



grey literature and data e-infrastructures

***14th International Conference on Grey Literature
Rome, 29 November 2012***

**Carlos Morais Pires
European Commission
DG CONNECT, unit C1
carlos.morais-pires[[@](mailto:carlos.morais-pires@ec.europa.eu)]ec.europa.eu**

Author's views do not commit the European Commission

presentation outline

- 1. grey literature... a bit of history**
- 2. e-infrastructures**
- 3. EC recent developments and future steps**

~14 billion years ago

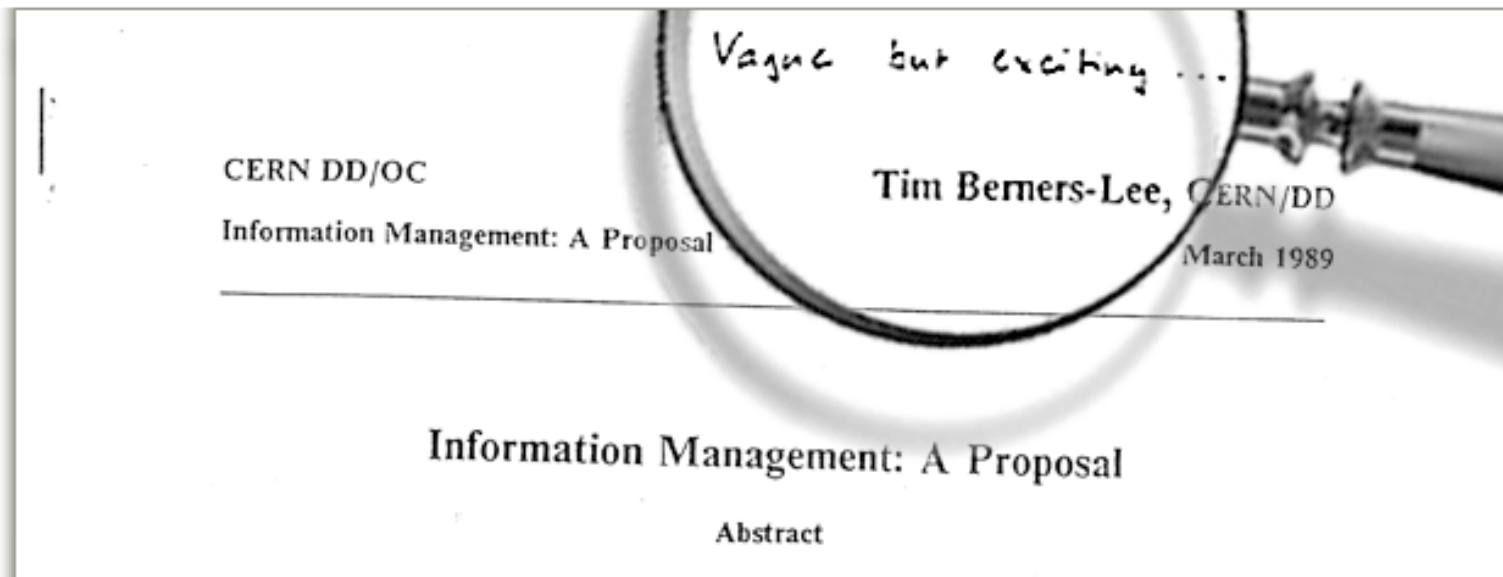
3-6 million years ago

1950... or 1990?

on-line research born... 20 years ago

In March 1989, Tim Berners-Lee submitted a proposal for an information management system to his boss, Mike Sendall. '*Vague, but exciting*', were the words that Sendall wrote on the proposal, allowing Berners-Lee to continue.

[Full text](#) of the proposal in html.



