Standard-based Interoperability amongst Local, Grid and Cloud Resources to enable an Italian Distributed Computing Infrastructure

Roberto Barbera (roberto.barbera@ct.infn.it)
Univ. of Catania and INFN

GARR Workshop – Rome, 29 November 2012

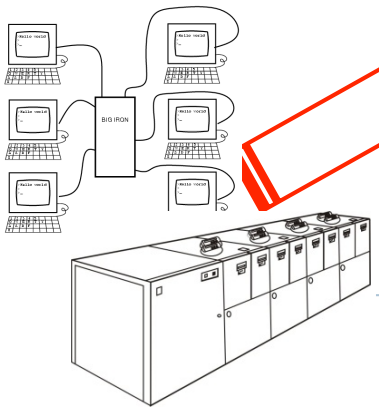
Outline

- Introductory considerations
- The Catania Science Gateway framework
- The CHAIN worldwide interoperability demo
- Summary and conclusions

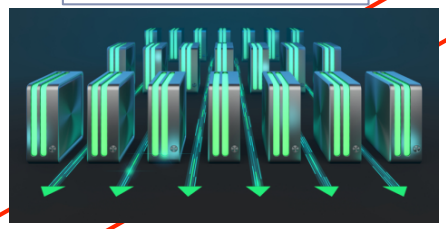
Evolution of distributed computing

Cost of hw
Cost of networks

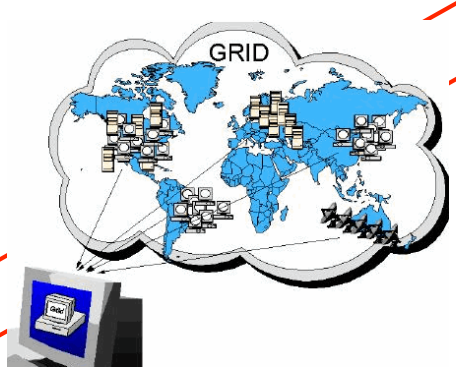
Mainframe Computing



