

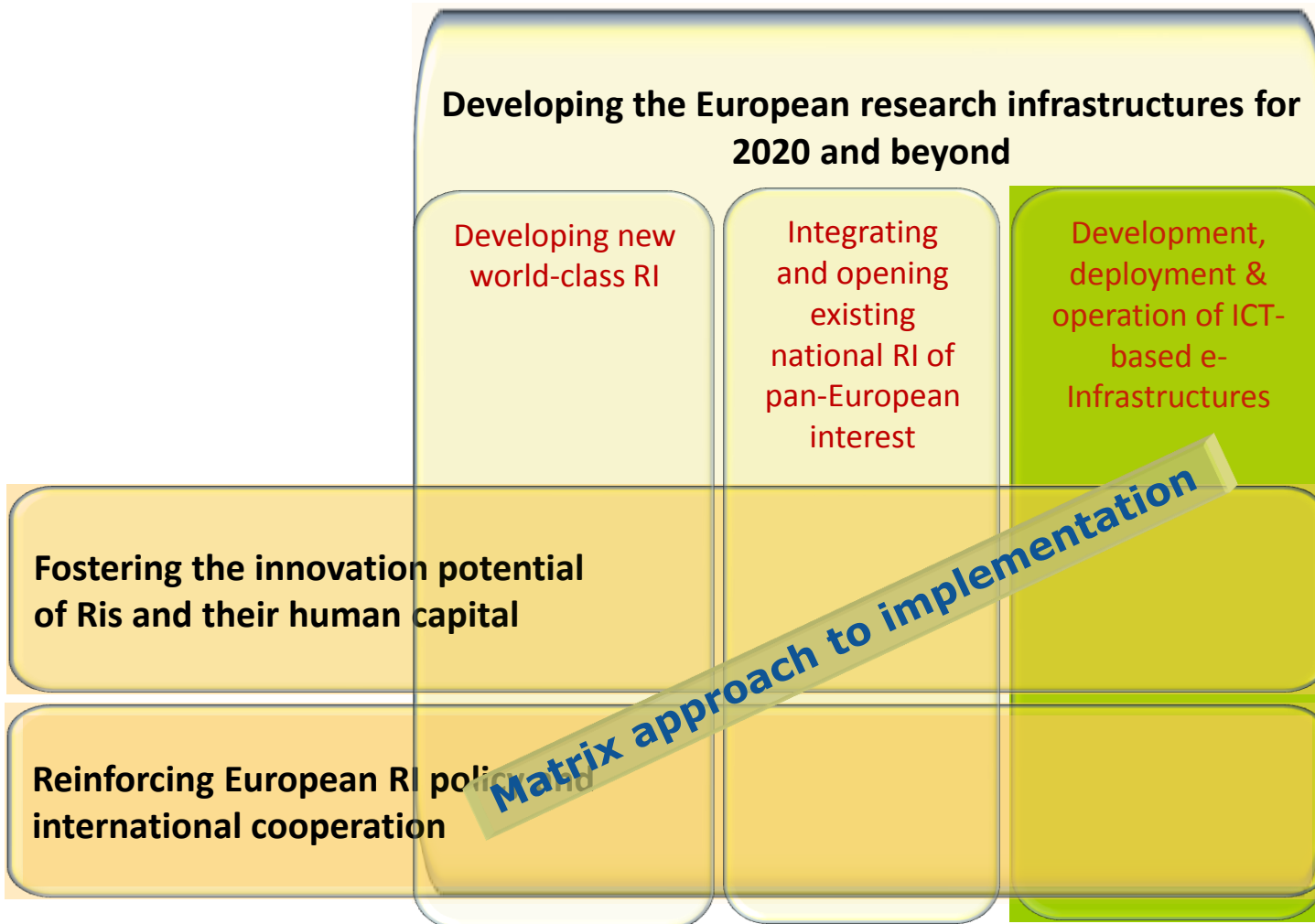
# e-Infrastructures in Horizon 2020

## e-Infrastructures for data and computing



**Kostas Glinos**  
**Head of Unit, e-Infrastructure**  
**European Commission – DG CONNECT**

# Research Infrastructures in Horizon 2020





# **e-infrastructure approach in Horizon 2020**

## **Transversal**

**Cutting across disciplines and sectors**

## **Support tomorrow's science**

**Open science, open access, best solutions**

## **Enabling innovation everywhere**

**Developing and testing innovative solutions**

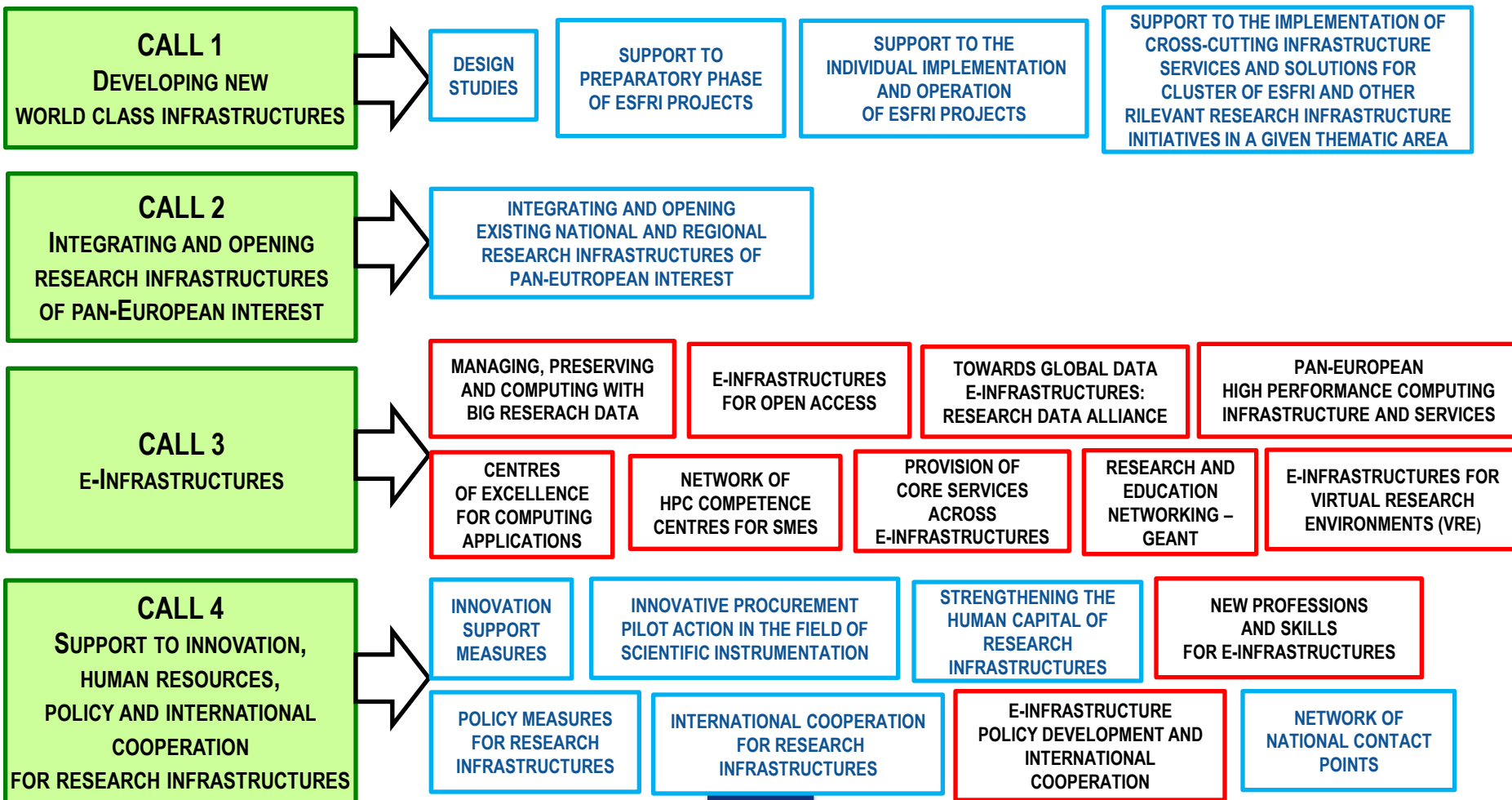
**Servicing industry and SMEs**

**Spinning out technologies**

## **Skills development across all actions**



# RESEARCH INFRASTRUCTURE Work Programme 2014-2015

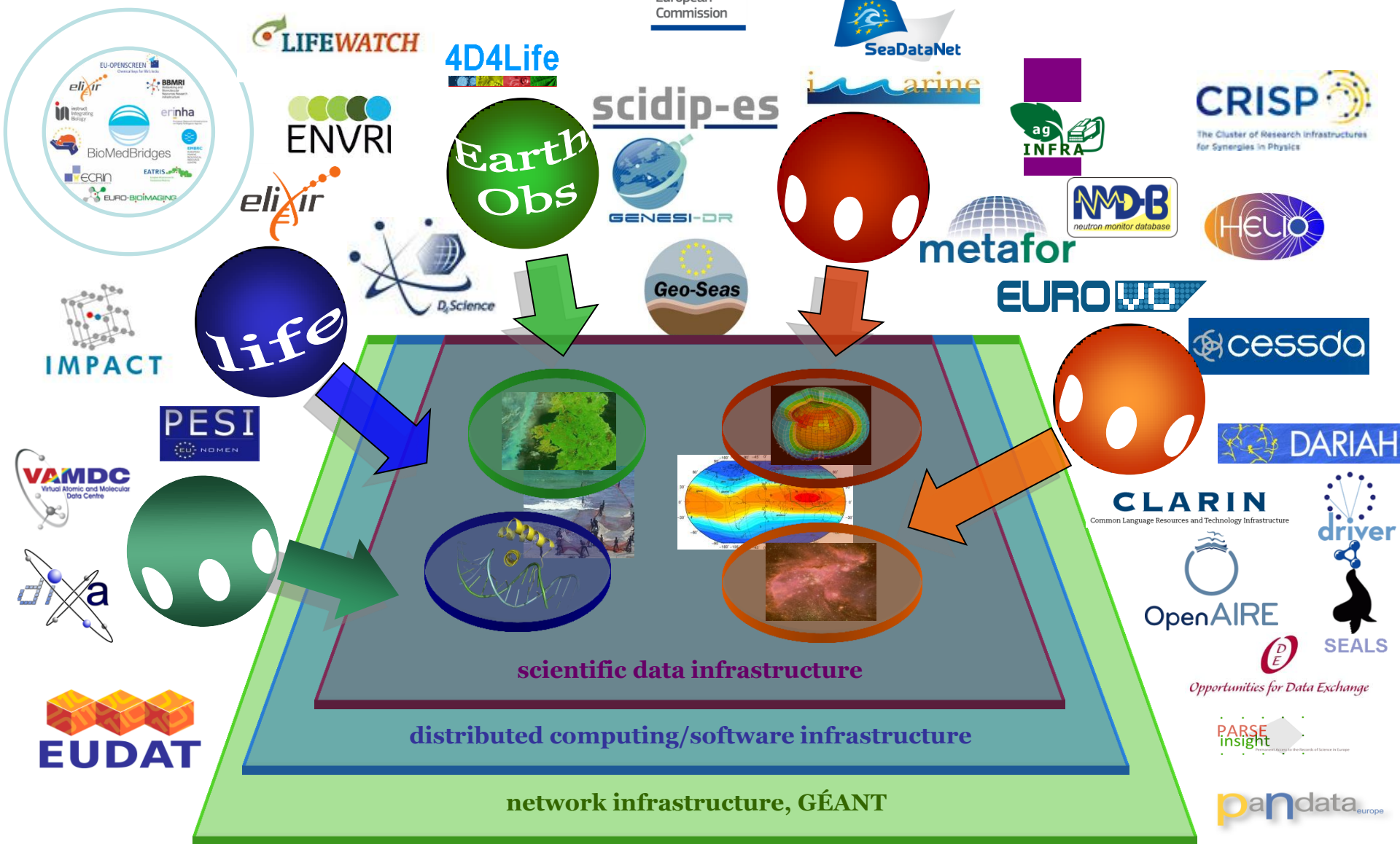