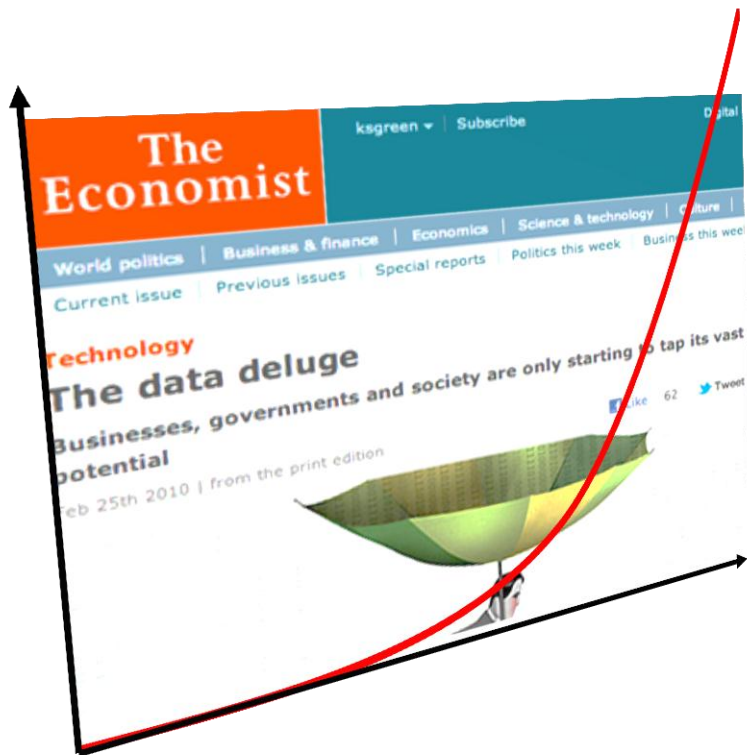
António Gedeão, 1950

A minha aldeia

**Minha aldeia é todo o mundo.
Todo o mundo me pertence.
Aqui me encontro e confundo
com gente de todo o mundo
que a todo o mundo pertence.**

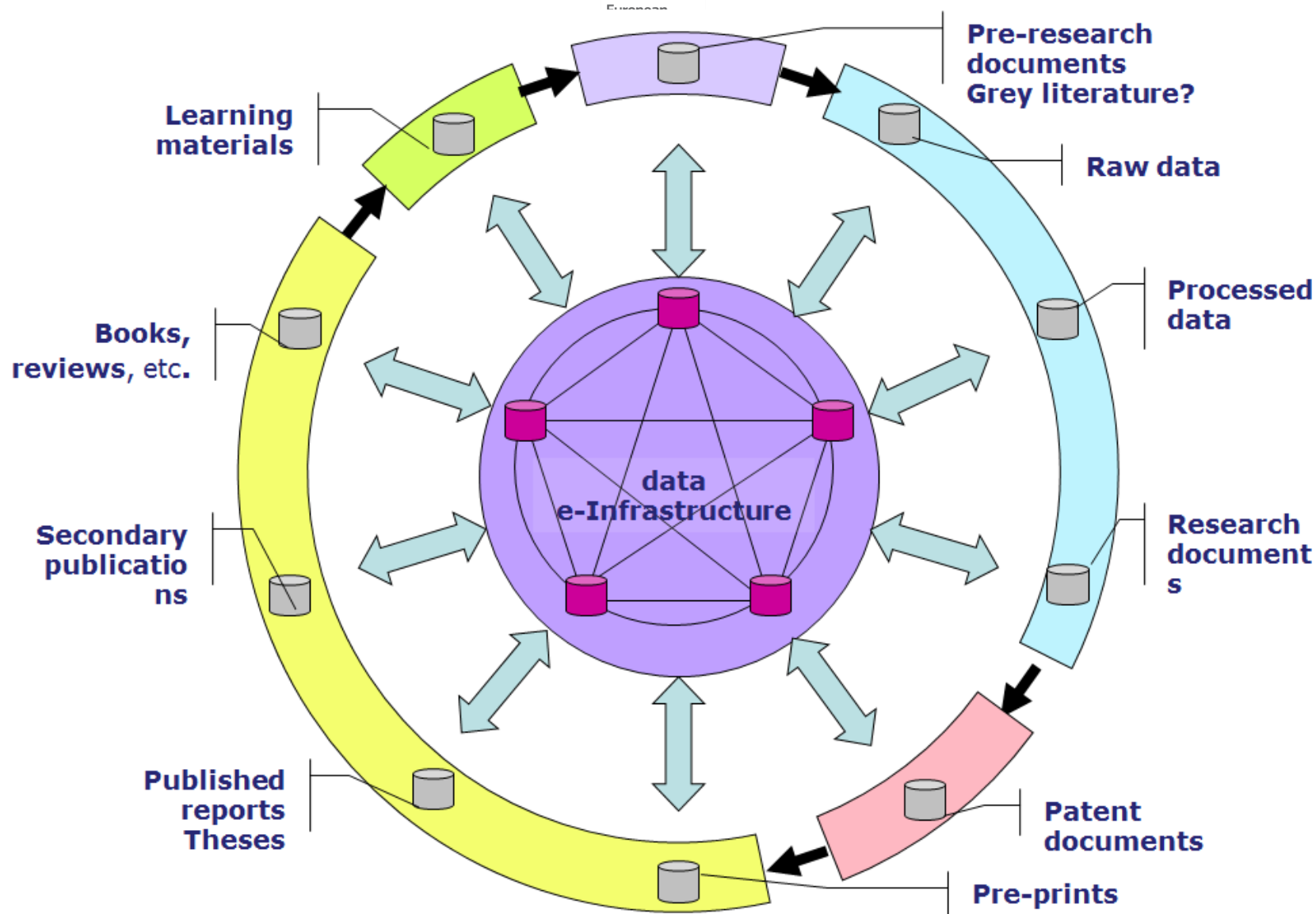
**Bate o sol na minha aldeia
com várias inclinações.
Angulo novo, nova ideia;
outros graus, outras razões.
Que os homens da minha aldeia
são centenas de milhões.**

