclo de vida) vs. global (requerimientos y políticas)



New England course

NSF Data Management Plan Requirements

Beginning January 18, 2011, proposals submitted to NSF must include a supplementary document of no more than two pages labeled "Data Management Plan" (DMP) . This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results. Proposals that do not include a DMP will not be able to be submitted. For more information about this new requirement, please see the Grant Proposal Guide, Chapter II.C.2.j and the Data Management and Sharing Frequently Asked Questions(FAQs).

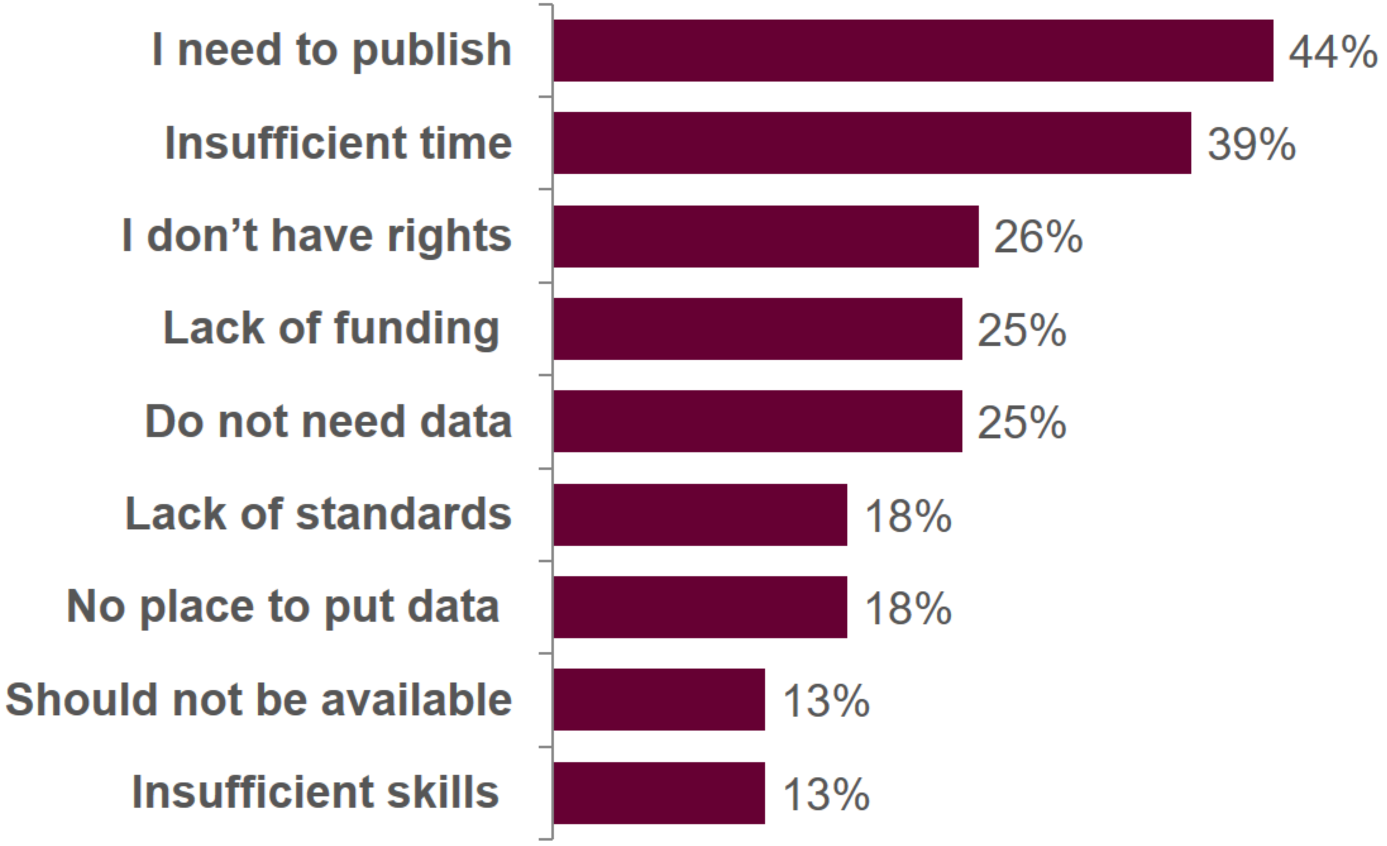
Please note: the Engineering Directorate (ENG) has additional guidance for proposals submitted to ENG programs, http://nsf.gov/eng/general/ENG_DMP_Policy.pdf. Questions and/or suggestions about this new requirement may be addressed to Dr. Maria K. Burka at mburka@nsf.gov.

Data during the research life-cycle. In the lifecycle diagram above, stages essential for all health and medical research are represented in blue, whilst additional research type specific processes are indicated in green. http://www.nhmrc.gov.au/grants-funding/policy/nhmrc-statement-data-sharing

Tenopir et al.

