

GESTÃO DE DADOS



CIENTÍFICOS

o papel das bibliotecas

Pedro Príncipe
Serviços de Documentação
da Universidade do Minho
Projeto OpenAIREplus



partilha, criatividade e engenho

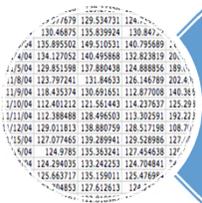
2º ENCONTRO de **BIBLIOTECAS DO ENSINO SUPERIOR**

Aveiro, 6 e 7 de junho de 2013

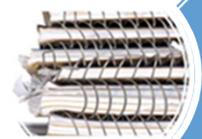
O mote para o workshop...

- compreender o **papel que as bibliotecas** têm hoje que desempenhar no domínio dos dados científicos produzidos nas instituições de investigação e ensino superior.
 - etapas da gestão de dados científicos,
 - ciclo de vida dos dados científicos,
 - criação de conjuntos de dados científicos,
 - metadados e repositórios de dados científicos.
- traçar uma **visão geral dos principais campos** de compreensão e intervenção dos profissionais e bibliotecas.

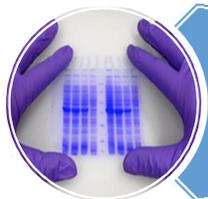
Tópicos:



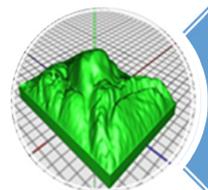
Dados científicos: definição, contexto e tipos



Gestão de dados científicos: relevância e atividades



O papel das bibliotecas na gestão de dados



Planeamento da gestão de dados científicos



Recursos, iniciativas, boas práticas

O que são dados científicos...

Data



Information



Presentation



Knowledge

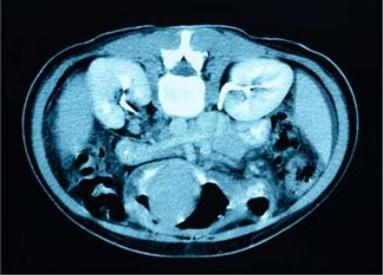


EpicGraphic.com

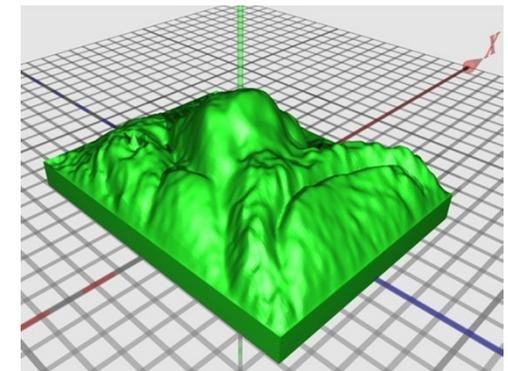
<http://epicgraphic.com/data-cake>



10/29/04	124.761606	129.27356	122.260995	209.662
11/1/04	119.977679	129.534731	124.739135	176.316
11/2/04	130.46875	135.839924	130.84732	168.289
11/3/04	135.895502	149.510531	140.795689	120.686
11/4/04	134.127052	140.495868	132.823819	206.138
11/5/04	129.851598	137.880438	124.888856	189.675
11/8/04	123.797241	131.84633	126.146789	202.496
11/9/04	118.435374	130.691651	112.877008	140.366
11/10/04	112.401212	121.561443	114.237637	125.298
11/11/04	112.388488	128.496503	113.302591	192.223
11/12/04	129.011813	138.880759	128.517198	108.701
11/15/04	127.077465	139.289941	129.528986	127.406
11/16/04	124.9785	135.363241	127.454638	129.669
11/17/04	124.294035	133.242253	124.704841	244.567
11/18/04	125.663717	135.159011	125.476984	169.271
11/19/04	123.704853	127.612613	124.25382	170.401
11/22/04	118.926697	122.818967	115.379664	134.970



Todo o tipo de coisas produzidas no decurso da investigação.



DADOS CIENTÍFICOS

- **Definição** da OCDE:
 - *“registos factuais usados como fontes primárias na investigação científica, e que são geralmente aceites na comunidade científica como necessários para validar os resultados de investigação”.*

DADOS CIENTÍFICOS

- Podem assumir várias **formas** (texto, números, imagens fixas, imagens em movimento, etc.) e
- **dimensões**, desde registos de observações individuais ou ensaios de pequenos laboratórios que não ultrapassarão algumas centenas de kilobytes, até aos dados produzidos pelo *Large Hadron Collider (LHC)* do CERN, que pode gerar várias dezenas de petabytes por dia.

DADOS CIENTÍFICOS

- Analógicos ou digitais
- Livros de laboratório ou software
- Criados na forma digital ('born digital')
- Ou convertidos para a forma digital ('digitised')

DADOS CIENTÍFICOS

- Os dados científicos são produzidos ou utilizados no contexto de investigação científica.
 - **recolhidos ou criados para efeito de processamento científico**, como os dados atmosféricos usados para previsão meteorológica, ou os dados recolhidos de sensores para monitorizar o estado de um edifício.
 - **obtidos como resultados do processamento automático de objetos** (por exemplo, uma coleção de imagens processada para obter os respetivos histogramas de cor, que constituem assim novos dados).
 - **dados que não são produzidos para investigação mas que acabam por ser objeto dela**, como as contribuições que os utilizadores de uma rede social fazem na forma de textos, fotografias ou outros objetos e que acabam por ser utilizados para estudos sociológicos.

Tipos de datos científicos

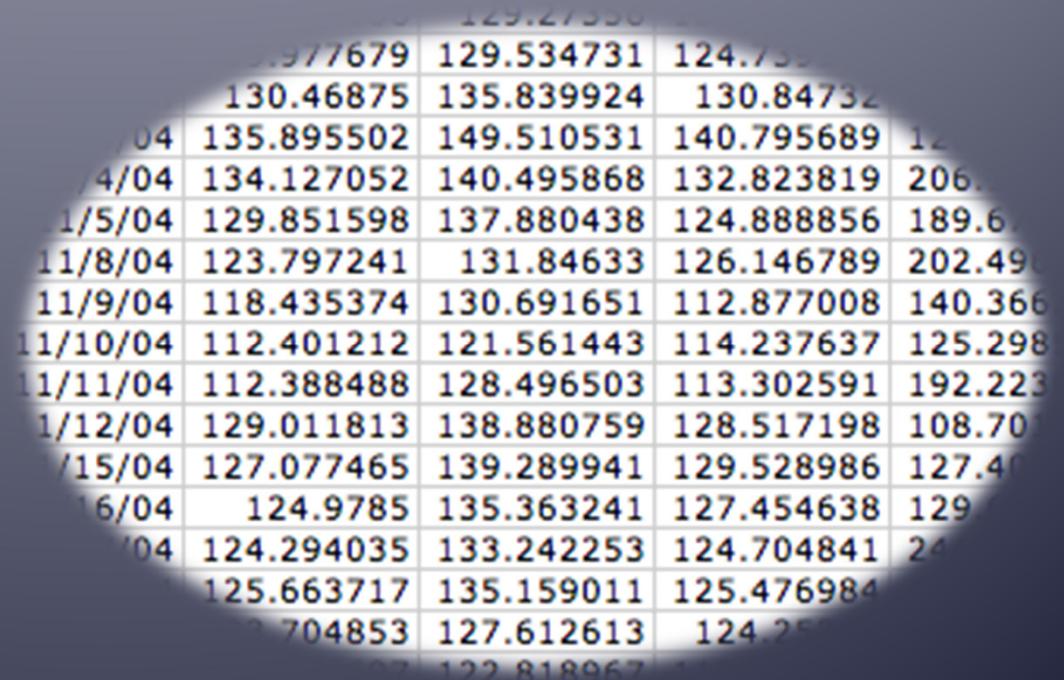
- Instrument measurements
- Experimental observations
- Still images, video and audio
- Text documents, spreadsheets, databases
- Quantitative data (e.g. household survey data)
- Survey results & interview transcripts
- Simulation data, models & software
- Slides, artefacts, specimens, samples
- Sketches, diaries, lab notebooks...

