# Data e-infrastructures in Horizon 2020

Brussels, 7 May 2014

Carlos Morais Pires
European Commission
e-Infrastructures, DG CNECT.C1

Author's views do not commit the European Commission

# summary

- strategic relevance of research infrastructures
- e-infrastructures: moving, processing and managing data
- coherent strategies and opportunities to implement them

# Research Infrastructures

*"Men of science […] could formerly work in isolation as writers still can.*

*Cavendish and Faraday and Mendel depend hardly at all upon institutions and Darwin only in so far as the government enabled him to share the voyage of the Beagle.*

*But this isolation is a thing of the past.*

*Most research requires **expensive apparatus** […]. Without facilities provided by a government or a university, few men can achieve much in modern science."*

*from Bertrand Russell in BBC Reith Lectures,1949*

# data has been and remains key to science

Need for "**expensive apparatus**" is something that modern science intensified (need for more powerful telescopes, light sources, research boats, geological probes etc)

Intrinsic to the ambition that European researchers remain at the vanguard of scientific discovery

But there is something about research data:

**information opens new worlds for science**

# research logic machines

**Research Data** collected at observation or experimentation phase were registered in the **scientists notebooks**, which used to be paper books

Now research data is stored in digital form. Easier to be processed by "**logic machines**" programmed with complex models able to dig into the data

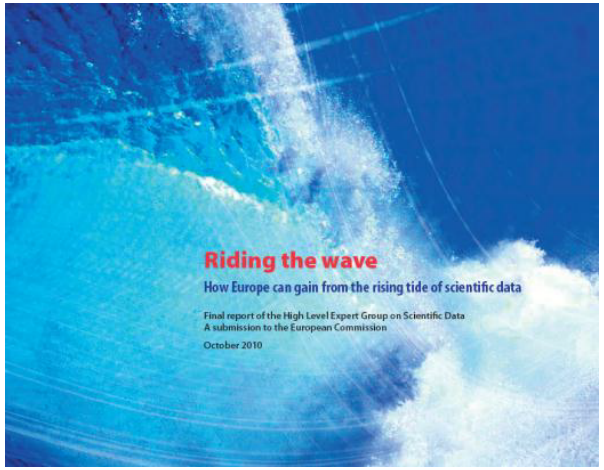
Logic machines are made of **human scientific knowledge and creativity**, **software** and the underlying **hardware**

Scientist notebooks can now be **linked** to a huge amount of other **data resources** (including scientific papers), **computers** with unprecedented capacity, eventually connected to **global networks**

# Europe riding the research data wave

**Vision**: "data e-infrastructure that supports seamless access, use, re-use, and trust of data.

In a sense, the physical and technical infrastructure becomes invisible and the **data themselves become the infrastructure** a valuable asset on which science, technology, the economy and society can advance".