# **Examples of past (FP7) work on computing and data e-infrastructure**



# Data e-infrastructure in FP7 bridging islands

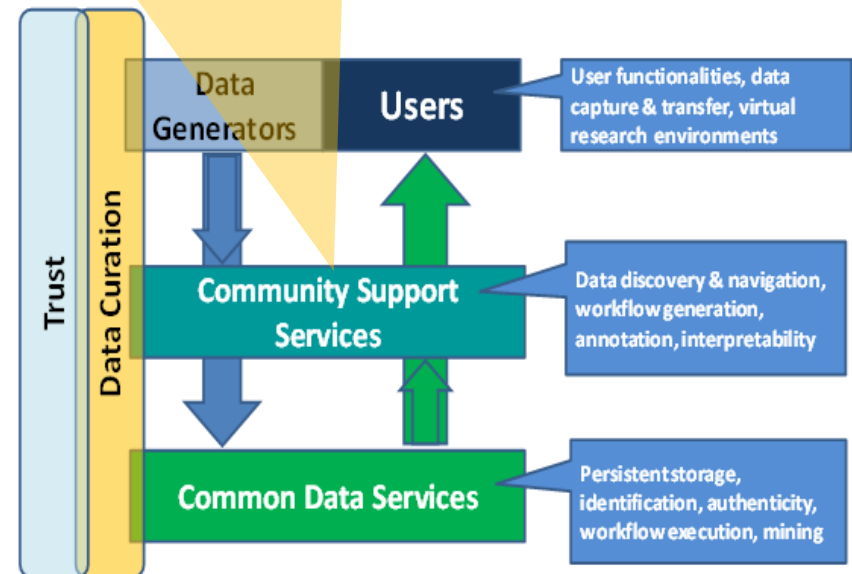
European  
Commission



# Implementing interoperable data infrastructures

- (a) **data generators**; research projects, big research infrastructure, installations or medium size laboratories, simulation centres, surveys or individual researchers
- (b) **discipline-specific data service providers**, providing data and workflows as a service
- (c) **providers of generic common data services** (computing centres, libraries)
- (d) **researchers as users**, using the data for science and engineering

community driven data infrastructure, including ESFRI, ESFRI clusters and others



Open. Share. Re-use.

