

# EGI: advanced computing for research

**Yannick Legré**

Managing Director

[yannick.legre@egi.eu](mailto:yannick.legre@egi.eu)

[www.egi.eu](http://www.egi.eu)



The EGI-Engage project is co-funded by the European Union (EU) Horizon 2020 program under grant number 654142





# EGI: advanced computing for research

EGI is a federation of over 300 computing and data centres spread across 56 countries in Europe and worldwide

EGI delivers services to support scientists, multinational projects and research infrastructures

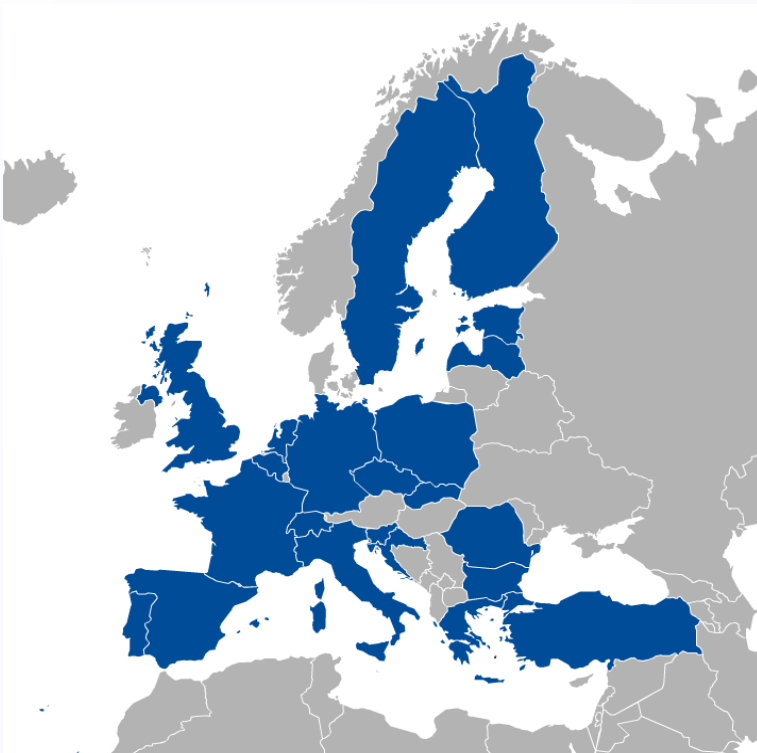


EGI's mission is to create and deliver  
open solutions for science and research  
by federating digital capabilities,  
resources and expertise  
between communities and  
across national boundaries.