DADOS CIENTÍFICOS

- Para se constituírem como verdadeiramente úteis, os dados científicos devem possuir estrutura e organização.
 - Os conjuntos de dados (“*datasets*”) são uma das unidades essenciais.
 - Os conjuntos de dados são coleções de informações ou factos relacionados entre si e registados num formato comum.
 - Por exemplo, os resultados de um estudo de opinião por entrevista numa investigação sociológica constituem um conjunto de dados, composto pelos registos individuais das entrevistas.

GESTÃO DE DADOS CIENTÍFICOS

Relevância, atividades, princípios



		129.27358			
		129.977679	129.534731	124.75	
		130.46875	135.839924	130.84732	
	04	135.895502	149.510531	140.795689	12
	4/04	134.127052	140.495868	132.823819	206.
	1/5/04	129.851598	137.880438	124.888856	189.6
	11/8/04	123.797241	131.84633	126.146789	202.49
	11/9/04	118.435374	130.691651	112.877008	140.366
	11/10/04	112.401212	121.561443	114.237637	125.298
	11/11/04	112.388488	128.496503	113.302591	192.223
	1/12/04	129.011813	138.880759	128.517198	108.70
	/15/04	127.077465	139.289941	129.528986	127.4
	16/04	124.9785	135.363241	127.454638	129
	/04	124.294035	133.242253	124.704841	24
		125.663717	135.159011	125.476984	
		127.704853	127.612613	124.25	
			122.818967		

RELEVÂNCIA DO TEMA

É cada vez mais relevante o entendimento da **necessidade de gerir o acesso e a utilização dos dados** produzidos ou recolhidos no âmbito das atividades de investigação, garantindo a sua preservação.

RELEVÂNCIA DO TEMA

- Cresce a **sensibilidade dos investigadores** para os efeitos da verdadeira explosão na produção de dados científicos:
 - crescimento global das atividades de investigação
 - consequência dos novos métodos e instrumentos de pesquisa e registo que originam cada vez maiores volumes de dados.
 - Onde o trabalho de investigação é baseado na análise de dados recolhidos de forma distribuída parece ter sido mais precoce e mais profundo.

RELEVÂNCIA DO TEMA

- Os **organismos que produzem e financiam ciência** têm manifestado nestes últimos anos redobrada atenção à necessidade de assegurar meios e mecanismos para a gestão dos dados científicos.

RELEVÂNCIA DO TEMA

- OCDE
 - OECD. *Declaration on Access to Research Data From Public Funding*, Paris, 2004.
http://www.oecd.org/document/15/0,3343,en_2649_34487_25998799_1_1_1_1,00.html
 - *OECD Principles and Guidelines for Access to Research Data from Public Funding*. Paris, 2007.
<http://www.oecd.org/dataoecd/9/61/38500813.pdf>
- Várias iniciativas dos Conselhos de Investigação do Reino Unido
 - RCUK. (2011). Common Principles on Data Policy -
<http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>
 - Jones, Sarah. (2011). Summary of UK research funders' expectations for the content of data management and sharing plans <http://www.dcc.ac.uk/>
- Iniciativas no Contexto da União Europeia produzidas pela Comissão Europeia no âmbito da Agenda Digital Europeia e ERA
 - CE (2012): Communication on a reinforced European Research Area partnership for excellence and growth
 - Communication Towards better access to scientific information
 - Recommendation on access to and preservation of scientific information
- EUA com iniciativas da National Science Foundation - <http://www.nsf.gov>

Relevância do tema

Riding the wave

How Europe can gain from the rising tide of scientific data

Final report of the High Level Expert Group on Scientific Data
A submission to the European Commission

October 2010

“To make progress in science, we need to be open and share... sharing data, and having the forum to openly use and build on what is shared, are essential to science. They fuel the progress and practice of scientific discovery. ”

Neelie Kroes, Rome, 11 April 2012



RELEVÂNCIA DO TEMA

- Multiplicam-se as **iniciativas** de *Open Access, Open Data*.
 - *Panton Principles for Open Data in Science*: <http://pantonprinciples.org>



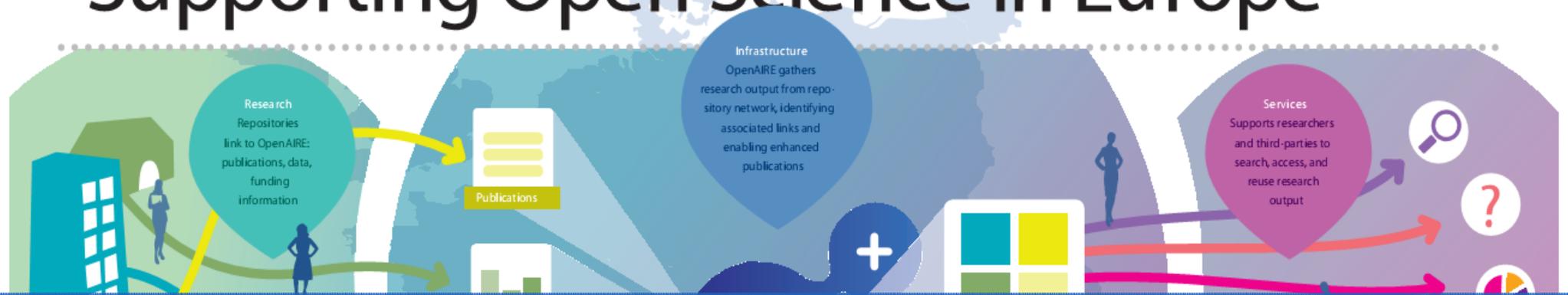
“Entendemos por dados abertos, em ciência, aqueles que estão livremente disponíveis na Internet pública permitindo a qualquer utilizador baixar, copiar, analisar, reprocessar, usá-los com um software ou utilizá-los para qualquer outra finalidade sem barreiras financeiras, legais, técnicas ou outras além do acesso à Internet. **Para este fim, os dados relacionados com ciência publicada devem ser explicitamente colocados em domínio público.**”

RELEVÂNCIA DO TEMA



»Modern science needs the free flow of knowledge...in an e-infrastructure that is open across national borders, disciplines and scientific communities«
Needie Kroes (European Commission, 2012)

Supporting Open Science in Europe



OpenAIRE implementa as políticas Open Access da UE

Who benefits from OpenAIRE?

- EU researchers who access, deposit and link to research output
- National Open Access initiatives
- Repository managers
- Policy makers and funders who monitor funded work
- Publishers who wish to raise visibility of output
- Potential data providers who want to explore linking up their research

What is OpenAIRE?

- A Participatory European Open Access infrastructure to manage scientific publications and associated information via repository networks
- Harvests and indexes FP7 Open Access publications
- Harvests subsets of related data, and other contextual information, cross-linking them to demonstrate Enhanced Publications
- The OpenAIRE portal provides a suite of services
 - deposit and access
 - guidelines and a helpdesk
- OpenAIRE runs a series of workshops, and produces

Why is OpenAIRE important?