*Science. Set Free.* ■

*Research results. Linked.*

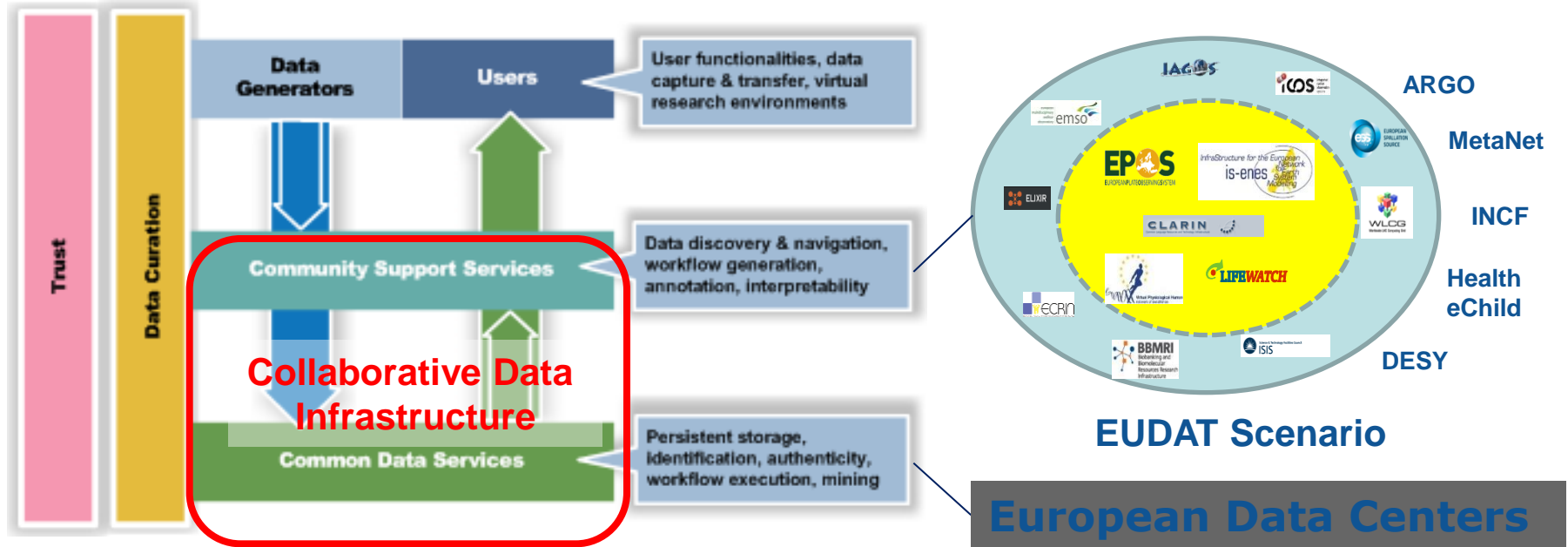


OpenAIRE

Open Access Infrastructure for Research in Europe



Data driven research across disciplinary and geographical boundaries  
Register relevant data objects stored in certified repositories  
Virtually integrate data objects in trusted federations  
Foster advancements in interoperability of object content  
Fragmentation and heterogeneity of data require standardization vs.  
innovation dynamics