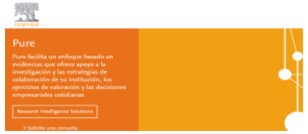
DATA SHARING

Barriers for scientists



- Research data is arguably the most visible aspect of the evolution of the scholarly record, presenting both challenges and opportunities for academic libraries, funders, publishers and other stakeholders. Pag.12

- Research data management is also an important element of new OCLC Research work in the area of research information management (RIM): “the integrated management of information about the research lifecycle, and about the entities which are party to it”



Informes de la OCLC: *The Evolving Scholarly Record* (ESR), *Stewardship of the Evolving Scholarly Record: From the Invisible Hand to Conscious Coordination*

Bryant, Rebecca, Brian Lavoie and Constance Malpas. 2017. *A Tour of the Research Data Management (RDM) Service Space*. The Realities of Research Data Management, Part 1. Dublin, Ohio: OCLC Research. doi:10.25333/C3PG8J

El reto actual son los datos de investigación

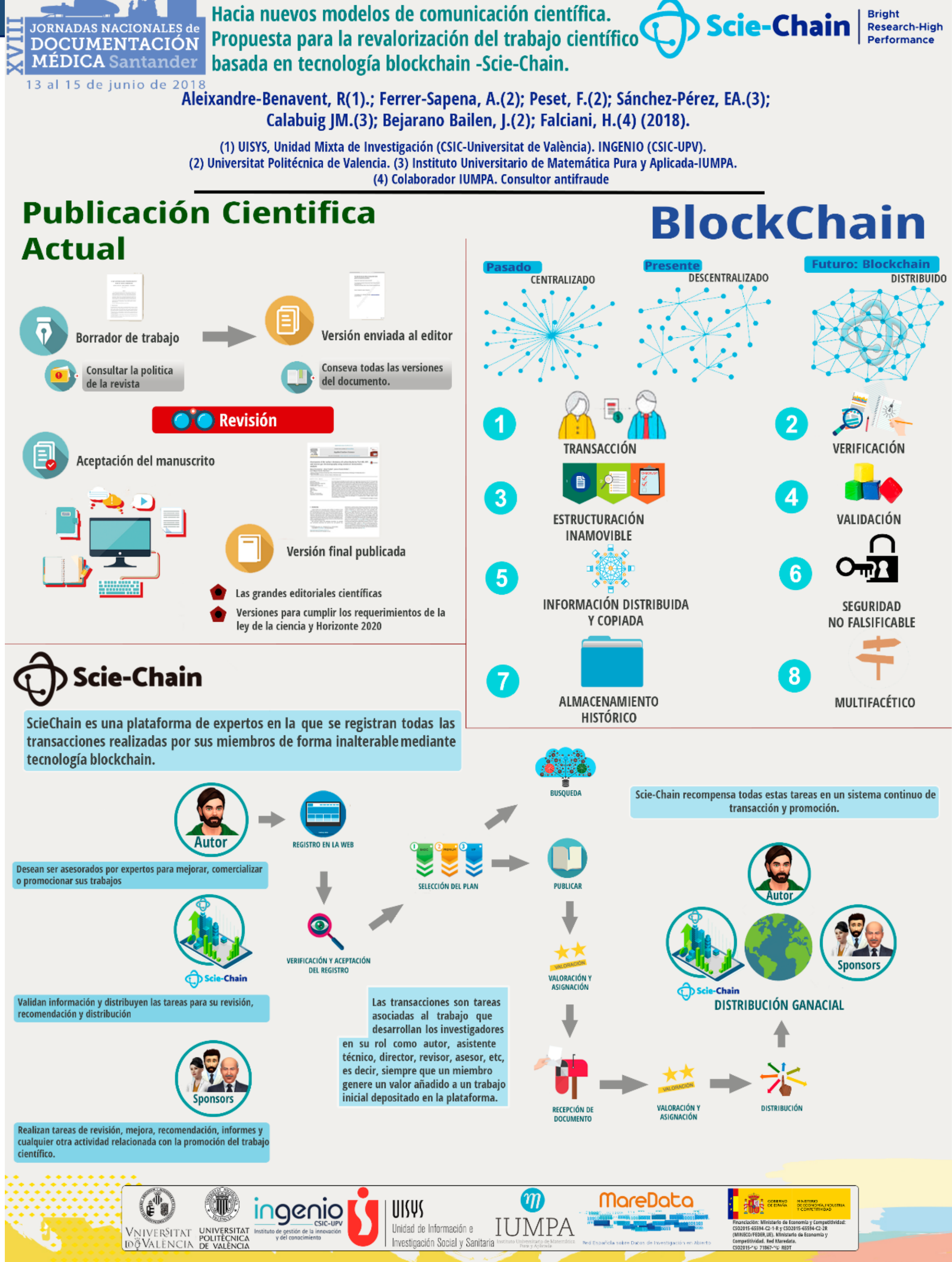
2013
- H2020
- Piloto

2015
- Open as default
- Open Aire+

2017
- as open as possible, as closed as necessary
- Analítica

FUTURO

TABOOW
ICOSYSTEM GUARDIAN
LA CRIPTOMONEDA NO ES TABÚ
¿Te interesa acceder a las criptomonedas?
Tecnologías éticas y seguras que te acercan a los mercados financieros
Participa del ICOsistema
www.taboow.org
En 2017 se han conseguido más de seis mil millones de dólares de financiación
¿Quieres encontrar financiación con blockchain?
TABOOW
ICOSYSTEM GUARDIAN
Scie-Chain