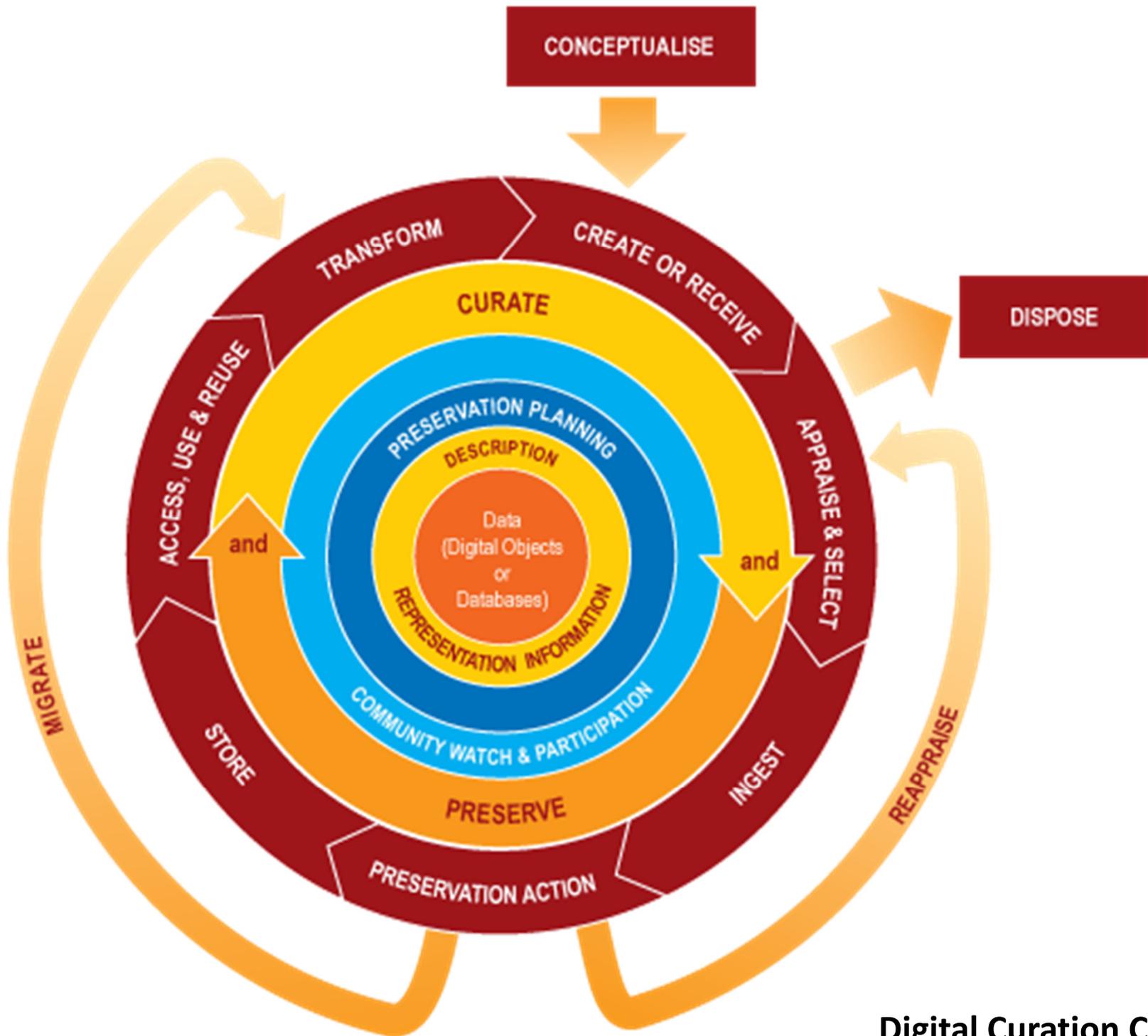
- By facilitating Open Science and Open Access, OpenAIRE allows scientists to access, reuse and enhance and research output
- OpenAIRE provides a cross-discipline support service for European Scientists
- Tools such as publication usage statistics
- OpenAIRE is based on
 - versatile technology and innovative research
 - European outreach effort which advocates Open Access

Who is OpenAIRE?

- OpenAIRE is an FP7 funded project, now in its second phase of funding until May 2014
- 41 project partners include 3 scientific communities: EBI, DANS and BADC
- Collaboration with EuroCRIS, EUDAT, DataCite, COAR, LIBER, SPARC Europe
- Contact
Project Coordinator:
Mike Hatzopoulos, mike@dluaogr
- www.openaire.eu

O que está envolvido na gestão de dados científicos:

- Planeamento e gestão de dados
- Criação de dados
- Documentar os dados
- Acesso e utilização dos dados
- Armazenamento e backups
- Partilha dos dados
- Preservação dos dados



Curadoria

- Curadoria de dados é a atividade de gestão e utilização de dados desde o momento da sua criação para garantir o seu armazenamento, a sua pesquisa e respetiva reutilização.

“the active management and appraisal of data over the lifecycle of scholarly and scientific interest”

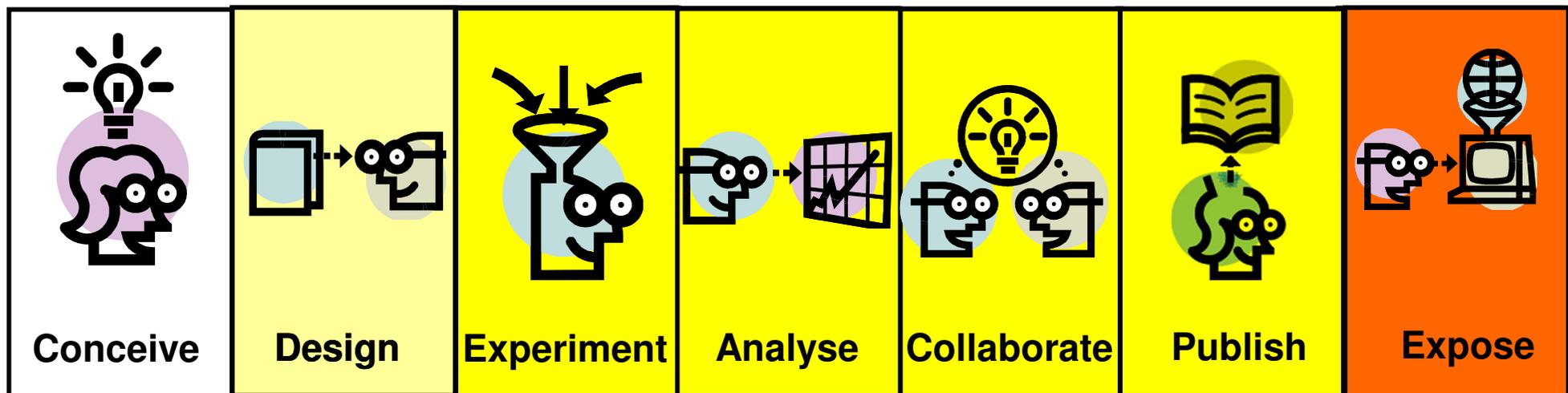
Digital Curation Centre

O ciclo de vida dos dados e a sua gestão:

Gestão e
Planeamento
de dados

Plataforma de
Gestão de dados científicos

Repositórios disciplinares
Ou institucionais e
Revistas científicas



Razões para os investigadores se preocuparem com a gestão dos dados

- Debate entre os participantes...

Razões para as instituições se preocuparem com a gestão dos dados

- Debate entre os participantes...

As bibliotecas neste processo

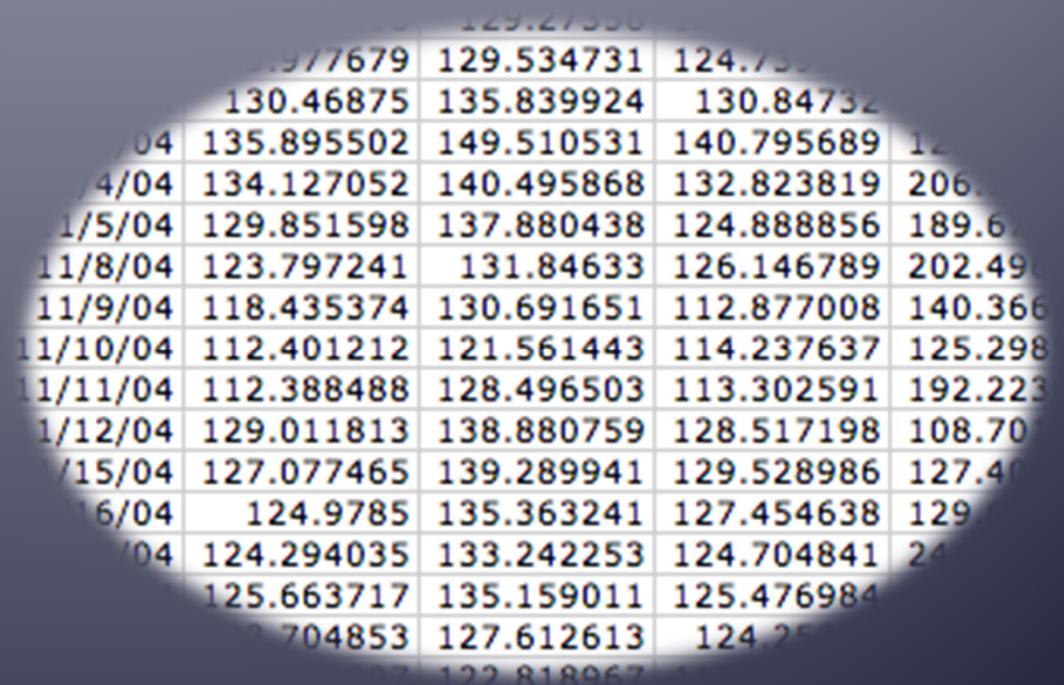
- Apoiar e aconselhar estudantes, investigadores, grupos de investigação sobre os meios e formas mais adequadas para **preparar os dados** para o seu armazenamento, a pesquisa e a reutilização.