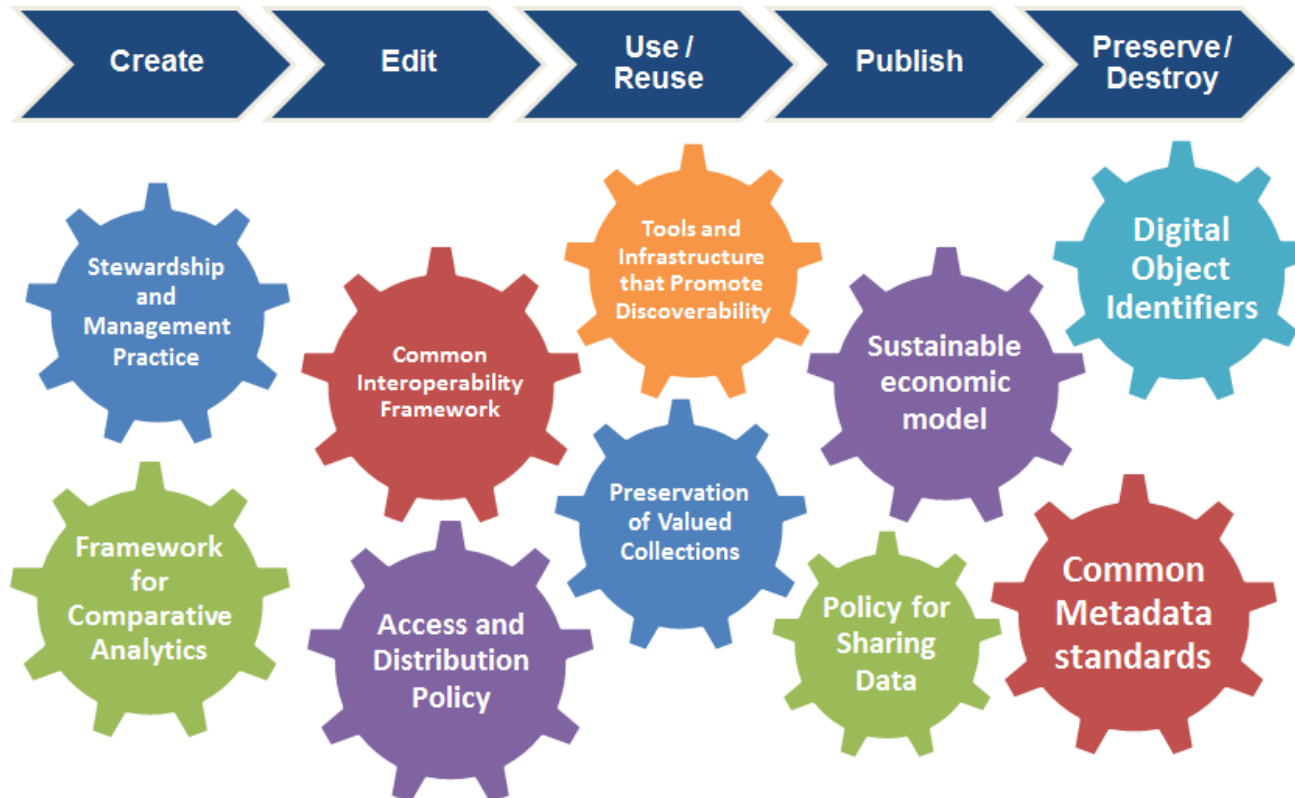
**launched 18 March**



## Research Data Alliance:

Common Infrastructure, Policy and Practice

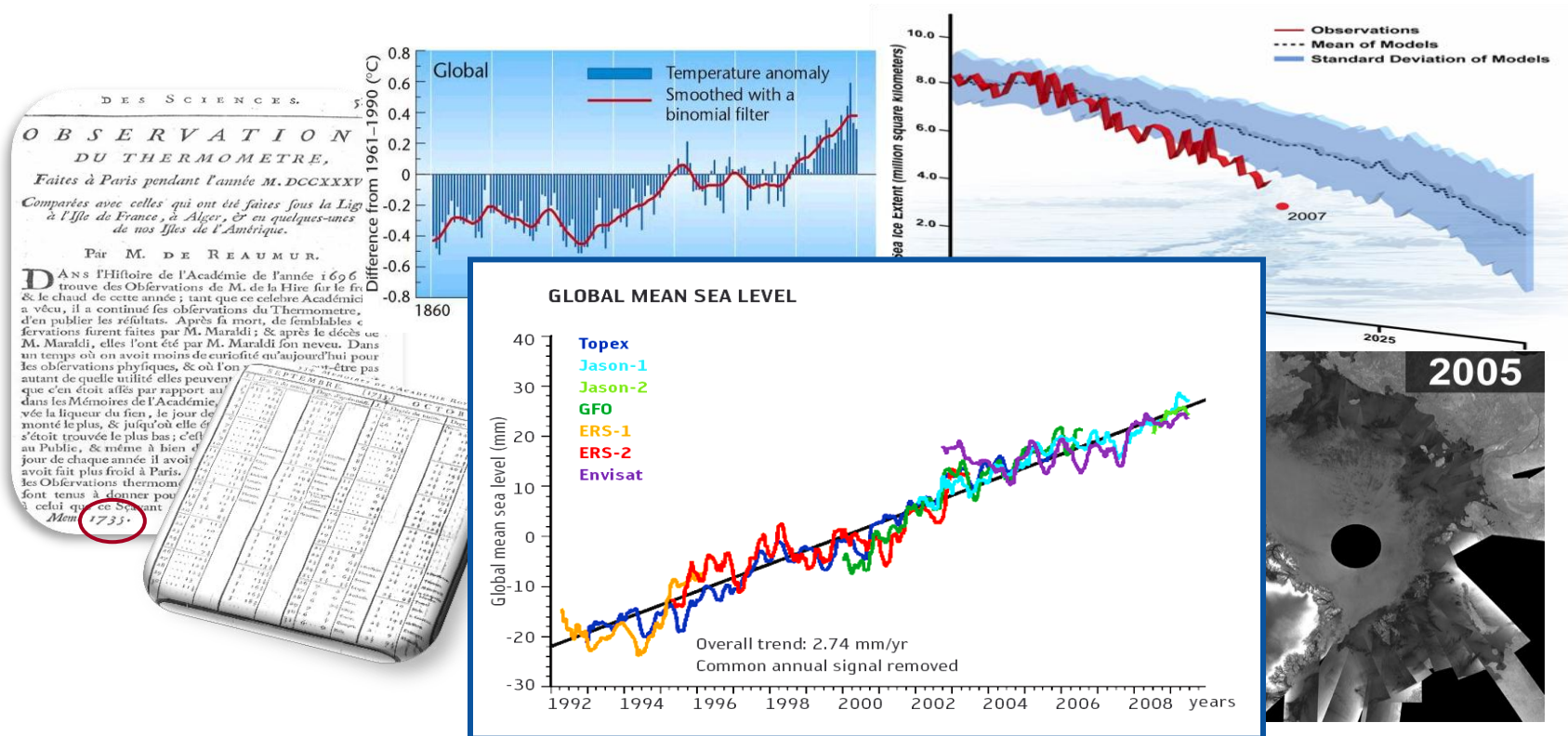
Drives Data Sharing and Exchange throughout the Data Life Cycle



*From Prof. Fran Berman and Prof. John Wood, Members of the RDA Council*

# community-driven data e-infrastructures

## SCIDIP-ES (Earth Observation Long Term Data Preservation)

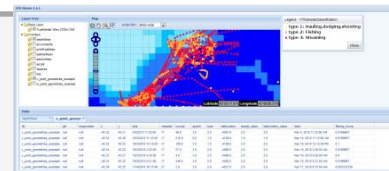


Adapted from a slide of Dr. Mirco Albani (ESA), project leader of SCIDIP-ES

# community-driven data e-infrastructures



- \* Vessel Activity Analysis
- \* Biological Niche Modeling
- \* Fishery Country Profile Product
- \* Global Catch Statistics Quality Improvement
- \* ...



## Simplified access and exploitation of tools for data analysis and processing

harmonization, aggregation, access to heterogeneous, multi-disciplinary and multi-format data

GEOSseas



EMODNET



GENESI-DR



GBIF



# community-driven data e-infrastructures



The **Virtual Observatory** is a community-led response to the challenges the astronomical community faces in data management and storage.



## Query VO resources for a given region of a sky

**Note:** DataScope V2.1 released March 26, 2007 (many cosmetic changes and some bug fixes)

What do we know about a given point or region in the sky?

To find out, just enter a target or position. The NVO DataScope will show you the results from hundreds of resources.

Position:

Use a target name (e.g., 3c273) or position (e.g., 10 10 10.1, 20 20 20.2)

Size:  (in degrees, max is 2)

Run query:

Skip cache? ☐ Refresh registry? ☐

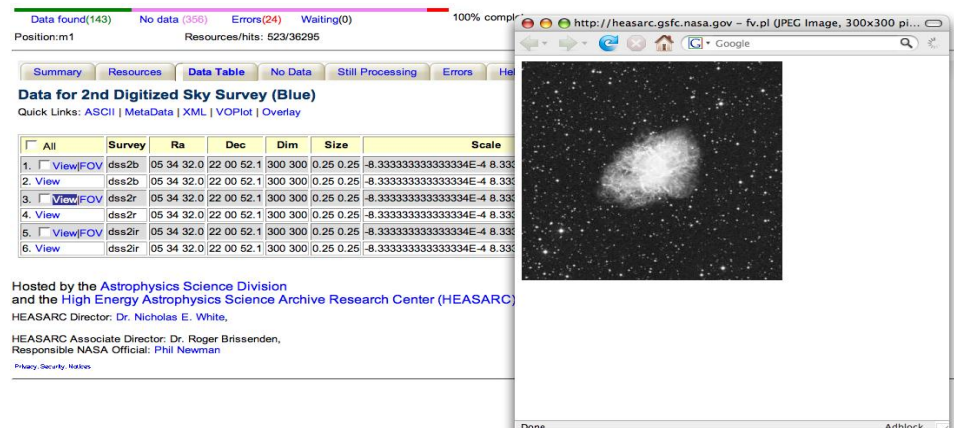
Do not add to list of recent queries? ☐

Some recent queries:

CGCG 456-050 (0.25)  
30 dor (0.25)  
186.66, -63.13 (0.0833)  
VCC 2062 (0.25)  
M87 (0.25)

**Positions** may be entered in decimal (dd.f, sdd.f) or sexagesimal (hh mm ss.f, dd mm ss.f) notation or as targets recognized by NED or SIMBAD.