# EGI Foundation

- 22 national e-Infrastructure providers
- Two EIROs: **CERN** and **EMBL-EBI**

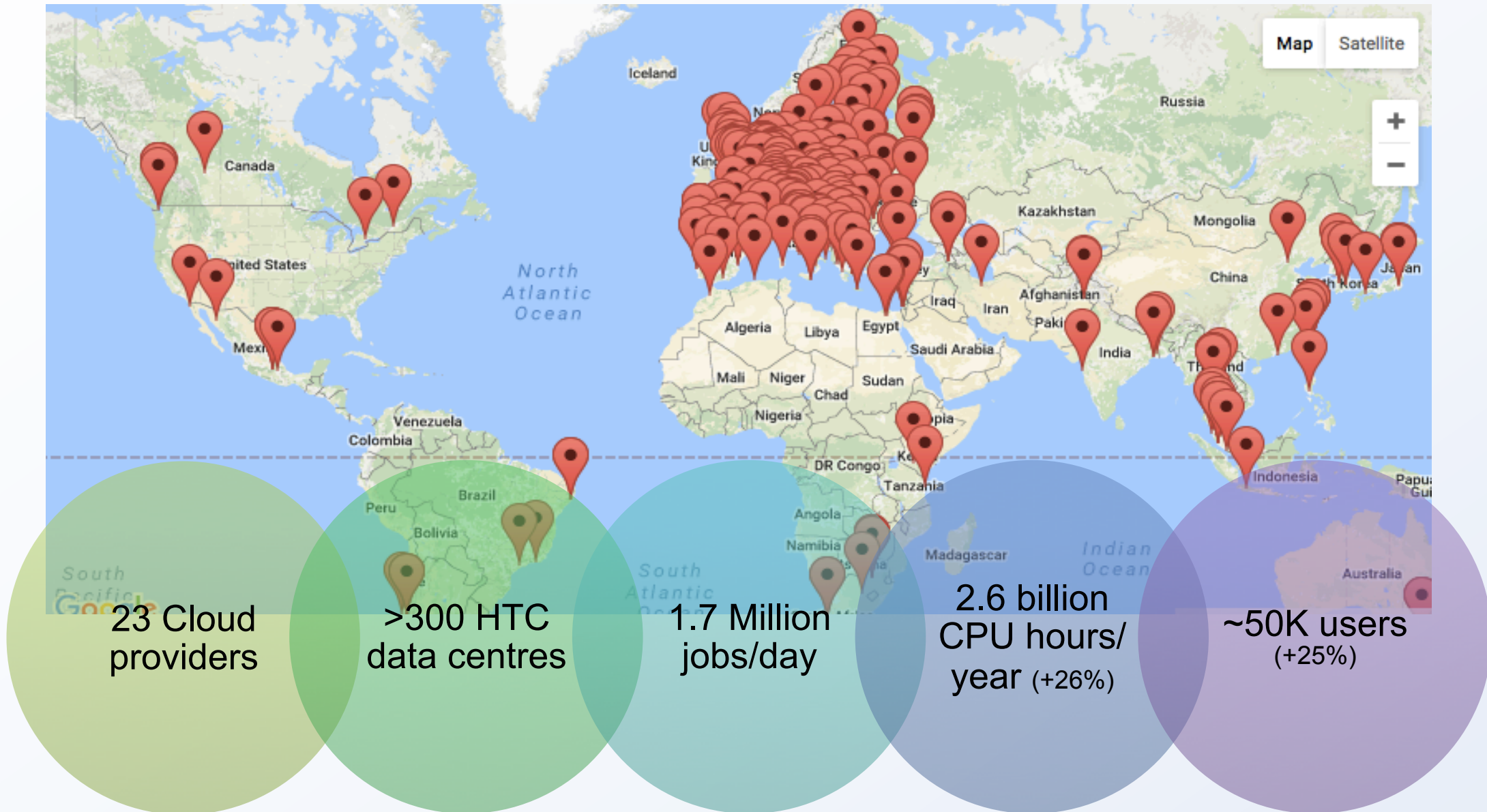


[www.egi.eu/about/egi-foundation/](http://www.egi.eu/about/egi-foundation/)



# EGI Federation, 2016 QR4

The largest distributed compute e-Infra worldwide

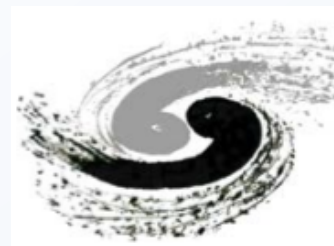


# A global system of e-Infrastructures

Africa-Arabia, Asia and Pacific region, China, Europe, India  
North and South America, Ukraine



Europe



China:  
Institute of HEP,  
Chinese Academy of Sciences



Africa and Arabia:  
Council for Scientific and  
Industrial Research, South Africa