Algumas leituras e recursos de referência

- Knowledge-Exchange (2011), "[A Surfboard for Riding the Wave Towards a four country action programme on research data](#)"
- Ball, A and Duke, M (2011) How to Cite Datasets and Link to Publications. Digital Curation Centre http://www.dcc.ac.uk/webfm_send/525
- Green, T (2009), "We Need Publishing Standards for Datasets and Data Tables", OECD Publishing White Paper, OECD Publishing. doi: 10.1787/603233448430 <http://dx.doi.org/10.1787/603233448430>
- MaRDI-Gross project *DMP Planning for Big Science Projects* (2012). <http://arxiv.org/abs/1208.3754v1>
- Digital Curation Centre – <http://www.dcc.ac.uk>
- International Journal of Digital Curation - <http://www.ijdc.net/>
- Data Science Journal - <http://www.jstage.jst.go.jp/browse/dsj>
- Open Knowledge Foundation: <http://okfn.org/>
- OpenAIRE: <http://www.openaire.eu>

O PAPEL DAS BIBLIOTECAS NA GESTÃO DE DADOS CIENTÍFICOS

As 10 recomendações da LIBER



A circular highlight is centered over a table of numerical data. The table consists of 10 columns and 10 rows of data. The numbers are arranged in a grid, with some values appearing to be averages or totals for each row. The values range from approximately 108.70 to 206.10.

		129.27350			
		129.77679	129.534731	124.75	
		130.46875	135.839924	130.84732	
	04	135.895502	149.510531	140.795689	12
	4/04	134.127052	140.495868	132.823819	206.1
	1/5/04	129.851598	137.880438	124.888856	189.6
	11/8/04	123.797241	131.84633	126.146789	202.49
	11/9/04	118.435374	130.691651	112.877008	140.366
	11/10/04	112.401212	121.561443	114.237637	125.298
	11/11/04	112.388488	128.496503	113.302591	192.223
	1/12/04	129.011813	138.880759	128.517198	108.70
	/15/04	127.077465	139.289941	129.528986	127.4
	16/04	124.9785	135.363241	127.454638	129
	/04	124.294035	133.242253	124.704841	24
		125.663717	135.159011	125.476984	
		127.704853	127.612613	124.25	
			122.818967		

Recomendações para as Bibliotecas



Ten recommendations for libraries to get started with research data management

Final report of the LIBER working group on E-Science / Research Data Management

1. Apoio na gestão dos dados científicos

- Offer research data management support, including data management **plans for grant applications**, intellectual property **rights** advice and **information** materials. Assist faculty with data management plans and the integration of data management into the curriculum.



Ten recommendations
for libraries to get
started with research
data management

Final report of the LIBER working group on E-Science / Research Data Management

Importância do plano de GDC



MANTRA - Ellie Bates - Importance of data management plan in a PhD work

<http://youtu.be/1stLmJUO81A>

2. Desenvolvimento de serviços e metadados

- Engage in the development of **metadata** and data **standards** and provide metadata **services** for research data.



Ten recommendations
for libraries to get
started with research
data management

Final report of the LIBER working group on E-Science / Research Data Management



DataCite - International Data Citation

DataCite Metadata Scheme for the Publication and Citation of Research Data

Version 2.0

January 2011

3. Desenvolvimento de competências

- Create **Data Librarian** posts and develop professional **staff skills** for data librarianship.



Ten recommendations
for libraries to get
started with research
data management

Final report of the LIBER working group on E-Science / Research Data Management

RDMRose Learning Materials

On this page:

- [Session 1: Introductions, RDM, and the Role of LIS](#)
 - [Session 1.1: Introduction to the RDMRose module](#)
 - [Session 1.2: RDM basics](#)
 - [Session 1.3: The LIS role in RDM](#)
 - [Session 1.4: Reflection and reflective writing](#)
- [Session 2: The Nature of Research and the Need for RDM](#)
 - [Session 2.1: The social organisation of research](#)
 - [Session 2.2: Research, information practices and data](#)
 - [Session 2.3: The RDM agenda](#)
 - [Session 2.4: The research data interview and audit](#)
 - [Session 2.5: Reflection on research](#)
- [Session 3: The DCC Curation Lifecycle Model](#)
 - [Session 3.1: Exploring the DCC Curation Lifecycle Model](#)
 - [Session 3.2: Data Management Plans](#)

ABOUT RDMROSE

RDMRose is a JISC funded project to produce taught and continuing professional development (CPD) learning materials in Research Data Management (RDM) tailored for Information professionals.

GIVE US FEEDBACK

If you have used any of the RDMRose learning materials, be it for your own use as self-supported CPD or as an educator, we would greatly welcome your feedback. Please fill out this [evaluation form](#). It will help us further enhance the learning materials!

MANTRA

Research Data Management Training

[Home](#) | [Software practicals](#) | [Project overview](#) | [University of Edinburgh guidance](#) | [Testimonials](#) | [Acknowledgements](#) | [Feedback](#)

Online learning units

 [Introduction to the course](#)

 [Research data explained](#)

 [Data management plans](#)

 [Organising data](#)

 [File formats & transformation](#)

 [Documentation & metadata](#)

 [Storage & security](#)

 [Data protection, rights & access](#)
NEW

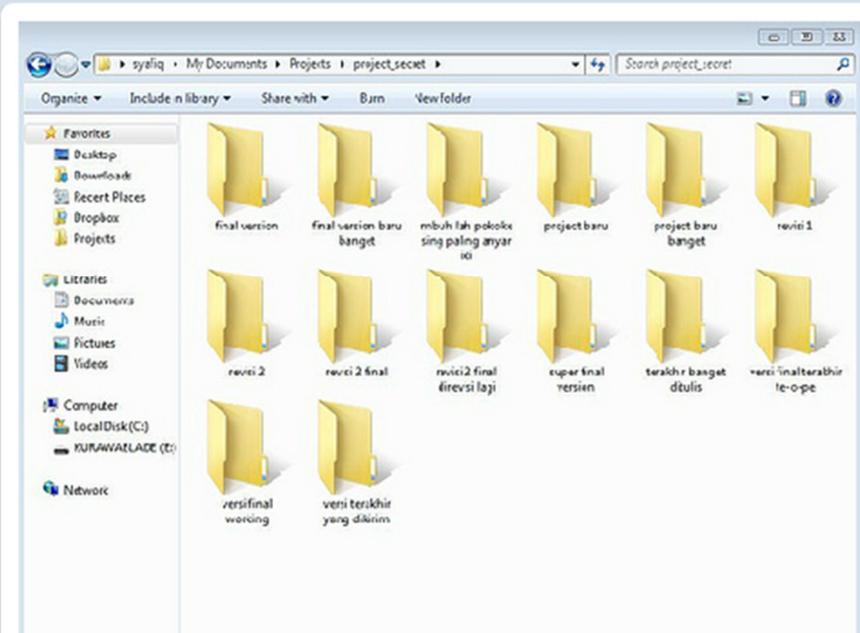
 [Preservation, sharing & licensing](#)

Organising data

Versioning

It is important to identify and distinguish versions of research data files consistently. This ensures that a clear audit trail exists for tracking the development of a data file and identifying earlier versions when needed. Thus you will need to establish a method that makes sense to you that will indicate the version of your data files.