The **Size** should be entered in decimal degrees.



Data found(143) No data (356) Errors(24) Waiting(0) 100% complete  
Position:m1 Resources/hits: 523/36295

Summary Resources Data Table No Data Still Processing Errors Help

**Data for 2nd Digitized Sky Survey (Blue)**  
Quick Links: ASCII | MetaData | XML | VOPlot | Overlay

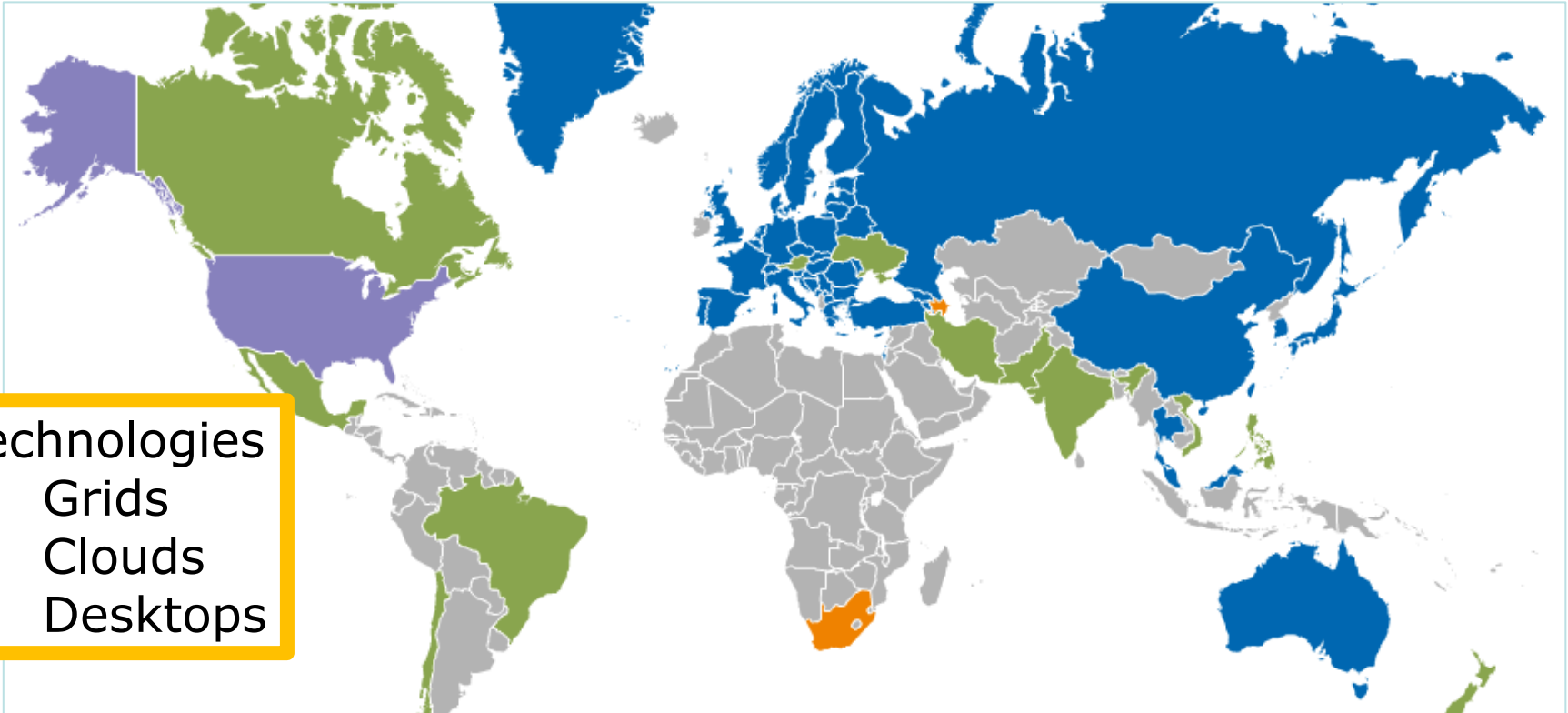
|    | All                     | Survey | Ra         | Dec        | Dim     | Size      | Scale                                    |
|----|-------------------------|--------|------------|------------|---------|-----------|--|
| 1. | <a href="#">ViewFOV</a> | dss2b  | 05 34 32.0 | 22 00 52.1 | 300 300 | 0.25 0.25 | -8.33333333333334E-4 8.33333333333334E-4 |
| 2. | <a href="#">View</a>    | dss2b  | 05 34 32.0 | 22 00 52.1 | 300 300 | 0.25 0.25 | -8.33333333333334E-4 8.33333333333334E-4 |
| 3. | <a href="#">ViewFOV</a> | dss2r  | 05 34 32.0 | 22 00 52.1 | 300 300 | 0.25 0.25 | -8.33333333333334E-4 8.33333333333334E-4 |
| 4. | <a href="#">View</a>    | dss2r  | 05 34 32.0 | 22 00 52.1 | 300 300 | 0.25 0.25 | -8.33333333333334E-4 8.33333333333334E-4 |
| 5. | <a href="#">ViewFOV</a> | dss2r  | 05 34 32.0 | 22 00 52.1 | 300 300 | 0.25 0.25 | -8.33333333333334E-4 8.33333333333334E-4 |
| 6. | <a href="#">View</a>    | dss2r  | 05 34 32.0 | 22 00 52.1 | 300 300 | 0.25 0.25 | -8.33333333333334E-4 8.33333333333334E-4 |

Hosted by the Astrophysics Science Division  
and the High Energy Astrophysics Science Archive Research Center (HEASARC)  
HEASARC Director: Dr. Nicholas E. White,  
HEASARC Associate Director: Dr. Roger Brissenden,  
Responsible NASA Official: Phil Newman  
Privacy Security Notice

Done AdBlock



# The European Grid Infrastructure today



## Technologies

- Grids
- Clouds
- Desktops

From 14 regional to 34 operations centres in 53 countries  
From 188,000 jobs/day with 80,000 cores on 250 Resource Centres  
to 1,200,000 jobs/day with 430,000 cores on 337 Resource Centres