XVIII JORNADAS NACIONALES de DOCUMENTACIÓN MÉDICA Santander
13 al 15 de junio de 2018
Hacia nuevos modelos de comunicación científica.
Propuesta para la revalorización del trabajo científico basada en tecnología blockchain -Scie-Chain.
Scie-Chain
Bright Research-High Performance
Aleixandre-Benavent, R(1).; Ferrer-Sapena, A.(2); Peset, F.(2); Sánchez-Pérez, EA.(3); Calabuig JM.(3); Bejarano Bailen, J.(2); Falciani, H.(4) (2018).
(1) UISYS, Unidad Mixta de Investigación (CSIC-Universitat de València). INGENIO (CSIC-UPV).
(2) Universitat Politécnica de Valencia. (3) Instituto Universitario de Matemática Pura y Aplicada-IUMPA.
(4) Colaborador IUMPA. Consultor antifraude
Publicación Cientifica Actual
Borrador de trabajo
Versión enviada al editor
Consultar la politica de la revista
Conseva todas las versiones del documento.
Revisión
Aceptación del manuscrito
Versión final publicada
Las grandes editoriales científicas
Versiones para cumplir los requerimientos de la ley de la ciencia y Horizonte 2020
BlockChain
Pasado
CENTRALIZADO
Presente
DESCENTRALIZADO
Futuro: Blockchain
DISTRIBUIDO
1
TRANSACCIÓN
2
VERIFICACIÓN
3
ESTRUCTURACIÓN INAMOVIBLE
4
VALIDACIÓN
5
INFORMACIÓN DISTRIBUIDA Y COPIADA
6
SEGURIDAD NO FALSIFICABLE
7
ALMACENAMIENTO HISTÓRICO
8
MULTIFACÉTICO
Scie-Chain
ScieChain es una plataforma de expertos en la que se registran todas las transacciones realizadas por sus miembros de forma inalterable mediante tecnología blockchain.
Autor
REGISTRO EN LA WEB
BUSQUEDA
SELECCIÓN DEL PLAN
PUBLICAR
Desean ser asesorados por expertos para mejorar, comercializar o promocionar sus trabajos
VERIFICACIÓN Y ACEPTACIÓN DEL REGISTRO
VALORACIÓN Y ASIGNACIÓN
Scie-Chain
Validan información y distribuyen las tareas para su revisión, recomendación y distribución
Las transacciones son tareas asociadas al trabajo que desarrollan los investigadores en su rol como autor, asistente técnico, director, revisor, asesor, etc, es decir, siempre que un miembro genere un valor añadido a un trabajo inicial depositado en la plataforma.
Sponsors
Realizan tareas de revisión, mejora, recomendación, informes y cualquier otra actividad relacionada con la promoción del trabajo científico.
RECEPCIÓN DE DOCUMENTO
VALORACIÓN Y ASIGNACIÓN
DISTRIBUCIÓN
Scie-Chain recompensa todas estas tareas en un sistema continuo de transacción y promoción.
Autor
Sponsors
Scie-Chain
DISTRIBUCIÓN GANACIAL
VNIVERSITAT DE VALÈNCIA
UNIVERSITAT POLITÈCNICA DE VALÈNCIA
ingenio
CSIC-UPV
UISYS
Unidad de Información e Investigación Social y Sanitaria
IUMPA
MareData

Conclusiones

- Actuar sobre los componentes mínimos del ecosistema:
  - Políticas que reconozcan el trabajo y clarifiquen las obligaciones
  - DMP, planificar desde el inicio
  - Identificadores, normas y estándares comunes
  - Repositorios sostenibles y certificados

Campañas de sensibilización y Data Literacy

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Peset, Fernanda y González, Luis-Millán (2015). Openscience: la comunicación científica en abierto. http://www.youtube.com/watch?v=nTTGiAW4BLU

Peset, Fernanda (2015). Dataverse, un caso de éxito a escala mundial. http://www.youtube.com/watch?v=nnL3LaIuT0k

Peset, Fernanda (2016). Introducción a Data Citation Index. http://youtu.be/EZEdDwpA6eg

Peset, Fernanda (2016). Dataverse: uso institucional. http://youtu.be/HLKW3yd2k-g

Peset, Fernanda (2015). Figshare: una herramienta para gestionar y preservar datos de investigación de forma individual y para las organizaciones. http://www.youtube.com/watch?v=giXU2Nq8zDU

Algunos audiovisuales del curso Introducción a la Gestión de Datos de Investigaci

Peset, F. (2014). Datasea. https://polimedia.upv.es/visor/?id=a48abcad-21bc-ab4a-ac6a-af987b3e7f1a

Gracias por dejarme aprender con vosotros

La ciencia es una actividad internacional, pero que se realiza inmersa en una cultura de trabajo nacional por lo que necesita que las estrategias nacionales estén ajustadas al marco común internacional.

Boulton (2015)