preservation, volumes, costs, etc



The cost of data in \$/€/£ per byte from (*)	
"Reliable" code / Curated data	10
"Production" code	1
Book	1
[Movie]	10^{-1}
Big physics (e.g. LHC) data	10^{-3}
	10^{-7}

(*) Peter Buneman, Univ . Edinburgh,
Linz April 2006,



Digital Agenda
101011101110000100 2010-2020
for Europe



Neelie Kroes

Digital Agenda

Digital (information) single market



Open Science means optimal sharing of research results and tools such as publications, research data, software, educational resources and infrastructures across institutional, disciplinary and national boundaries.

reports and studies: european

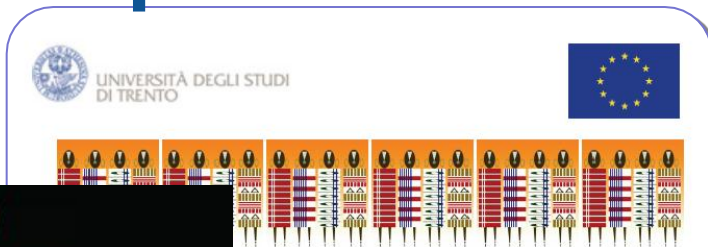


Global Research Data Infrastructure
Towards a 10-year vision for global research
data infrastructures