# A European Cloud Partnership: big science teams up with big business



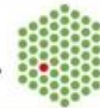
## Strategic Plan

- ▶ Establish multi-tenant, multi-provider cloud infrastructure
- ▶ Identify and adopt policies for trust, security and privacy
- ▶ Create governance structure
- ▶ Define funding schemes



To support the computing capacity needs for the ATLAS experiment

EMBL



Setting up a new service to simplify analysis of large genomes, for a deeper insight into evolution and biodiversity



To create an Earth Observation platform, focusing on earthquake and volcano research

Atos

Capgemini  
CONSULTING TECHNOLOGY OUTSOURCING

CloudSigma

CSA  
cloud security alliance

egi

interoute  
from the ground to the cloud

logica  
the brilliant together

OpenNebula.org  
The Open Source Toolkit for Cloud Computing

orange

Business  
Services

SAP

the SERVER LABS  
the IT architects

sixsq

Telefonica

terradue 20

THALES

Trustit

..T..Systems

# **Computing and data infrastructure in WP 2014-2015**

# Public Consultation on Research Data Infrastructures

*closed 27/3/2013*

ec.europa.eu/digital-agenda/en/content/consultation-research-data-infrastructures-framework-action



Google

[About](#) | [Contact](#) | [Legal notice](#) | [Search](#) | [Login](#)



## DIGITAL AGENDA FOR EUROPE

A Europe 2020 Initiative

[European Commission](#) > [Digital Agenda for Europe](#) > [Digital Life](#) > [Get involved](#) > [Consultations](#)

[Home](#)

[Our Goals](#)

[Digital Life](#)

[Business & Funding](#)

[Science and Technology](#)

[Telecoms & the Internet](#)

[Creativity & media](#)

[Around the World](#)

### Get involved

- › [About](#)
- › [Consultations](#)
  - › [Help us improve our analysis and measurement!](#)
  - › [Consultation on Research Data Infrastructures](#)
  - › [Consultation on directions for](#)

## Consultation on Research Data Infrastructures: Framework for Action

Share

Search the site



### Introduction

In preparation to H2020, the European Commission (DG CONNECT) has put together a [Framework for Action](#) addressing the area of **Research Data e-Infrastructures**.

The "Action Fiches" propose concrete domains of action. They aim at realising the European vision for a global interoperable data e-infrastructure supporting open, digital-driven, science (note that High Performance Computing and Distributed Computing as such are outside

## About the public consultation

### What

Relevance, Strengths, Weaknesses

Propose additional areas of actions

### How many

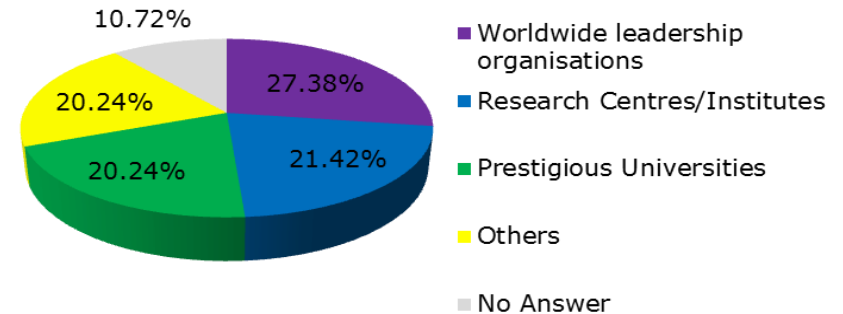
80+ replies from 100+ organisations

### Who

Research organisations and associations, academia,...

LERU, LIBER, CNRS, COAR, EIROforum,, OpenAIRE, CERN, APA, Volker Mehrmann  
TU Berlin, European Bioinformatics Institute, Max Planck Society, Observatoire  
Astronomique de Strasbourg, Museum f. Naturkunde Berlin, Pensoft Publishers,  
University of Edinburgh, University of Göttingen, University of Florence, etc.

**The number of different types of institution which the survey participants come from**







European  
Commission

## INTEGRATED e-INFRASTRUCTURE SERVICES

### VRE

E-INFRASTRUCTURES FOR  
VIRTUAL RESEARCH  
ENVIRONMENTS (VRE)

PROVISION OF  
CORE SERVICES  
ACROSS  
E-INFRASTRUCTURES

### DATA

### COMPUTING

NETWORK OF  
HPC COMPETENCE  
CENTRES FOR SMES

MANAGING, PRESERVING  
AND COMPUTING WITH  
BIG RESEARCH DATA

RESEARCH AND  
EDUCATION  
NETWORKING –  
GEANT

CENTRES  
OF EXCELLENCE  
FOR COMPUTING  
APPLICATIONS

E-INFRASTRUCTURES  
FOR OPEN ACCESS

PAN-EUROPEAN  
HIGH PERFORMANCE COMPUTING  
INFRASTRUCTURE AND SERVICES

TOWARDS GLOBAL DATA  
E-INFRASTRUCTURES  
RESEARCH DATA ALLIANCE

### CONNECTIVITY

### SUPPORT

E-INFRASTRUCTURE  
POLICY DEVELOPMENT AND  
INTERNATIONAL  
COOPERATION

NEW PROFESSIONS  
AND SKILLS  
FOR E-INFRASTRUCTURES



## **CALL 3**

### **1. Managing, preserving and computing with big research data**

MANAGING, PRESERVING  
AND COMPUTING WITH  
BIG RESEARCH DATA

Development and deployment of integrated, secure, permanent, on-demand service-driven and sustainable e-infrastructures for scientific computing and data

### **2. e-Infrastructure for Open Access**

TOWARDS GLOBAL DATA  
E-INFRASTRUCTURES  
RESEARCH DATA ALLIANCE

Robust e-infrastructure supporting Open Access policies in Europe, providing reliable and permanent access to digital scientific records