Latin America:  
Universida de Federal do  
Rio de Janeiro



India:  
Centre for Development of  
Advanced Computing

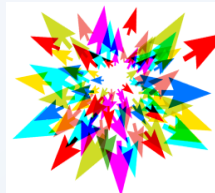


Ukraine:  
Ukrainian National Grid



Open Science Grid

USA

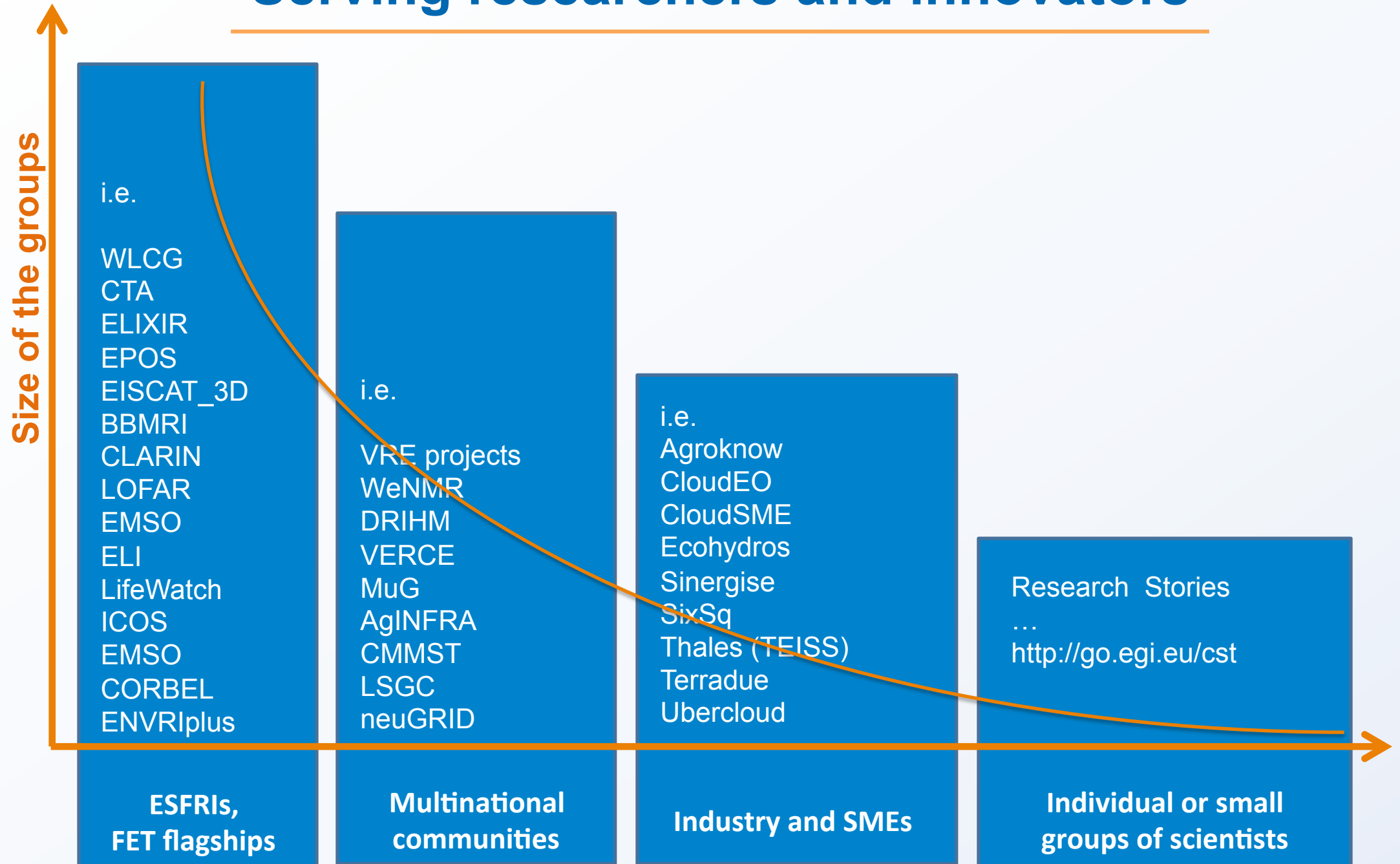


compute  
canada

calcul  
canada

Canada

# Serving researchers and innovators



# Projects related to Big Data

---



- Accelerating the implementation of the **Open Science Commons** expanding the capabilities of a European backbone of **federated services for compute, storage, data, communication, knowledge and expertise**, complementing community-specific capabilities.



- Developing a **data/computing platform targeted at scientific communities**, deployable on multiple hardware, and provisioned over hybrid (private or public) e-Infrastructures



- Defining the **competences and skills framework**, body of knowledge and model curriculum for the new profession of the **data scientist**.



- Creating a more coherent, interdisciplinary and interoperable **cluster of Environmental Research Infrastructures** across Europe.



- Building a **hybrid cloud platform** on top of which a competitive marketplace of European cloud players can develop their own services for a wider range of users.



- Implementing a **federated access for Earth Observation data**, including user friendly tools for data mining, discovery, access and exploitation



**A wide offer of  
services for  
Big Data**

# EGI Live Services

## Compute



**Cloud Compute** : Run virtual machines on demand with complete control over resources



**Cloud Container Compute** : Run Docker containers in a lightweight virtualised environment



**High-Throughput Compute** : Execute thousands of tasks to analyse large datasets

## Storage and Data



**Online Storage** : Store, share and access your files and their metadata on a global scale



**Archive Storage** : Back-ups for long term and future use in a secure environment



**Data Transfer** : Transfer large sets of data from one place to another

## Training



**FitSM training** : Learn how to manage IT services with a pragmatic and lightweight standard



**Training infrastructure** : Dedicated computing and storage for training and education

# EGI Services in the pipeline

## Operations

**Service Monitoring:** Monitor a wide range of services across your infrastructures

**Configuration Database:** Manage the configuration information of federated e--infrastructure assets and their functional relations