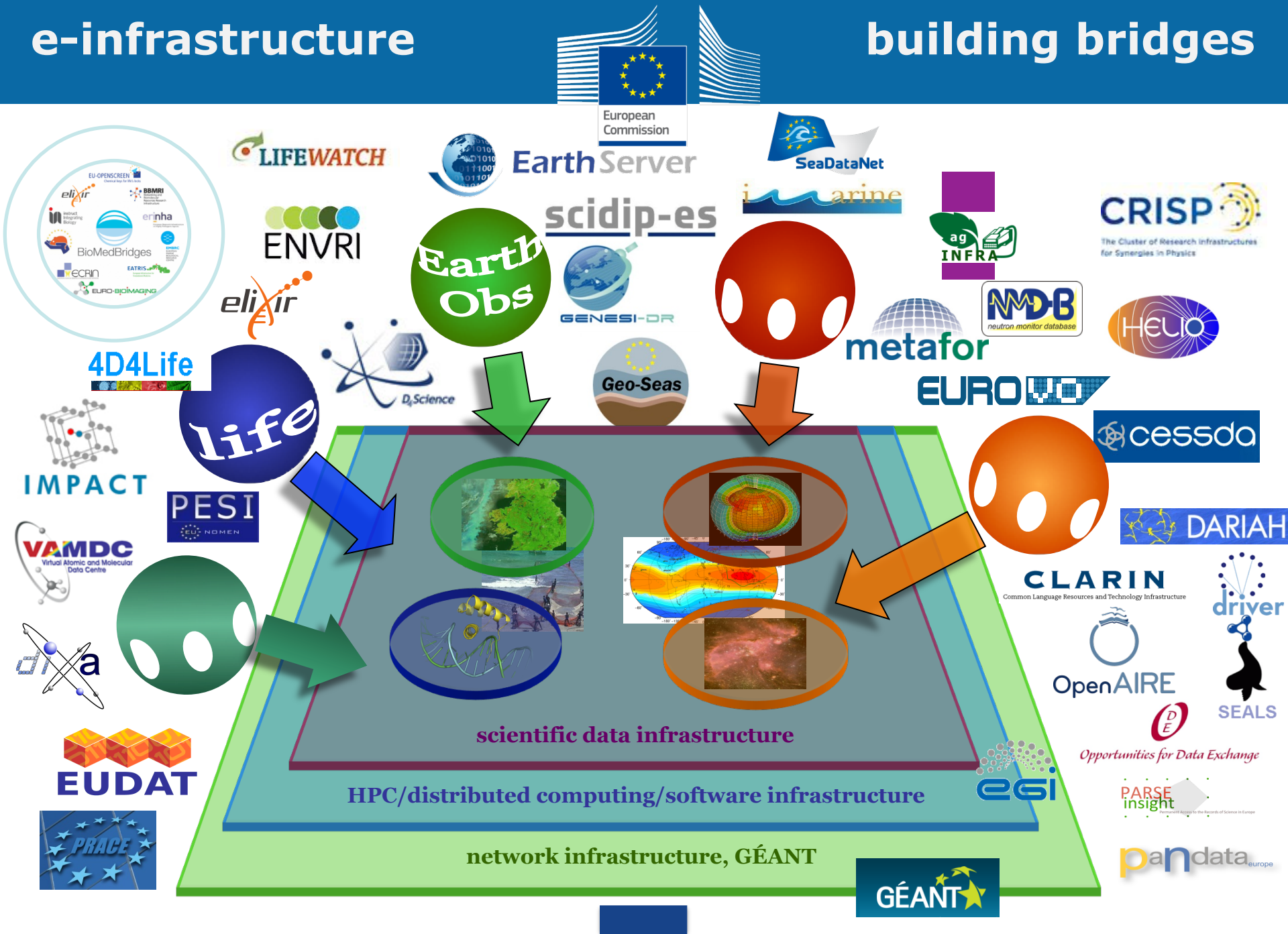



*The High Level Expert Group on Scientific Data presented Riding the Wave in October 2010*

Russell's quote could be extended:

"**without data and computing infrastructures few men can achieve much in 21st century's science**"

# e-infrastructure building bridges



# who is involved?

**data generators**

research projects, big research infrastructure, installations or medium size laboratories, simulation centres, surveys or individual researchers

**discipline-specific data services**

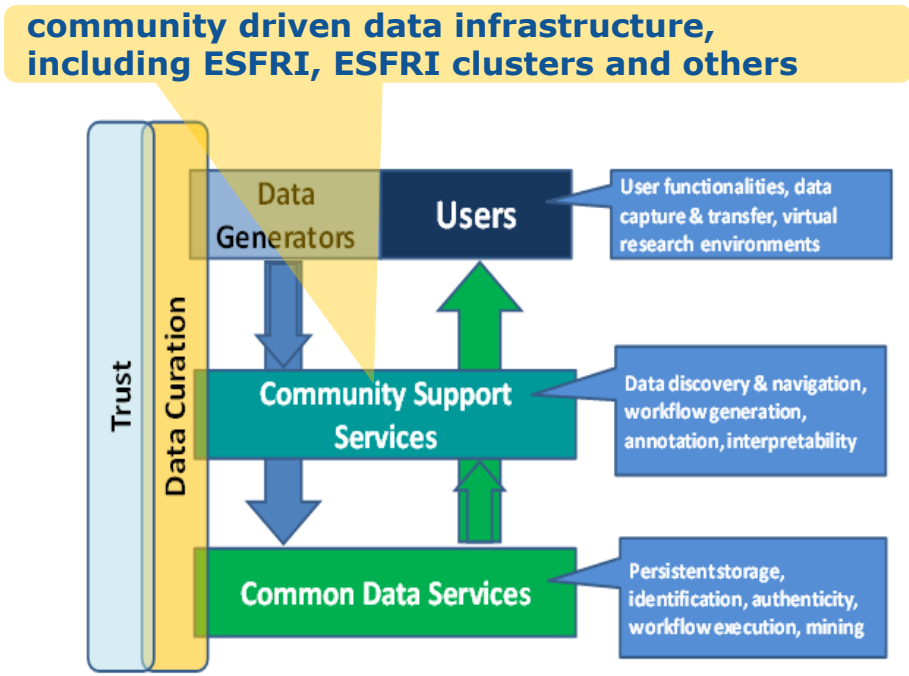
providing data and workflows as a service