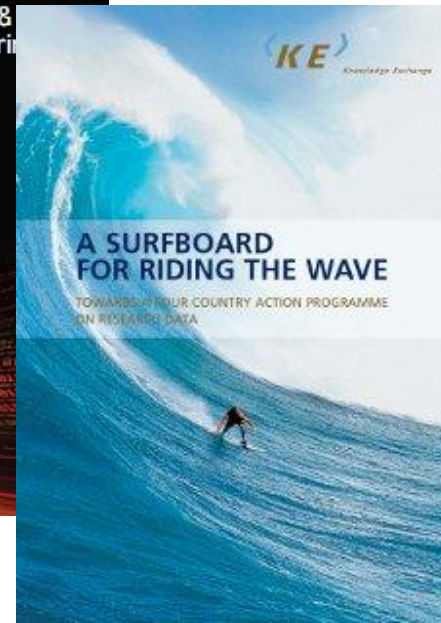
Technological & Organisational Aspects
of a Global Research Data Infrastructure
A view from the experts

www.grdi2020.eu | info@GRDI2020.eu

Funded under the Seventh Framework Programme (FP7) - Infrastructures,
"Capacities - Research Infrastructures" - Project Number: 246682



European Commission (e-Infrastructures Unit)
invites you to attend the workshop on

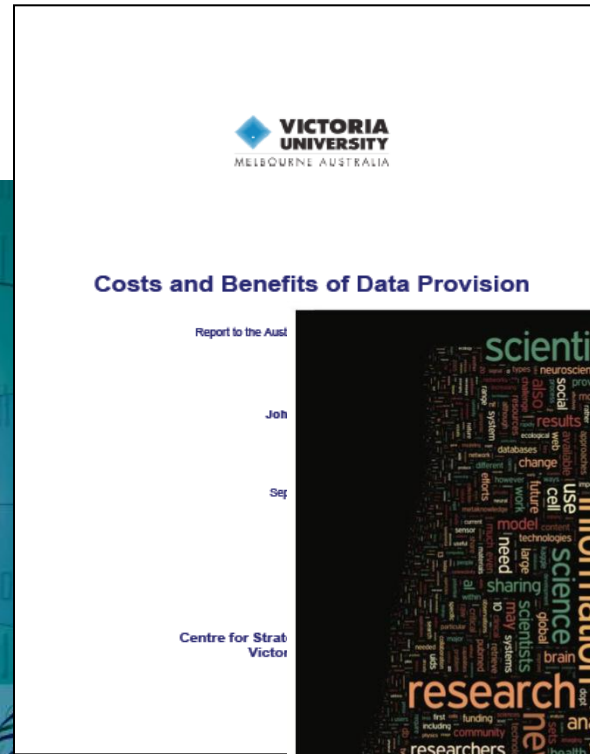
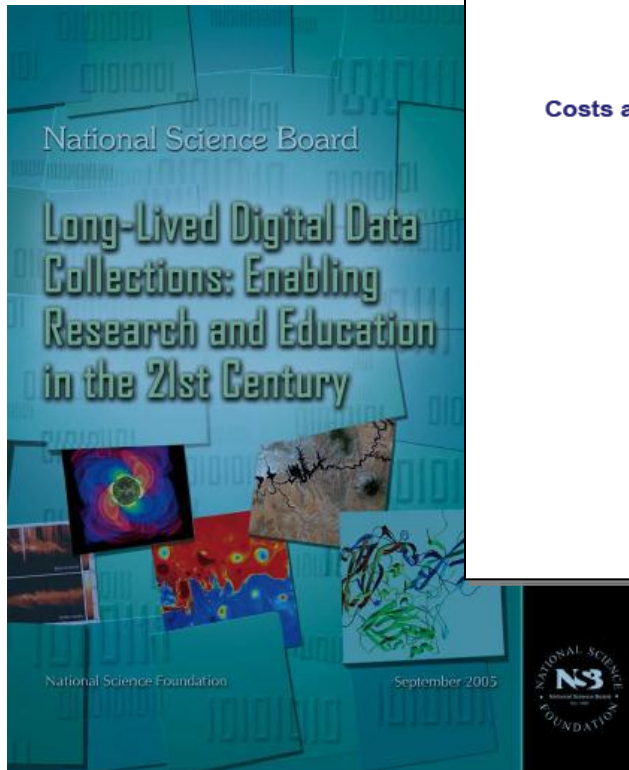