- A common form for expressing data file versions is to use ordinal numbers (1,2,3 etc.) for major version changes and the decimal for minor changes e.g v1, v1.1, v2.6
- Beware of using confusing labels: revision, final, final2, definitive_copy as you may find that these accumulate
- Record every change irrespective of how minor that change may be



4. Políticas institucionais de gestão dos dados.

- Actively participate in institutional research **data policy development**, including resource plans. Encourage and adopt **open data** policies where appropriate in the research **data life cycle**.



Ten recommendations
for libraries to get
started with research
data management

Final report of the LIBER working group on E-Science / Research Data Management

Políticas institucionais de GDC



Professor Jeff Haywood
Vice Principal Knowledge Management

0:01 / 4:37

Research data management policies - Jeff Haywood
http://youtu.be/V8IldfBAr_0

The image shows a video player interface. The main content is a video of Professor Jeff Haywood, a middle-aged man with short hair, wearing a grey suit, white shirt, and patterned tie. He is looking slightly to the right of the camera. Below the video frame, there is a semi-transparent white box containing his name and title in bold black text. At the bottom of the video frame, there is a black control bar with a play/pause button, a volume icon, a progress bar showing 0:01 / 4:37, and several other icons (list, settings, full screen, etc.). Below the video player, the video title "Research data management policies - Jeff Haywood" is displayed in black text, followed by a purple underlined URL: http://youtu.be/V8IldfBAr_0.

5. Trabalhar em articulação com os parceiros para fomentar infraestruturas

- **Liaise and partner** with researchers, research groups, data archives and data centers to foster an interoperable infrastructure for data access, discovery and data sharing.



Ten recommendations
for libraries to get
started with research
data management

Final report of the LIBER working group on E-Science / Research Data Management

6. Disponibilizar serviços de curadoria no ciclo de vida dos dados científicos

- **Support the lifecycle** for research data by providing services for **storage, discovery** and **permanent access**.



Ten recommendations
for libraries to get
started with research
data management

Final report of the LIBER working group on E-Science / Research Data Management

A Digital Curation Centre
'working level' guide



How to Develop a Data Management and Sharing Plan

Sarah Jones (DCC)



Digital Curation Centre, 2011.
Licensed under Creative Commons Attribution 2.5 Scotland:
<http://creativecommons.org/licenses/by/2.5/scotland/>

DMP*online*
The  Data Management Planning Tool

<https://dmponline.dcc.ac.uk>

www.dcc.ac.uk/resources/how-guides/develop-data-plan

7. Citação dos dados e links e identificadores permanentes

- Promote research **data citation** by applying persistent identifiers to research data.



Ten recommendations
for libraries to get
started with research
data management



Helping you to find,
access, and reuse data

DataCite

DataCite Metadata Version 3.0 Available for Comment and Testing

Published by Sergio Ruiz on 3 June 2013 - 9:17pm

The DataCite Metadata Working Group is pleased to present a new version of the DataCite Schema for review, comment and testing. We will be giving greatest priority to feedback from DataCite clients and members, but we invite the scholarly and research community to contribute to this discussion as well.

With Version 3.0, we have worked to respond appropriately to a large number of feature requests while adhering to our goal of remaining domain agnostic. As well, we continue to strive for simplicity wherever possible.

Tags:

[datacite](#), [Metadata](#), [Schema](#), [identifiers](#), [DOI](#), [standards](#)

[Read more](#)

DataCite Summer Meeting - 19-20 September 2013 - Washington DC

Published by Sergio Ruiz on 31 May 2013 - 12:55pm

Hold the date!

The DataCite Summer Meeting 2013 will take place at the National Academy of Sciences (NAS) in Washington, D.C., on September 19th - 20th. The meeting will begin at 13:30 on September 19th and will be collocated with the [DataCite Summer Meeting \(2013\)](#) (DOI: 10.5284/1000164).

Why cite data?

What is DataCite?

What do we do?

DOI resolver

Resolve a DOI string (e.g. 10.5284/1000164) by entering it below:

Metadata Search

8. Repositórios de dados

- Provide an **institutional Data Catalogue** or **Data Repository**, depending on available **infrastructure**.



Ten recommendations
for libraries to get
started with research
data management

Search 731 records for.

euro-basin

[Q Search](#)

[Any Collection](#)

[Any Author](#)

[Any Year](#)

Showing records 1 to 10 out of 38 results.

1

01 January 2013 **Journal article** **Open access**

Capturing quantitative zooplankton information in the sea: Performance test of laser optical plankton counter and video plankton recorder in a *Calanus finmarchicus* dominated summer situation

[View](#)

[Basedow, Sünnje L.](#) ; [Tande, Kurt S.](#) ; [Norrbin, M. Fredrika](#) ; [Kristiansen, Stian A.](#)

We compared two optical plankton counters, the Laser Optical Plankton Counter (LOPC) and the Video Plankton Recorder (VPR) for their abundance estimates of *Calanus finmarchicus* during an early summer situation (June 2008) in two North Norwegian fjords. [...]

Uploaded by [Thomé](#) on 23 May 2013.

2

[View](#)

9. Envolve-se e pratique a gestão de dados científicos

- Get involved in subject specific **data management practice.**



Ten recommendations
for libraries to get
started with research
data management

Final report of the LIBER working group on E-Science / Research Data Management

10. Oferecer ou mediar serviços de armazenamento

- Offer or mediate secure storage for dynamic and static research data in co-operation with institutional IT units and/or seek exploitation of appropriate cloud services.



Ten recommendations
for libraries to get
started with research
data management

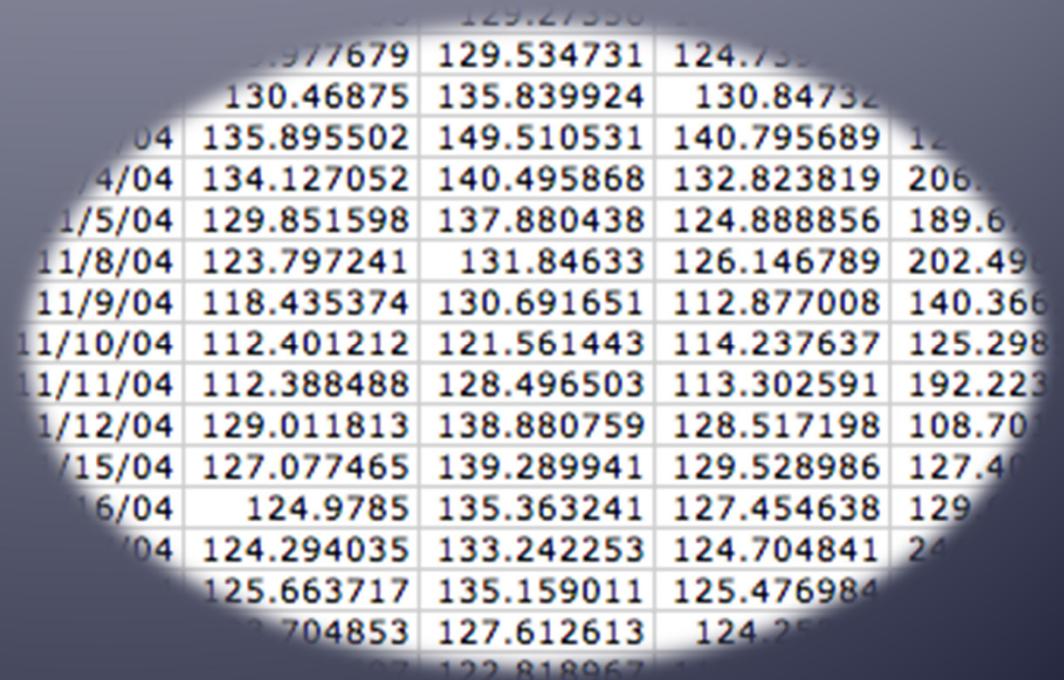
Final report of the LIBER working group on E-Science / Research Data Management

Exemplos

- University of Edinburgh policy
 - <http://www.ed.ac.uk/is/research-data-policy>
- Research data guidance
 - <http://www.ed.ac.uk/is/data-management>
- MANTRA online training
 - <http://datalib.edina.ac.uk/mantra/>
- Edinburgh University Data Library
 - <http://www.ed.ac.uk/is/data-library>