**generic common data services**

computing centres, libraries,…

**researchers as users**

using (and producing) data for science and engineering



# issues to be addressed (policy framework)

Are publically funded **research data a public good**?
How do we ensure **preservation** and **access**?
How to we make data **discoverable** and **exchangeable**?
How to ensure **integrity** and **reliability** of data?
How do we ensure appropriate **recognition**?
How do we manage **intellectual property**?
How do we deal with **privacy** in the research context?
How do work the long term funding and **cost/benefit**?
How to work at **European and global** levels?
How to foster **cooperation with developing countries**?
etc…

# Policy context

# Open Science

**A Reinforced European Research Area Partnership for Excellence and Growth,** COM(2012) 392 – July 2012

**Towards better access to scientific information: boosting the benefits of public investments in research,** COM(2012) 401 final - July2012

**Commission, Recommendation on access and preservation of scientific information,** C(2012) 4890 final – July 2012


## Horizon 2020

- **Open Access to Scientific Publications**
  - OpenAIRE supporting infrastructure
- **Pilot on research data**
  - Data Management Plan is required

# issues to be addressed (e-infrastructure)

The EC in coordination with EU Member States is looking after research data as an infrastructure

As a valuable and a strategic resource, research data opens at least three key issues to be addressed(*):

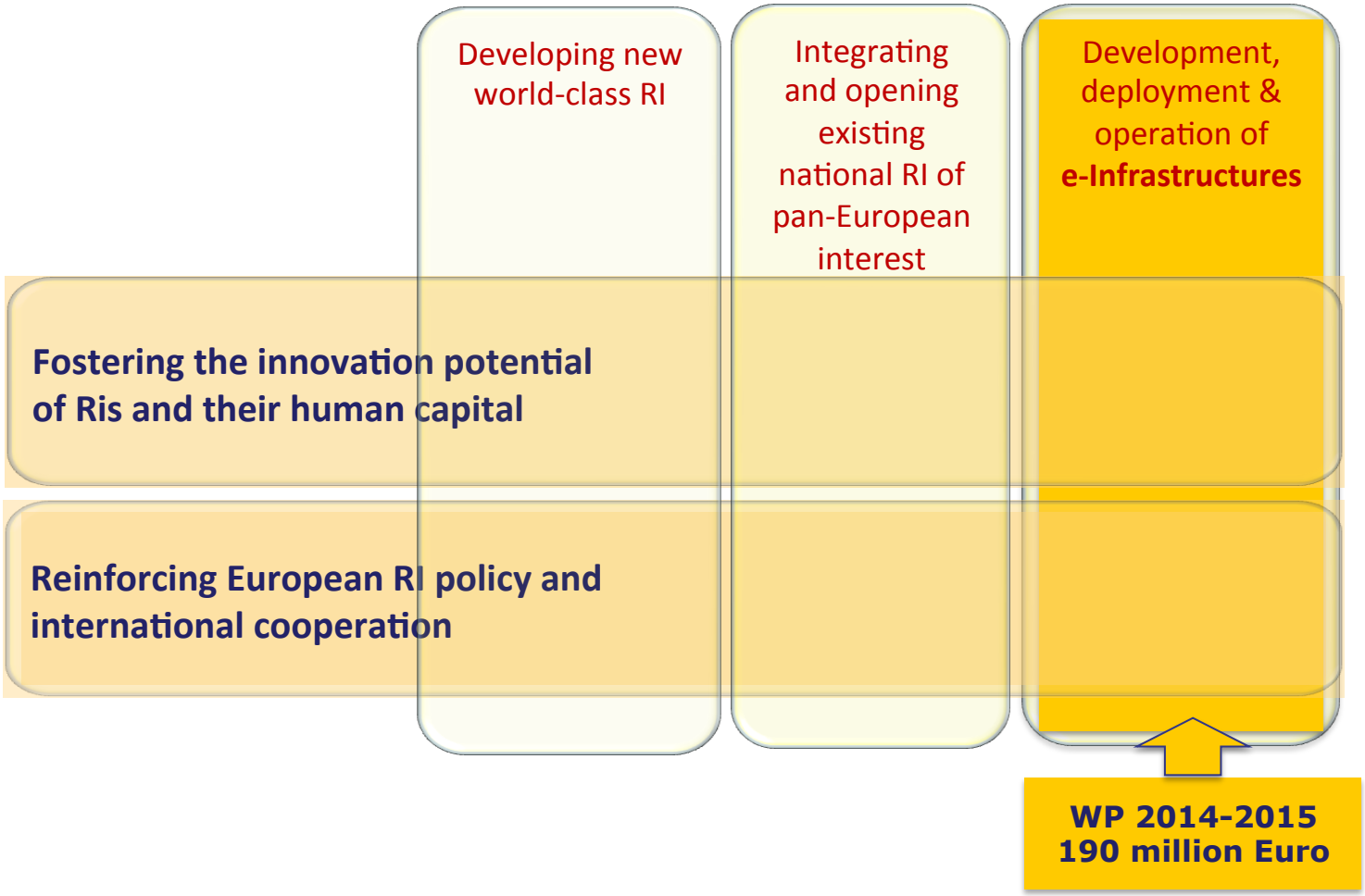
- How data can be networked
- How to envision and set up data governance on a global scale
- How the EU can play a leading role in helping start and steer this global trend