THOR
and access

Digital preservation of
research data in Europe

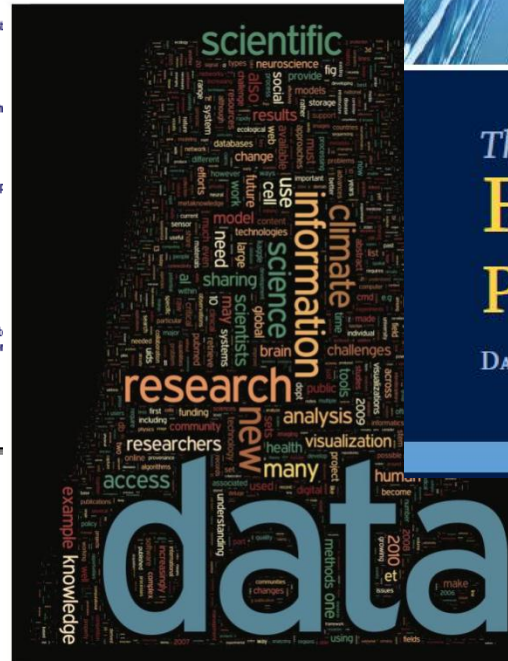


White Paper



DATA-INTENSIVE SCIENTIFIC DISCOVERY

EDITED BY TONY HEY, STEWART TANSLEY, AND KRISTIN TOLLE



Open Infrastructures for Open Science

Open Scientific Content

data, computational resources and software resulting from public funded research

Open Culture

career systems should support and reward those who participate in the culture of sharing

Open Infrastructures

reliable, high-performance and economically efficient infrastructures

"To make progress in science, we need to be open and share. [...] With the right infrastructure and the right approach, we can bring on a new age of scientific practice and discovery"

Neelie Kroes on "Open infrastructures for Open Science", Rome 11 April 2012



Open Infrastructures for Open Science

COM and REC on Scientific Information, July 2012

Open Access,

Long term preservation,

Capacity building with data infrastructures



recommendation:

[...] hereby recommends that member states

COMMISSION RECOMMENDATION

on access to and preservation of scientific information

THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Article 292 thereof,

Whereas:

- (1) The Communication from the Commission Europe 2020¹ puts forward the development of an economy based on knowledge and innovation as a priority.
- (2) The targets set by the Europe 2020 strategy are given in more detail in particular in the Flagship Initiatives 'Digital Agenda for Europe'² and 'Innovation Union'³. Among the actions to be taken under the 'Digital Agenda', publicly funded research should be widely disseminated through

open access publication of scientific data and papers. The 'Innovation Union' initiative calls for a European Research Area (ERA) framework to be set up to help remove obstacles to mobility and cross-border cooperation. It states that open access to publications and data from publicly funded research should be promoted and access to publications made the general principle for projects funded by the EU research Framework Programmes.

- (3) On 14 February 2007, the Commission adopted a Communication on scientific information in the digital age: access, dissemination and preservation⁴, accompanied by a staff working paper. This provided an overview of the state of play in Europe regarding scientific publishing and the preservation of research results, examining relevant organisational, legal, technical and financial issues.
- (4) The Communication was followed in November 2007 by Council Conclusions on scientific information in the digital age: access, dissemination and preservation. The Conclusions invited the Commission to experiment with open access to scientific publications resulting from projects funded by EU research framework programmes and included a set of actions to be undertaken by the Member States. There have been advances in some of the areas dealt with in the Conclusions, but not all targets have been met and progress has been uneven among Member States. EU action is needed to make the most of Europe's research potential.
- (5) 'Open access policies' aim to provide readers with access to peer-reviewed scientific publications and research data free of charge as early as possible in