PLANEAMENTO DA GESTÃO DE DADOS CIENTÍFICOS

Elementos, financiadores, políticas institucionais



A circular highlight is centered over a table of numerical data. The table has five columns and multiple rows. The data is as follows:

	129.27358			
	129.977679	129.534731	124.75	
	130.46875	135.839924	130.84732	
04	135.895502	149.510531	140.795689	12
4/04	134.127052	140.495868	132.823819	206.
1/5/04	129.851598	137.880438	124.888856	189.6
11/8/04	123.797241	131.84633	126.146789	202.49
11/9/04	118.435374	130.691651	112.877008	140.366
11/10/04	112.401212	121.561443	114.237637	125.298
11/11/04	112.388488	128.496503	113.302591	192.223
1/12/04	129.011813	138.880759	128.517198	108.70
/15/04	127.077465	139.289941	129.528986	127.4
16/04	124.9785	135.363241	127.454638	129
/04	124.294035	133.242253	124.704841	24
	125.663717	135.159011	125.476984	
	127.704853	127.612613	124.25	
		122.818967		

PLANEAMENTO DA GDC

- Principais questões no planeamento da gestão de dados científicos:
 - Tipos de dados, formatos, normas e métodos de recolha
 - Questões legais e éticas, de propriedade intelectual
 - Acesso, Partilha e reutilização dos dados
 - Gestão e armazenamento de curto prazo
 - Depósito (arquivo) e preservação a longo prazo

Gestão de Dados Científicos

- As **condições em que os dados recolhidos ou produzidos** numa investigação podem, ou não, ser acedidos e reutilizados por outros investigadores, para além do contexto em que foram gerados, são questões importantes.
- A forma como são cuidados (**curadoria de dados**) e as **condições legais** associadas ao seu acesso e partilha constituem os dois elementos determinantes do futuro dos diversos conjuntos de dados científicos.
- É preciso garantir que os dados são **registados, mantidos e preservados** de forma adequada.

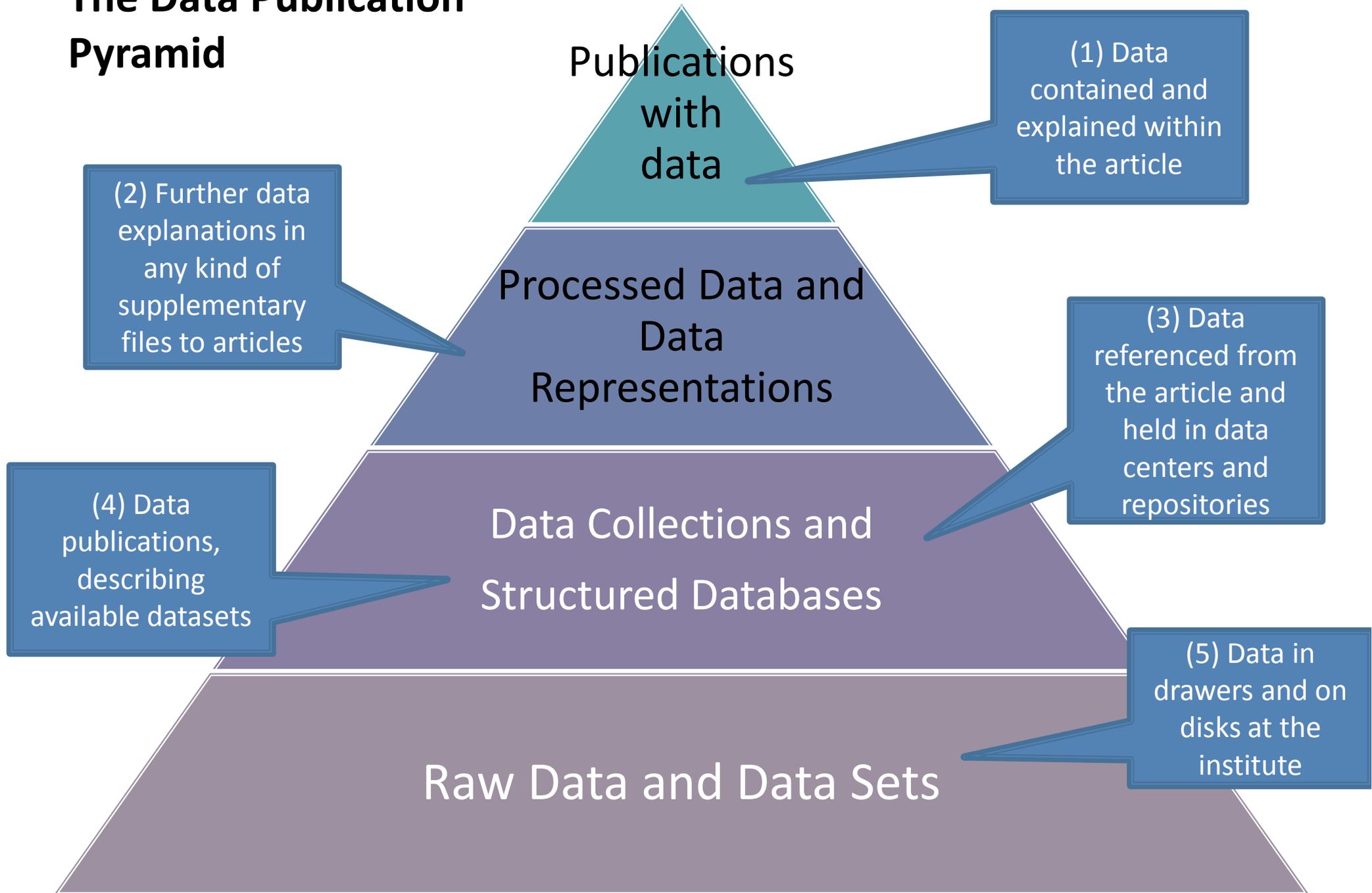
Gestão de Dados Científicos

- Um dos primeiros requisitos é que os conjuntos de dados sejam **acompanhados de informação que descreva** a sua origem (tempo ou espaço, métodos e instrumentos de recolha), âmbito, autoria, propriedade e condições de reutilização, ou seja, de metadados.
- Em paralelo com a **interoperabilidade** tecnológica, a existência de metadados adequados e normalizados é um requisito essencial para o acesso e reutilização dos dados científicos.

DOCUMENTAÇÃO E METADADOS

- Repositórios de dados científicos e a normalização dos metadados.
- Difícil uniformizar a descrição entre domínios científicos.
- A descrição dos conjuntos de dados ainda se pode considerar muito pouco desenvolvida.
- Ao nível do conjunto têm sido adoptados modelos genéricos como o do Dublin Core, já muito utilizado nos repositórios institucionais.
- Projecto DataShare (Edimburgo, Oxford e Southampton)