## Data Management

**Data Hub:** Selected big datasets of public interest made exploitable – <https://datahub.egi.eu>

**Federated Data Manager:** Share, discover, and process data federated from different sources

**Content Distribution:** Deliver data in the most efficient way

## Security

**Attribute Management:** Manage community membership and expose trusted information

**Check-In:** Handle transparent Single Sign-On from multiple heterogeneous identity providers



# EGI Services for Participants

## Operations

**Accounting** Track and report the usage of your services

**Helpdesk** Handle EGI service requests and incidents for distributed support teams

**Operations Tools** Integrate resources and operations

**Collaboration and Community Management Tools**

Coordinate activities to ensure seamless operations

**Configuration Database** Manage the configuration information of federated e--infrastructure

**Service Monitoring** Monitor EGI services and provide operational and business insights

**Validated Software and Repository** Manage high-quality software releases

## Security

**Attribute Management:** Manage community membership and expose trusted information

## Coordination

**Communications** Share your successes at a larger scale

**ITSM Coordination** Ensures professional service management for EGI IT services

**Community Coordination** A joint approach to user engagement

**Operations Coordination and Support** Coordinate activities to ensure seamless operations

**Security Coordination** Enhance local security for a safer global infrastructure

**Project Management and Planning** A joint approach to planning and management

**Strategy and Policy Development** Define common strategies and policies in Europe and worldwide

**Technical Coordination** Progress and innovation through collaboration

# Services offered to SMEs and Industry

---

✓ Pilots

✓ Proof of concepts

**Access**

**Reuse**

✓ Datasets

✓ Tools/Software

✓ Applications

✓ EU visibility

✓ Networking

**Market**

**Co-design**

✓ Products/Services

✓ Integrate

✓ Share expertise

# Big Data

# Use Cases



# WLCG: Online Storage and Data Transfer at unprecedented scales

---



ATLAS control room at CERN

In 2016, the Worldwide LHC Computing Grid (WLCG) transferred on average **80 Petabytes of data per month.**

This corresponds to more than **1 billion files per month** transferred to **thousands of particle physicists** working across the world.

The WLCG is the largest resource provider and service consumer of the EGI Federation

# Cherenkov Telescope Array: the world's leading gamma-ray observatory

---



Credit: Akihiro Ikeshita Mero-TSK

1350 scientists from 32 countries

## **CTA EGI usage (2013-2016):**

- 360 million HS06 CPU hours
- 11 PB of data transferred
- 2 PB currently in storage
- 11 million compute jobs

CTA used EGI's High-Throughput Compute and Online Storage services to manage its computational challenges during the array preparatory phase. <http://go.egi.eu/cta>

# Supporting the Geohazard Exploitation Platform

---



## Very fast access to large volumes of data

- Earth Observation data
- In-situ data

## Distributed processing

- Hybrid cloud infrastructure
- Access to HTC resources



# Making sense of ocean salinity & temperature data

---



<http://www.euro-argo.eu/>

## Euro-Argo

- Operates an array of ca. 800 floating sensors
- Enhances coverage in the Europe's regional seas
- Provides access to quality controlled data and data products

Hadoop clusters for data filtering are being deployed and evaluated in EGI Federated Cloud



# **Contribution to Big Data Value in Europe**

# Publicly-funded e-Infrastructures landscape

---



**Open research  
data**

**Data- and  
compute-  
intensive science**

**Research and  
education  
networking**

**High Performance  
Computing**

**Big data  
innovation**



# The European Cloud Initiative

---

- **European Open Science Cloud (EOSC)**
  - Integration and consolidation of e-infrastructures
  - Federation of existing research infrastructures and scientific clouds
  - Development of cloud-based services for Open Science
  - Connection of ESFRIs to the EOSC
- **European Data Infrastructure (EDI)**
  - Development and deployment of large-scale European HPC, data and network infrastructure
- **Widening access**
  - SMEs, Industry, Government

**E-INFRASTRUCTURES:**  
**MAKING EUROPE THE BEST PLACE**  
**FOR RESEARCH AND INNOVATION.**



<https://ec.europa.eu/futurium/en/content/e-infrastructures-making-europe-best-place-research-and-innovation>

# A new push to stimulate R&D+I

---

**Provide  
added  
value  
services**

**Consume  
services &  
resources**

**Widening  
access**

- SMEs,  
Industry,  
Government

**Enhance  
resource  
capacity**

- For innovation  
and  
production of  
new  
knowledge

<https://www.egi.eu/business/>



# We are open for collaborations!

---

## EGI is

An open  
ecosystem for  
data-/compute-  
intensive  
applications

Bridging  
research and  
innovation

An enabler for  
stimulating the  
data value  
chain

Follow us:

@EGI\_elnfra

@EGIEInfrastructure



EGI Foundation • Science Park 140 • 1098 XG Amsterdam  
The Netherlands

+31 (0)20 89 32 007 • [www.egi.eu](http://www.egi.eu)

# Discover more

---

Visit [www.egi.eu/](http://www.egi.eu/)

Services [www.egi.eu/services/](http://www.egi.eu/services/)

Publications [www.egi.eu/about/publications/](http://www.egi.eu/about/publications/)

Contact us:

- For public institutions: [contact@egi.eu](mailto:contact@egi.eu)
- For private organisations: [business@egi.eu](mailto:business@egi.eu)