¹ COM (2010) 2020 final of 3.3.2010, available at: <http://eur->

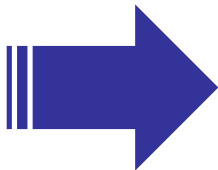
towards horizon2020



Preparation Draft Roadmap

Roadmap Discussion

Actions in H2020 documents



April 2012
H2020
Consultation

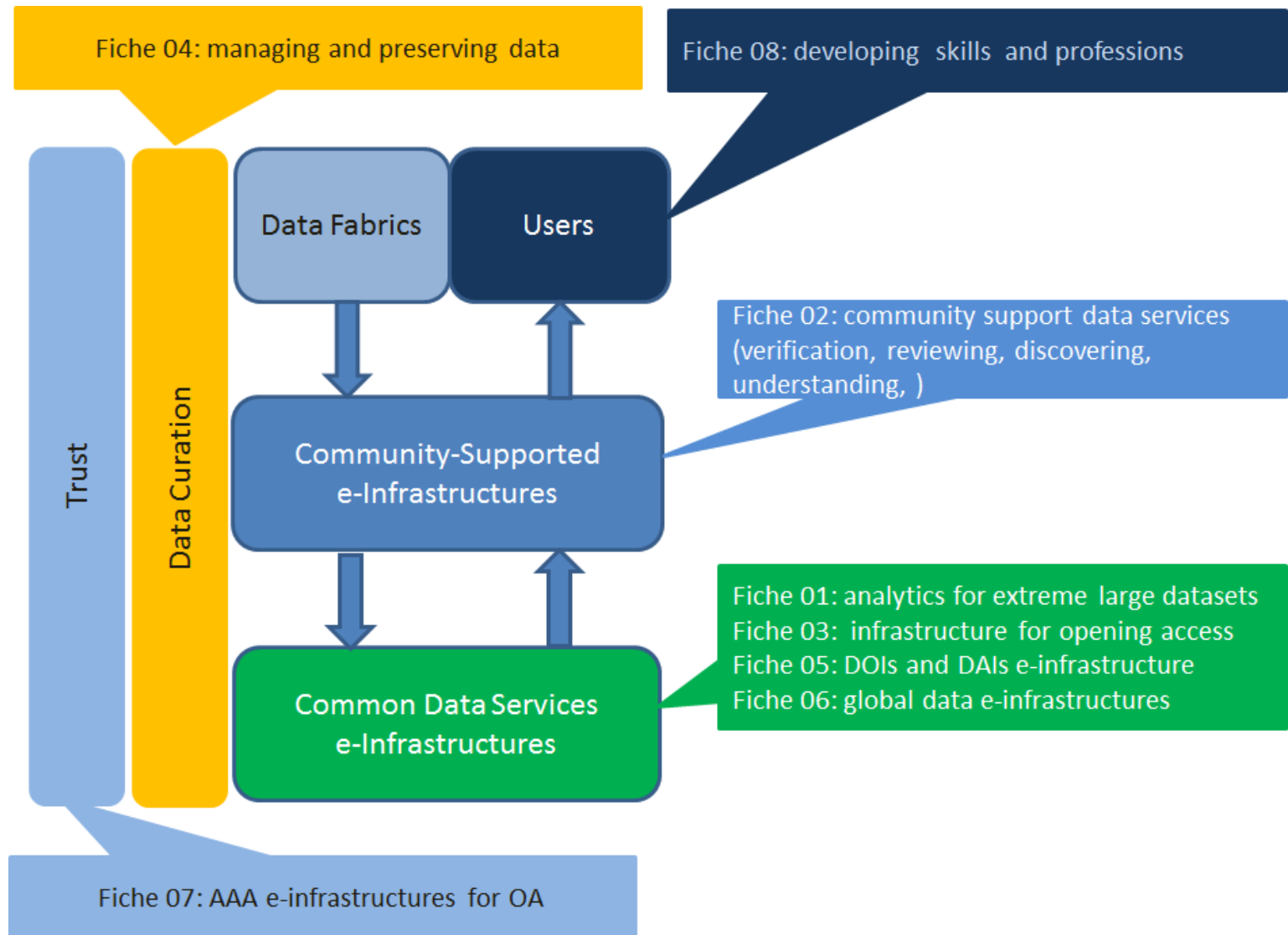
December 2012
Dissemination

February/March 2013

Framework for Action



1. analytics for very large research datasets
2. community support data services
3. infrastructure for Open Access
4. management, preservation, curation & persistence
5. discovery, provenance and persistence of data
6. towards global data e-infrastructures
7. Authentication and Authorisation e-infrastructures
8. developing skills and new professions