### **3. Towards global data e-infrastructures**

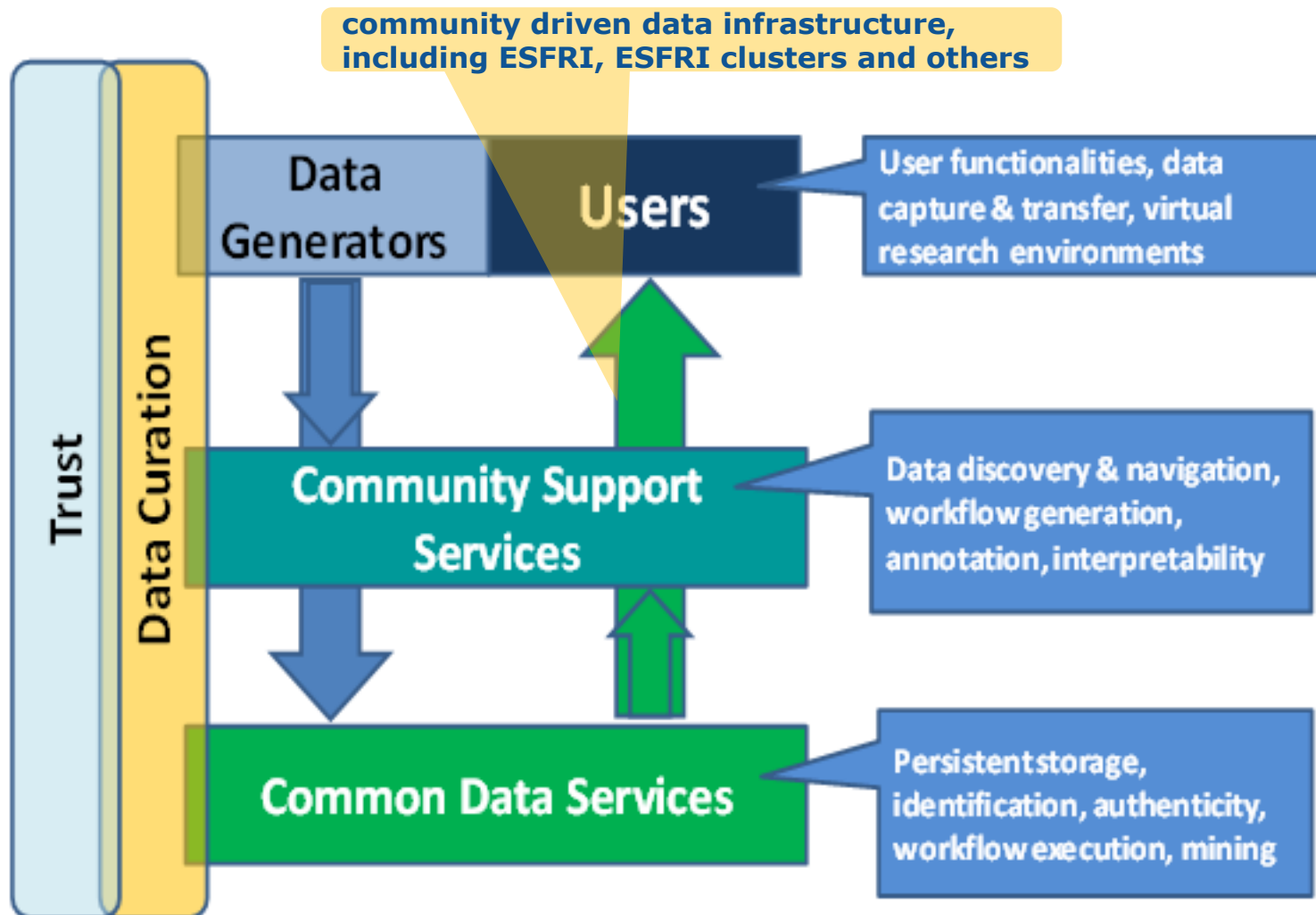
Consolidating Europe's contribution to the Research Data Alliance (RDA) and ensuring that RDA serves to foster research data interoperability at global level

E-INFRASTRUCTURES  
FOR OPEN ACCESS

### **7. Provision of core services across e-infrastructures**

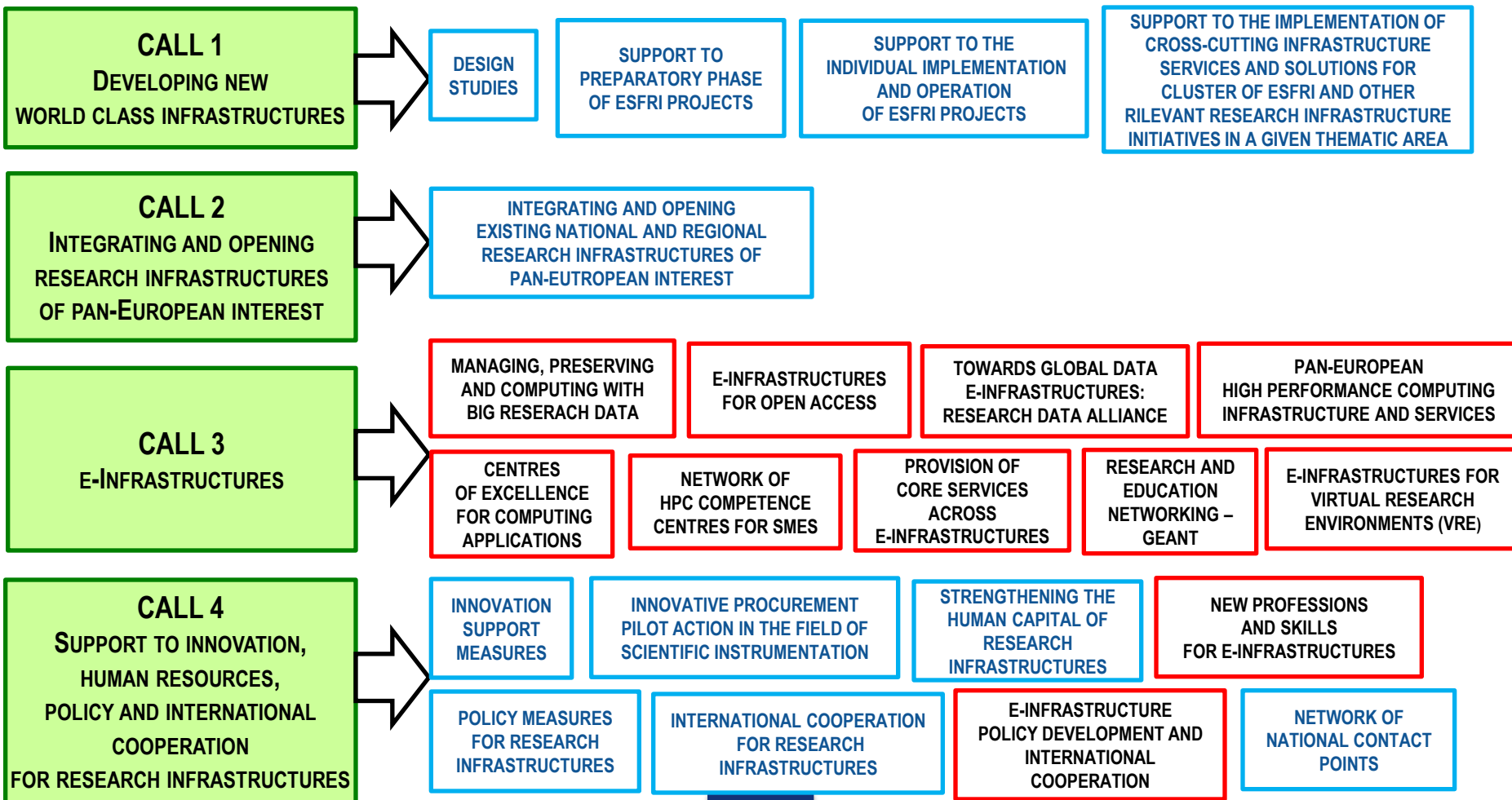
Harmonise and/or deploy core e-infrastructure services to ensure their effective use