The Data Publication Pyramid

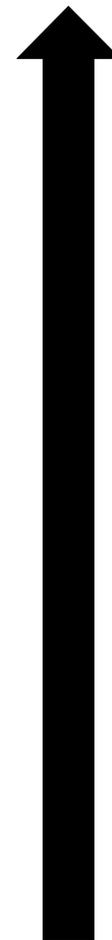
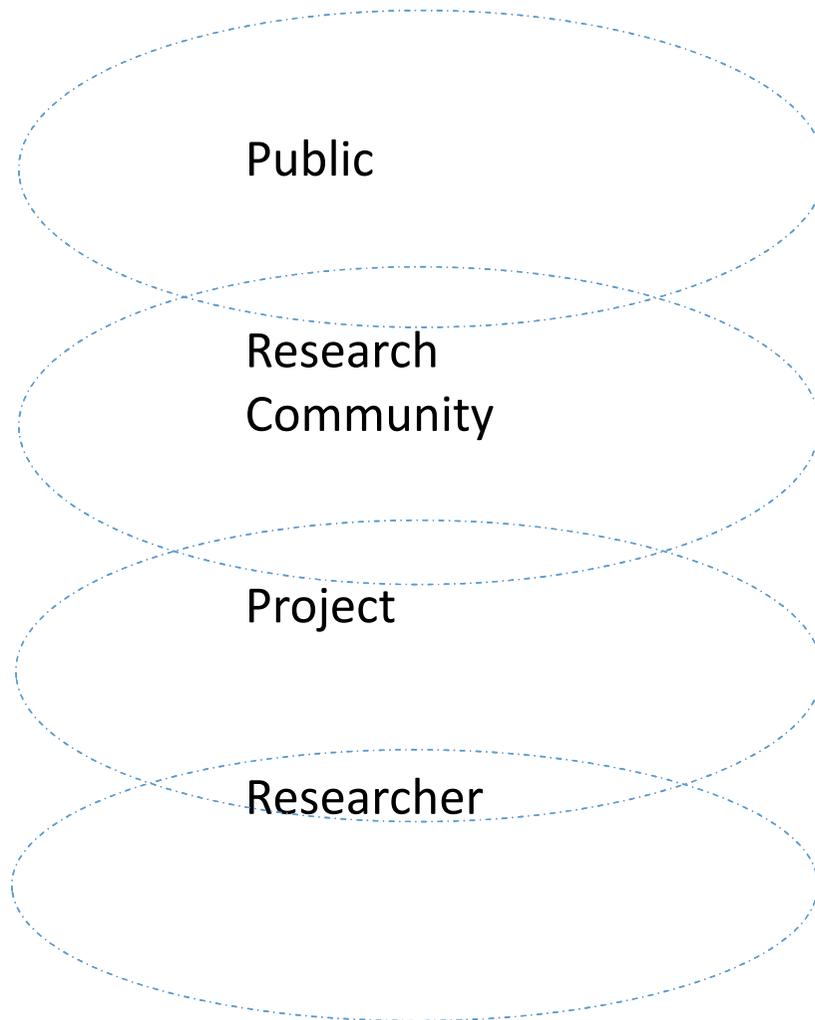


Importância de documentar os dados



http://youtu.be/7IN_SD5B43U

Necessidade dos metadados



Metadata may not be required if you are working alone on your own computer, but become crucial when data are shared online.

Metadata help to place your dataset in a broader context, allowing those outside your institution, discipline, or research environment to understand how to interpret your data.

Elementos do planeamento da gestão de dados científicos

	ANU	ANDS	DCC	FSSDA	Geo Aus	MIT	NSB	NSF Eng	QUT	RELU	Melbourne
Data description	•	•	•	•	•	•	•	•	•	•	•
Existing data	•	•	•						•	•	
Format	•	•	•	•	•	•	•	•	•	•	•
Metadata		•	•	•			•	•	•		•
Data organization	•	•	•						•		•
Quality Assurance			•	•		•			•	•	•
Storage and backup	•	•	•	•	•	•			•	•	•
Security	•	•	•	•	•	•		•		•	
Responsibility	•	•	•		•	•		•	•	•	•
Budget	•	•	•					•			
IP	•	•	•	•		•		•	•	•	•
Legal requirements	•	•	•								
Access and sharing	•	•	•	•	•	•	•	•	•		•
Audience	•	•	•			•				•	
Selection and retention periods	•	•	•			•	•	•	•		•
Archiving and preservation		•	•	•		•	•	•			•
Ethics and privacy			•	•		•		•	•		•

REFERÊNCIAS

- Australia National University's [Information Literacy Program DPM Template](#) is a formatted template for any discipline.
- The Australian National Data Service created [Data Management Planning](#), a document that lists the questions that should be answered in a data management plan.
- The Digital Curation Centre created its [Data Management Plan Content Checklist](#) as "a comprehensive list of the details that researchers may be asked to include in such plans."
- The Finnish Social Science Data Archive's [Data Management Planning Website](#) lists questions that should be answered in a data management plan. It is aimed at social science researchers in particular.
- Geoscience Australia's [Guide to Preparation of Data Management Plans](#).
- MIT Libraries' [Data Management Webpage](#) provides a list of questions that should be answered in data management plan.
- The National Science Board's [Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century](#) is one of the foundational documents in the US' current push for data sharing. It gives broad guidelines for what should be included.
- The National Science Foundation Directorate for Engineering's [Data Management for NSF Engineering Directorate Proposals and Awards](#) is the first document to directly address the coming NSF requirement
- The Queensland University of Technology [QUT Data Management Checklist](#) is a highly structured, popuable template.
- The UK Rural Economy and Land Use Programme's [Data Management Plan](#) is a form that must be filled out by RELU award holders at the outset of their projects.
- The University of Melbourne's [Research Data Management Plan Template](#) is a te,plate aimed at university researchers.

Bons exemples

- <http://monash.edu/library/researchdata/about/strategy/>
- Monash University's [Research Data Management Strategy and Strategic Plan 2012-2015 \[pdf 196kB\]](#) was publicly released on 13 April 2012.

3. Skills and knowledge

By 2015 Monash University's researchers – including Higher Degree by Research students and, over time, Honours and undergraduate researchers – will have well-developed data management awareness, knowledge and skills. Data management skills will be seen as essential graduate skills that are necessary in the research sector and transferable to a range of other workplaces. The University will offer a range of professional development opportunities – both stand-alone and embedded in the curriculum, delivered through as many channels as possible – that meet the needs of researchers in different disciplines and at different career stages. The development needs and career paths of data management professionals in central units will also be addressed, and the University will be able to recruit, retain and develop high quality staff that contribute both to the University and national capability.

Goal	Initiatives	Key measures	Responsibility/key stakeholders
3.1 Develop the data management skills and knowledge of Monash researchers	<ul style="list-style-type: none"> Project to re-develop, expand and better coordinate the program of research data skills development opportunities available to Monash researchers Work with faculties and the new Monash Institute of Graduate Research to explore embedding research data management skills development in coursework curricula and professional development offerings 	<ul style="list-style-type: none"> Continuous improvement in program of learning opportunities, in terms of number and range of opportunities Increased attendance (e.g. enrolment / attendance statistics) High quality training (e.g. participant feedback, impact surveys) 	<ul style="list-style-type: none"> University Librarian MeRC eSolutions Director, Monash Institute of Graduate Research Associate Deans of Research and Research Training
3.2 Develop the skills and knowledge of professional staff <ul style="list-style-type: none"> Library staff Technical staff – MeRC and eSolutions 	<ul style="list-style-type: none"> Provide targeted development opportunities for professional staff Support opportunities for professional staff to network and participate in communities of practice – Monash, national, international Proactively address staff retention and career planning Promote research data management career paths at entry level through guest lectures, placements, and participation in curriculum review 	<ul style="list-style-type: none"> Continuous improvement in program of opportunities to gain skills and knowledge and to network with other data management professionals Retention of key staff Feedback from staff 	<ul style="list-style-type: none"> University Librarian CIO Director, MeRC
3.3 Integrate research data skills as part of research-led teaching	<ul style="list-style-type: none"> Integrate data management concepts and services into the Graduate Certificate of Academic Practice (GCAP) Map data management knowledge, skills and attributes to educational frameworks such as the Research Skill Development (RSD) framework, the Monash graduate attributes, Australian Qualifications Framework, professional accreditation schemes and other relevant frameworks. Develop action plan for increasing access to, and re-use of, research data by Honours and undergraduate researchers. 	<ul style="list-style-type: none"> GCAP contribution made in first semester 2012 and evaluated for inclusion in future years Skills mapping work completed by June 2012 Action plan developed by December 2012 	<ul style="list-style-type: none"> University Librarian Library Office of the PVC Teaching and Learning Monash Institute of Graduate Research MeRC

4. Research data management at Monash University by 2015

Further developing the statement of intent, aims for research data management can be grouped around five themes:

Excellence and impact

- More research data discoverable and available for re-use
- Re-use of Monash data contributes to formal and informal measures of research quality and impact
- Systems and policies help make research data available more quickly and easily

World class infrastructure

- Systems and facilities that support data management are expanding and improving
- Researchers make more use of these systems and facilities
- A range of institutional and discipline-specific needs are catered for
- Local infrastructure leverages national and international services and facilities
- Infrastructure supports the management of data, regardless of format