Funders Perspective on Research Data Alliance

initial group of funders

Alan Blatecky (NSF), Carlos Morais Pires (EC)

***EUDAT conference
Barcelona, October 24, 2012***



IGoF and the RDA

Why/ Funders Motivation

How/ How do we see the process

What/ What do we expect

Four Threats to Establishing a Global Data Research Infrastructure

Not understanding the critical importance and the need to share data for next century science and education

Not understanding the urgency to address and create a global data infrastructure now

Relying on additional workshops, conferences, committees and so forth to study and provide more recommendations

Waiting for standards to be approved that will enable data sharing, interoperability, and support the entire data life cycle

IGoF Motivation

G8+O5 and Data infrastructures

South Africa (Nov 2011) and
Hamburg (April 2012)

Technical/Cultural

Creation of data

Curation & Preservation of data

Access to data

Computing infrastructures

International governance

vision that research data will

Unmanaged	→	Managed
Disconnected	→	Connected
Invisible	→	Findable
Single-use	→	Reusable



Transform research and usher in a
new era of discovery and innovation

Non Government Structures (NGS) Funded to support RDA

US:

Fran Berman – RPI

Bill Michener – DataOne

Beth Plale – Indiana

Sayeed Choudhury – Johns Hopkins

Australia:

Ross Wilkinson – ANDS

Andrew Treolar - ANDS

Europe:

Leif Laaksonen (iCORDI/CSC)

Peter Wittenburg (iCORDI/Max Planck Institute)

Juan Bicarregui (iCORDI/STFC)



Initial Council

US:

Fran Berman – RPI

Australia:

Ross Wilkinson – ANDS

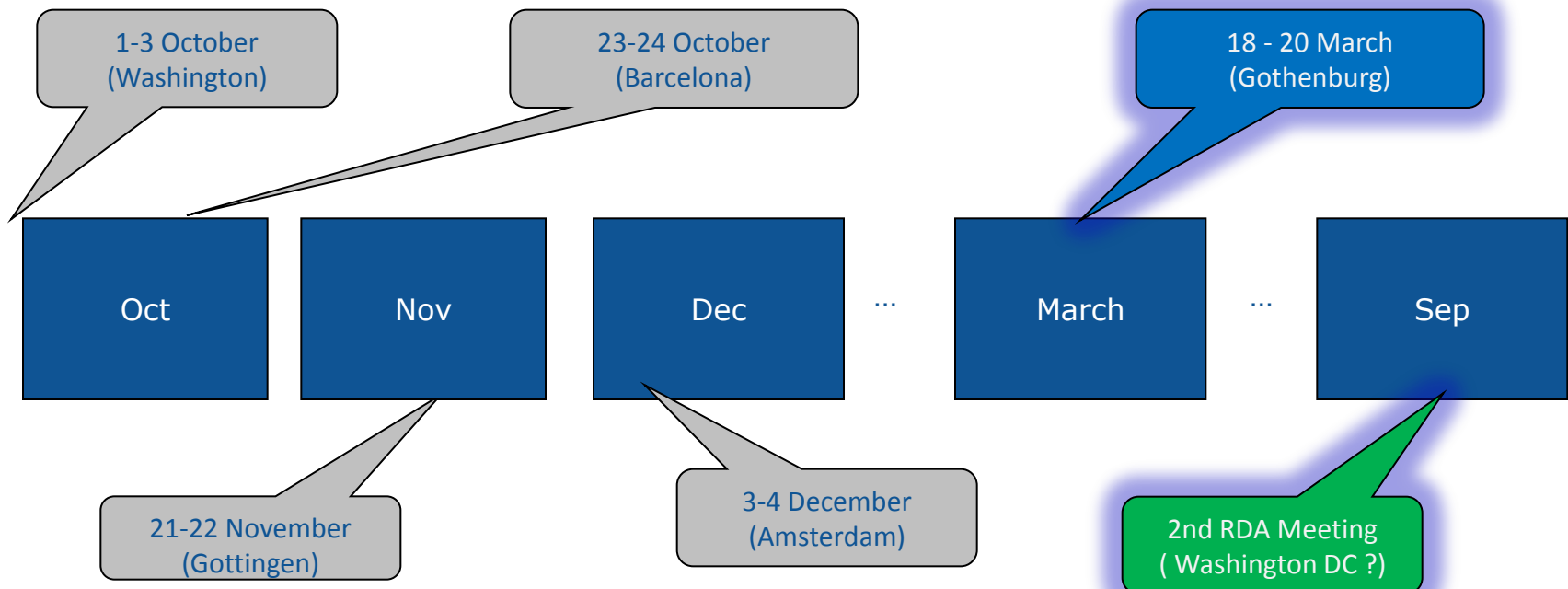
Europe:

John Wood, (iCORDI/ Commonwealth Universities)

4 more At-Large members to be appointed by March meeting
and will represent other sector stakeholders

Opening up RDA

Involve other science agencies in global research data infrastructures;
Leverage the G8+05 working group on data infrastructures;
First international meeting of RDA will be held in March 2013

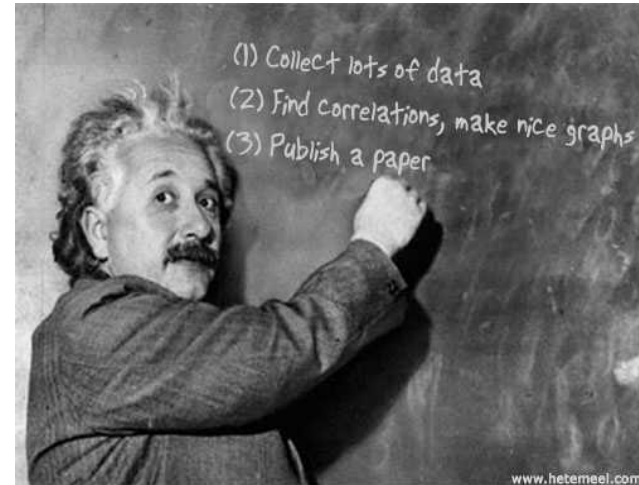


accessing scientific information

We don't know how scholar communication will adapt to new paradigms bringing closer human and machine readable information...

Opportunities for innovation in publishing

publication + data + software



scientific information “continuums”

experimental data and publications (new paradigm)

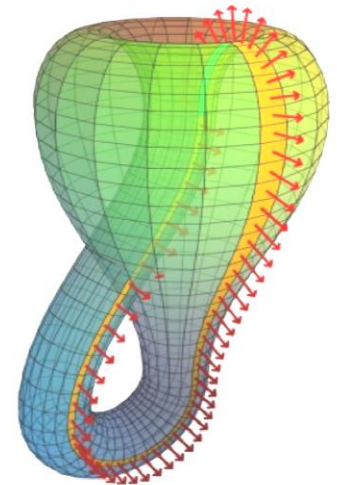
humans and computers (e-infrastructure)

different scientific disciplines (access, multidisciplinary)

past, present and future (preservation)

research and education (public mission)

different institutions (organisation)



Klein Bottle
<http://plus.maths.org/issue26/index.html>



Thank You!

Carlos Morais Pires

[carlos.morais-pires\(at\)ec.europa.eu](mailto:carlos.morais-pires(at)ec.europa.eu)