*(*) Fred Friend, Jean-Claude Guédon Herbert van Sompel "Beyond Sharing and Re-using: Toward Global Data Networking"*

# Research Infrastructures

European Commission

Horizon 2020



# RESEARCH INFRASTRUCTURE (E-INFRASTRUCTURE HIGHLIHGTED)
## Work Programme 2014-2015

**CALL 1**
DEVELOPING NEW WORLD CLASS INFRASTRUCTURES

- DESIGN STUDIES
- SUPPORT TO PREPARATORY PHASE OF ESFRI PROJECTS
- SUPPORT TO THE INDIVIDUAL IMPLEMENTATION AND OPERATION OF ESFRI PROJECTS
- SUPPORT TO THE IMPLEMENTATION OF CROSS-CUTTING INFRASTRUCTURE SERVICES AND SOLUTIONS FOR CLUSTER OF ESFRI AND OTHER RILEVANT RESEARCH INFRASTRUCTURE INITIATIVES IN A GIVEN THEMATIC AREA

**CALL 2**
INTEGRATING AND OPENING RESEARCH INFRASTRUCTURES OF PAN-EUROPEAN INTEREST

- INTEGRATING AND OPENING EXISTING NATIONAL AND REGIONAL RESEARCH INFRASTRUCTURES OF PAN-EUTROPEAN INTEREST

**CALLS IN 2014**
**DEADLINES SEPT 2014 AND JAN 2015**
**INITIATIVES STARTING IN 2015 UNTIL 2018**

**CALL 3**
**E-INFRASTRUCTURES**

- MANAGING, PRESERVING AND COMPUTING WITH BIG RESERACH DATA
- E-INFRASTRUCTURES FOR OPEN ACCESS
- TOWARDS GLOBAL DATA E-INFRASTRUCTURES: RESEARCH DATA ALLIANCE
- Pan-European High Performance Computing infrastructure and services
- Centres of Excellence for Computing applications
- Network of HPC Competence Centres for SMEs
- PROVISION OF CORE SERVICES ACROSS E-INFRASTRUCTURES
- RESEARCH AND EDUCATION NETWORKING – GEANT
- E-INFRASTRUCTURES FOR VIRTUAL RESEARCH ENVIRONMENTS (VRE)

**CALL 4**
SUPPORT TO INNOVATION, HUMAN RESOURCES, POLICY AND INTERNATIONAL COOPERATION FOR RESEARCH INFRASTRUCTURES

- INNOVATION SUPPORT MEASURES
- INNOVATIVE PROCUREMENT PILOT ACTION IN THE FIELD OF SCIENTIFIC INSTRUMENTATION
- STRENGTHENING THE HUMAN CAPITAL OF RESEARCH INFRASTRUCTURES
- NEW PROFESSIONS AND SKILLS FOR E-INFRASTRUCTURES
- POLICY MEASURES FOR RESEARCH INFRASTRUCTURES
- INTERNATIONAL COOPERATION FOR RESEARCH INFRASTRUCTURES
- E-INFRASTRUCTURE POLICY DEVELOPMENT AND INTERNATIONAL COOPERATION
- NETWORK OF NATIONAL CONTACT POINTS