Skills and knowledge

- Researchers have the knowledge and skills they need to manage data well, and understand the benefits of making data discoverable and available for re-use
- Data management skills are seen as essential for research and transferable to other workplaces
- Research data contributes to the educational outcomes of students from an early stage in their academic career
- Data management professionals have career paths and development opportunities
- Professional development opportunities meet the needs of researchers from different disciplines and at different career stages

Integrity and professionalism

- Managing data well is seen as a key part of research integrity and professional practice
- Compliance with Section 2 of the *Code for Responsible Conduct of Research* is improved
- All Monash researchers understand their obligations and take practical steps - as individuals and teams - to improve how research data is managed
- Research data management advisory and technical services are increasingly coordinated and integrated

Leadership and collaboration

- Monash University leads and actively participates in global, national and regional research data initiatives
- The University is regarded as a partner of choice for collaborative work in this area
- Data management technologies developed at Monash University are successfully adopted by other organisations
- Monash University is seen by other institutions as an authoritative source of information and advice

Requisitos dos financiadores

Research Funders	Policy Coverage		Policy Stipulations					Support Provided			
	Published outputs	Data	Time limits	Data plan	Access/sharing	Long-term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	●	●	●	●	●	◐	○	●	○	◐	○
BBSRC	●	●	●	●	●	●	●	●	●	◐	●
CRUK	●	●	●	●	●	●	●	◐	●	○	○
EPSRC	●	●	●	◐	●	●	●	◐	○	○	●
ESRC	●	●	●	●	●	●	●	●	●	●	◐
MRC	●	●	●	●	●	●	○	◐	●	○	◐
NERC	●	●	●	●	●	●	●	●	●	●	◐
STFC	●	●	●	●	●	●	●	◐	●	◐	○
Wellcome Trust	●	●	●	●	●	●	●	●	●	◐	●

www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies

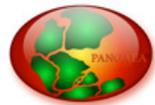
Papel da biblioteca – na prática...



Preparing data management plans for NSF grant applications

[National Science Foundation
http://youtu.be/Lc82pxxRkMo](http://youtu.be/Lc82pxxRkMo)

Observações finais (na prática)



All Water Sediment Ice Atmosphere

hermione

Search

hermione

more...

Not logged in (log in or sign up)

Always quote citation when using data!

486 datasets found on search for »hermione«

Show Map Google Earth Data Warehouse

<< PREV | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | NEXT >>

- Ramirez-Llodra, E; Company Claret, JB; Sardà, F et al. (2012):** Distribution of marine litter at bathyal and abyssal depths in the Mediterranean Sea
Related to: Ramirez-Llodra, E; Company Claret, JB; Sardà, F et al.: Distribution of marine litter at bathyal and abyssal depths in the Mediterranean Sea. *data submission PDI-1669, Progress in Oceanography*
Size: 0 datasets
Unpublished dataset #778322 - Score: 100% - Similar datasets
- Hunter, WR; Jamieson, AJ; Huvenne, VAI et al. (2013):** Sediment community responses to marine versus terrigenous organic matter in the Whittard canyon
Supplement to: Hunter, WR; Jamieson, AJ; Huvenne, VAI et al. (2013): Sediment community responses to marine vs. terrigenous organic matter in a submarine canyon. *Biogeosciences*
Size: 4 datasets
doi:10.1594/PANGAEA.803858 - Score: 16% - Similar datasets
- Grünke, S; Lichtschlag, A; de Beer, D et al. (2012):** Mats of psychrophilic thiotrophic bacteria associated with cold seeps of the Barents Sea
Supplement to: Grünke, S; Lichtschlag, A; de Beer, D et al. (2012): Mats of psychrophilic thiotrophic bacteria associated with cold seeps of the Barents Sea. *Biogeosciences*
Size: 26 datasets
doi:10.1594/PANGAEA.786576 - Score: 15% - Similar datasets
- Guilini, K; van Oevelen, D; Soetaert, K et al. (2011):** Results of an isotope tracer experiment with arctic deep-sea nematodes
Supplement to: Guilini, K; van Oevelen, D; Soetaert, K et al. (2010): Nutritional importance of benthic bacteria for deep-sea nematodes from the Arctic ice margin: Results of an isotope tracer experiment. *Limnology and Oceanography*
Size: 4 datasets
doi:10.1594/PANGAEA.772860 - Score: 14% - Similar datasets
- Grünke, S; Felden, J; Lichtschlag, A et al. (2013):** Niche differentiation among mat-forming, sulfide-oxidizing bacteria in chemosynthetic ecosystems of the Nile Deep Sea Fan (Eastern Mediterranean Sea)
Supplement to: Grünke, S; Felden, J; Lichtschlag, A et al. (2011): Niche differentiation among mat-forming, sulfide-oxidizing bacteria at cold seeps of the Nile Deep Sea Fan (Eastern Mediterranean Sea). *Geobiology*
Size: 9 datasets
doi:10.1594/PANGAEA.810017 - Score: 14% - Similar datasets
- Van Rooij, D; De Mol, P; Le Guilloux, F et al. (2012):** Multibeam echosoundings reveal physical mechanisms of deep-sea-bottom videos from the Bay of Biscay

<http://www.pangaea.de>

Observações finais (na prática)

The screenshot displays the OpenAIRE website interface. At the top, there is a navigation bar with links for 'Contact us', 'About', and 'Newsletter', along with a language selector (BG, CS, EE, EL, EN, ES, FR, HU, IT, LV, LT, NL, PL, PT, RO, SK, SV, TR) and a search box. The OpenAIRE logo and tagline 'Open Access Infrastructure for Research in Europe' are prominently displayed. A secondary navigation bar includes 'Home', 'Open Access in the EU', 'Participate Deposit, Provide Content', 'Discover Publications, Statistics', 'Get Support FAQ, Helpdesk, Guides', and 'My OpenAIRE My Deposits, My Alerts'. A social media section offers links to Twitter, LinkedIn, Facebook, YouTube, and RSS. A 'Sign In' button shows the user 'Hi Pedro Principe' with a 'logout' option.

The main content area is titled 'Project: HERMIONE' and includes the following details:

- Title:** Hotspot Ecosystem Research and Man's Impact on European seas
- Call:** FP7-ENV-2008-1
- Grant agreement number:** 226354
- Start date:** 01/04/2009
- End date:** 30/09/2012
- Special clause 39:** yes
- Scientific area:** Environment (including Climate Change)
- Programme:** SP1-Cooperation

A link for 'Detailed project information (CORDIS)' is provided. An 'App Box' on the right offers options to 'View publications details', 'Dynamically incorporate publications in your site (HTML)', 'View EC progress report (HTML)', and 'Download EC progress report (CSV)'. At the bottom of the project section are buttons for 'Claim publications' and 'Deposit OA'.

The 'Publications' section lists several research findings:

- Changes in polychaete standing stock and diversity on the northern side of Senghor Seamount (NE Atlantic)
- Chemosynthetic bacteria found in bivalve species from mud volcanoes of the Gulf of Cadiz
- Testing the protozoan hypothesis for Ediacaran fossils: a developmental analysis of Palaeopascichnus
- Limitations of microbial hydrocarbon degradation at the Amon mud volcano (Nile deep-sea fan)
- Diversity and distribution of cold-seep fauna associated with different geological and

A sidebar on the left titled 'In this section' contains links for 'Publications', 'Projects', 'Statistics', and 'Repositories'.

Observações finais (na prática)

zenodo

Research. Shared.

Search Communities Browse Upload Get started

Email

Password

Sign in

Search 731 records for:

euro-basin

Search

Any Collection

Any Author

Any Year

Showing records 1 to 10 out of 38 results.

1

01 January 2013 Journal article Open access

Capturing quantitative zooplankton information in the sea: Performance test of laser optical plankton counter and video plankton recorder in a *Calanus finmarchicus* dominated summer situation

View

Basedow, Sünne L. ; Tande, Kurt S. ; Norrbin, M. Fredrika ; Kristiansen, Stian A.

We compared two optical plankton counters, the Laser Optical Plankton Counter (LOPC) and the Video Plankton Recorder (VPR) for their abundance estimates of *Calanus finmarchicus* during an early summer situation (June 2008) in two North Norwegian fjords. [...]

Uploaded by Thomé on 23 May 2013.

2

View