# data infrastructure architecture





# RESEARCH INFRASTRUCTURE Work Programme 2014-2015



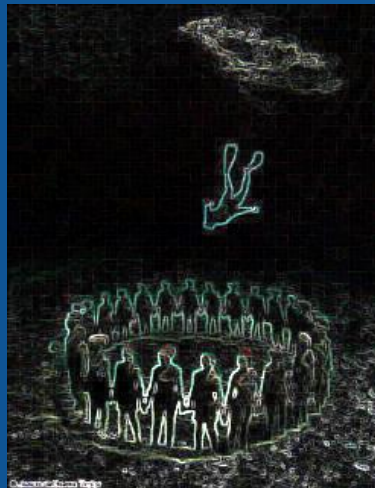


## Our challenges in 2014:

- *Bridging generic "service-driven" and "community-driven" data infrastructures*
  - Remove silos rather than make new ones!
- *Integration between data and computing*
  - How to use what we've got without extrapolating the past
- *Supporting policies; open data; open software*
- *Sharing basic operations services and building blocks*
- *Business plans for financial sustainability*
  - ...and partnerships with the private sector



# THANK YOU



European Commission – DG CONNECT  
eInfrastructures



## Call 3 Topic 1: Managing, preserving and computing with big research data (1/3)

### Challenges:

- **Provision of services** cutting across a wide-range of scientific communities and addressing diverse computational requirements
- ...based on **integrated, secure, permanent, on-demand service-driven, privacy-compliant and sustainable e-infrastructures** incorporating advanced computing resources and software
- **Capacity increase** to manage, store and analyse extremely large, heterogeneous and complex **research datasets**, including text mining of large corpora
- **Legal constraints and requirements**, system and service architectures, formats, types, vocabularies and legacy practices of scientific communities



## Call 3 Topic 1: Managing, preserving and computing with big research data (2/3)

### Scope:

- *Federated pan-European infrastructure for data manag't*
  - **Access, replication, annotation, search, compute, analysis, preservation**
- *Services for quality and reliability*
  - **including certification mechanisms and services**
- *Federating data management and curation tools*
  - **support development of Data Management Plans**
- *Large scale virtualisation of data/compute centre resources*
- *Standards-based open platform for scientific computing*
  - **deployed on different hardware and e-infrastructures**
  - **abstract application development and execution from available resources; for long-tail research communities**



## Call 3 Topic 1: Managing, preserving and computing with big research data (3/3)

### Scope (continued):

- *Support the evolution of the European Grid Infrastructure*
  - **Federation of ICT resources to offer compute/storage services (any kind of resource: public and private, grid and cloud, ...)**
- *Proof of concept & prototypes of data infrastructure-enabling software*
  - **extremely large or highly heterogeneous data sets**
  - **e.g. for databases or data mining**
- *Enabling aggregation of content for textual analysis*
  - **Platform for text mining, analytics, visualisation**
  - **Consulting and counselling services on the legal framework and permissions to text mine collections**





## Call 3 Topic 2: e-Infrastructure for Open Access

### Challenges:

- **Reliable and permanent access** to digital scientific records
- Based on already existing e-infrastructures (**institutional and thematic repositories, aggregators**, etc.)
- **Supporting Open Access policies**, also for Horizon 2020
- Some key elements:
  - **Link literature and data** in order to enable a more transparent evaluation of research and reproducibility of results
  - Provide service driven infrastructures to enable wide participation in the **Open Data Pilot**
  - Include scheme to provide financial support to **Gold Open Access post-grant**





# Call 3 Topic 2: e-Infrastructure for Open Access

## Scope:

### 1. Service-driven e-infrastructure for OA

- Linking institutional and thematic repositories across Europe
- Respond to requirements of researchers and organisations
- Helpdesks to support producers and users of information
- Technical solutions for sharing sensitive data (e.g. patient data)
- Bibliometrics and webometrics for open access

### 2. Prototyping new services in support of open science

- New forms of publishing, peer review, data mining services, ...
- Stimulation of publishing in open access journals
- Barriers to data sharing and pan-European data sharing agreements

### 3. Global interoperability of open access data e-infrastructures

- Linking platforms across the globe



## Call 3 Topic 3: Towards global data e-infrastructures – Research Data Alliance

### Challenges:

- Europe's contribution to the Research Data Alliance (RDA), ensuring that RDA fosters research data interoperability and exchange at global level
- Ensure Europe's role as a global player
- *RDA is an open international forum to create consensus on solutions and best practices to specific problems hampering data exchange and interoperability*



## Call 3 Topic 3: Towards global data e-infrastructures – Research Data Alliance

### Scope:

1. Active **participation of European stakeholders** (organisations and individual experts) in RDA; promotion of EU industry involvement and innovation.
2. Engagement of scientific communities in defining **best practices for data exchange**
3. RDA **governance** structures
4. Establishment of **coordination mechanisms** at European level and with international organisations dealing with **standardisation**, research data and education issues (IETF, W3C, CODATA, OECD, UNESCO, ...).





## Call 3 Topic 7: Provision of core services across e-Infrastructures

PROVISION OF  
CORE SERVICES  
ACROSS  
E-INFRASTRUCTURES

### Challenge:

Harmonise and deploy core infrastructure services

### Scope:

- Development and uptake of a **Digital Identifier** e-infrastructure **for digital objects** (articles, datasets, collections, software, nomenclature, etc.) **and authors**;
- Deployment and promotion of a pan-European **identity federation** for researchers, educators and students
  - **Implementing an ERA Communication action**

*NB! Core services: those needed across a broad range of communities and e-infrastructures and ensure their interoperation*



## Call 3 Topic 7: Provision of core services across e-Infrastructures

PROVISION OF  
CORE SERVICES  
ACROSS  
E-INFRASTRUCTURES

### Expected Impact:

- **Improved interoperability** of e-infrastructure services, hence easier access and collaboration between scientists
- **Reduce duplication of efforts** for common services
- DIs open **new prospects for services** and for encouraging **openness and trust**
- Europe-wide **single sign-on service** enables researchers to collaborate within secure and trusted virtual research environments
- AAI facilitates **sharing** of information resources at EU level
- Expansion of the **coverage of national identity federations** for network, services and applications. Easier participation in identity federations





## Call 3 Topics 1,2,3 and 7: Practicalities

### When (indicative)

Call publication date 11/12/2013

Call deadline for Topic 2 15/04/2014 at 17:00

Call deadline for Topics 1, 3 and 7 02/09/2014 at 17:00

### Budget (indicative)

Topic 1 Managing, preserving and computing with big research data: 55M€

Topic 2 e-Infrastructure for Open Access: 13M€

Topic 3 Towards global data e-infrastructures – RDA: 4 M€

Topic 7 Provision of core services across e-Infrastructures: 6 M€