# Research Data Alliance: a funder' perspective

Societal challenges of our time transcend borders

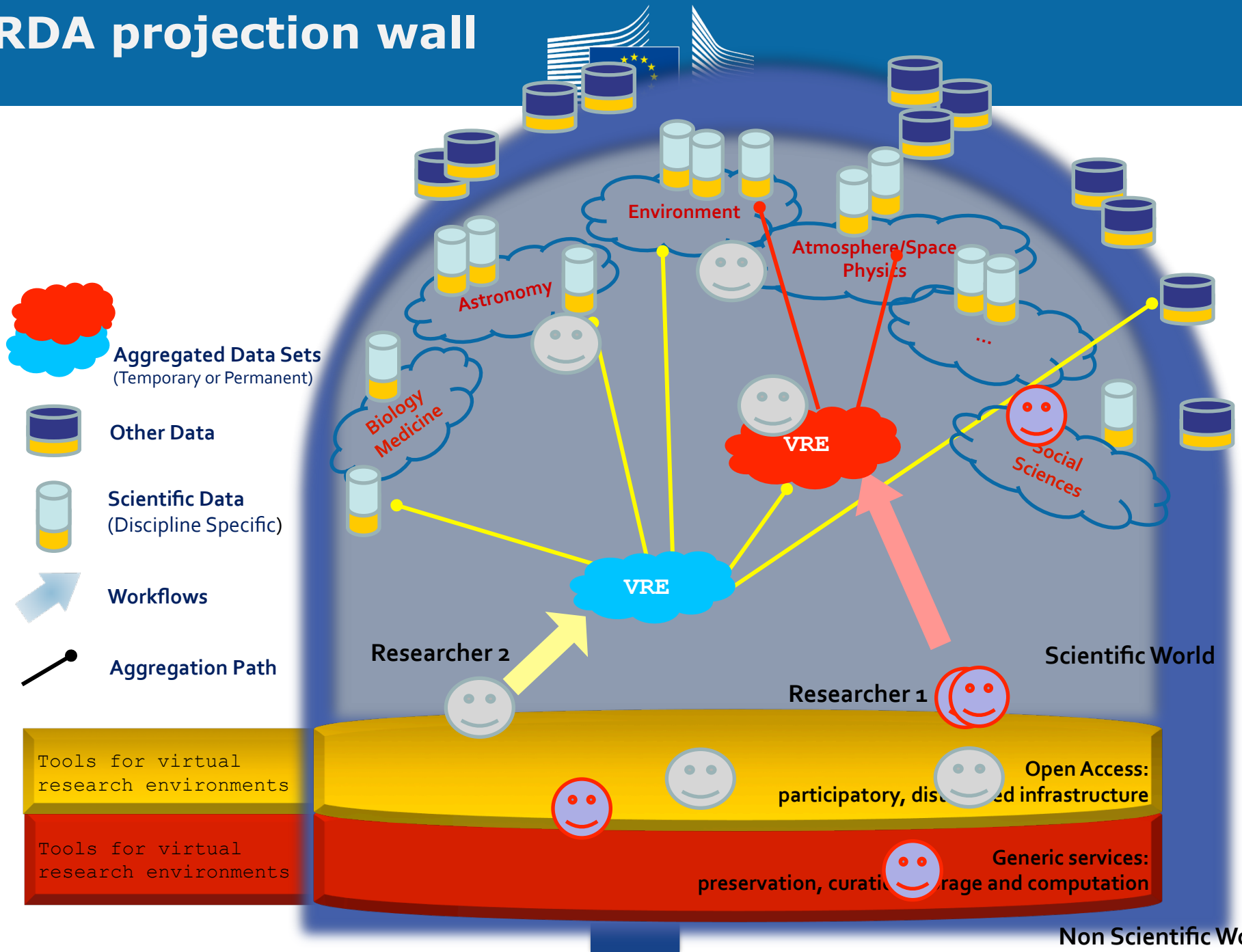
Data and computing intensive science is made of global collaborations

Research data are global

The European Commission has been supporting the set-up of the Research Data Alliance (RDA) to enable data exchange on a global scale

The initial phase of RDA has been supported by the collaboration between the European Commission, the US National Science Foundation and National Institute of Standards and Technology and the Australian Ministry of Research

# RDA projection wall



# take five

**5 principles** describing the benefits of a global research data infrastructure (G8+O6)

Data is:

**Discoverable** – IDs, Descriptive Metadata, ...

**Accessible** – Acknowledgment, License, Terms of Use, Intellectual Property, Legal ...

**Understandable** – Semantics, Analysis, Quality, Language translation ....

**Manageable** – Responsibility, Costs, Preservation ...

**People** (Usable) - Workforce, Cultural, Training, ...

# Final remarks

**Data e-Infrastructures** increase scope, depth and economies of scale of the scientific enterprise

**Horizon 2020 provides tools and opportunities** addressing data and computing e-infrastructures

If taken with appropriate resources and critical mass, can **project Europe into the new world of data driven science**

The objective is to **combine the expertise** of scientific communities with the expertise of ICT communities capable of exploring the limits of high bandwidth communication, high-performance computing, open scientific software and virtual research environments

**Carlos Morais Pires**
**carlos.morais-pires(at) ec.europa.eu**

# Thank you!