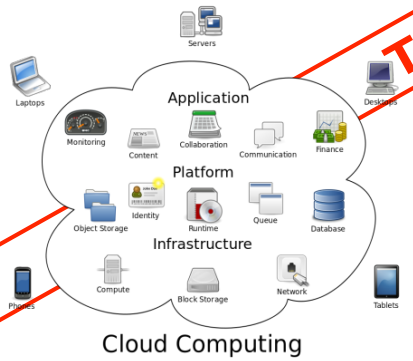
80's-90's Cluster Computing



90's-00's Grid Computing



00's-10's Cloud Computing

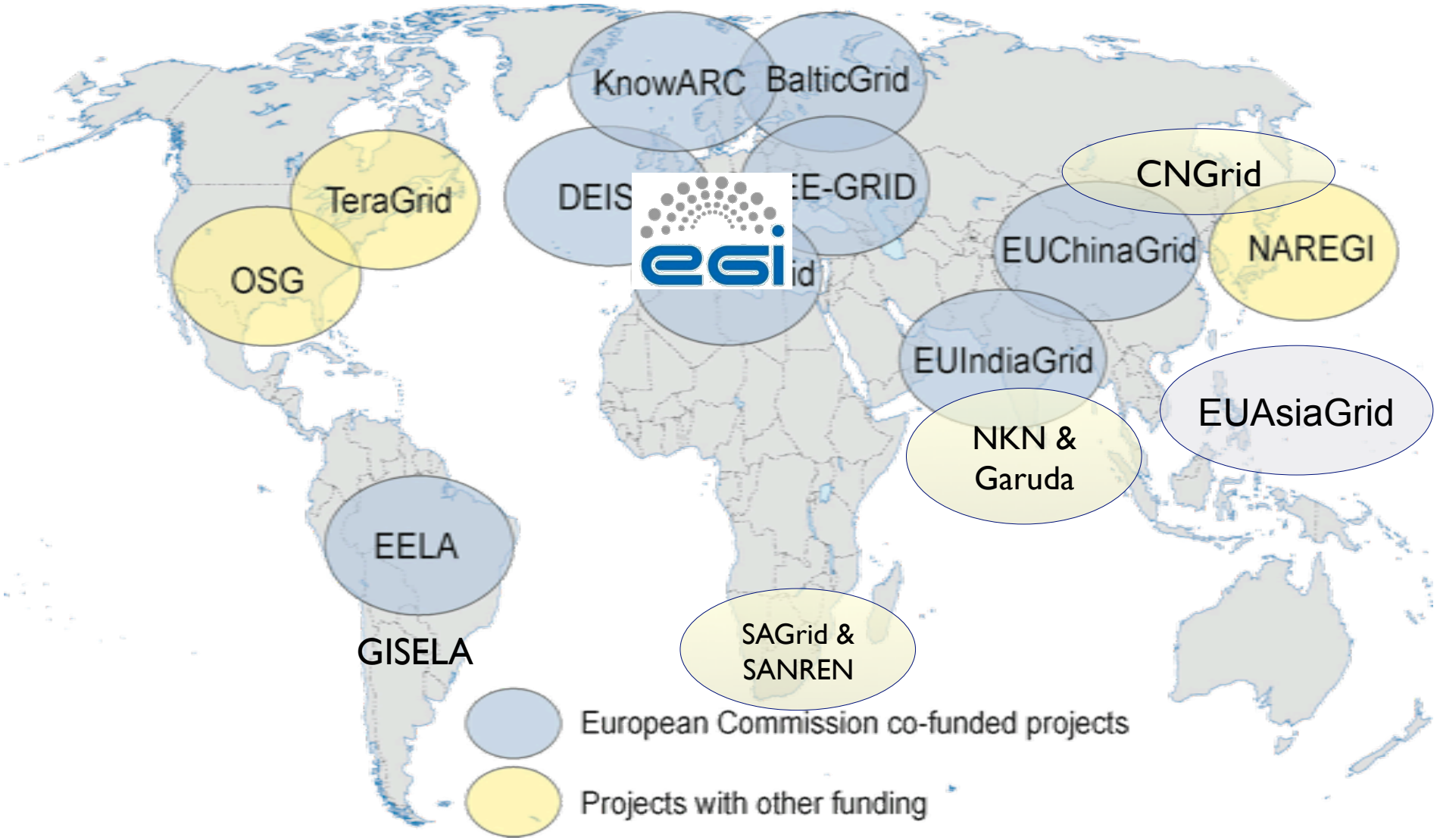


Time

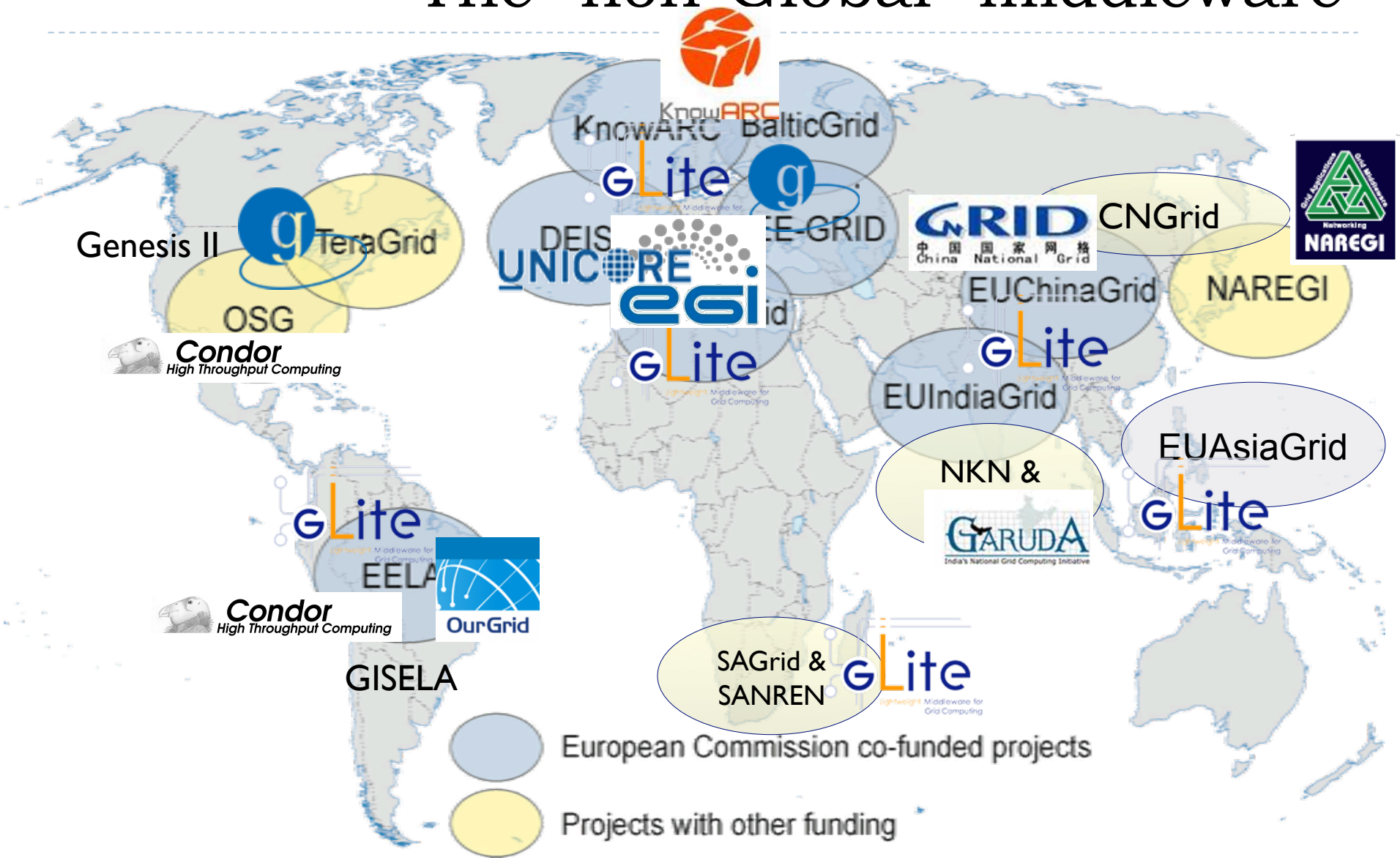
Power of COTS
WAN bandwidth

One architecture… many local resource managers

The “Global” Grid



The “non-Global” middleware



Using Grids is not straightforward ☹

JDL

```
Type = "Job";
JobType = "MPICH";
MPIType = "MVAPICH2_PGI706";
CpuNumber = 16;
MPIGranularity = 4;
Executable = "flash2";
StdOutput = "mpi.out";
StdError = "mpi.err";
InputSandbox =
```

Script

```
echo Staging Input Data \(Courtesy of European Space Agency\);
#edg-rm --vo=gilda copyFile lfn:$1.N1 file://$PWD/$1.N1;
lcg-cp --vo=gilda lfn:$1.N1 file://$PWD/$1.N1;
echo Staging Application;
gunzip beam20.tar.gz;
tar xvf beam20.tar;
cd beam-2.0/bin;
echo Starting Application;
./pds2jpg-ASAR-run.sh $1;
mv $1-b*.jpg ../..
cd ../..
rm -fr beam-2.0;
```

CLI

```
Exit code:      0
Status Reason:    Job terminated successfully
Destination:      grid010.ct.infn.it:2119/jobmanager-lcgpbs-gilda
Submitted:        Tue Jun 29 15:34:40 2010 CEST
```

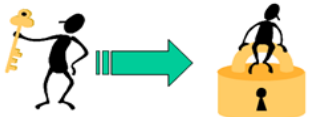
Users have to cope with complex security procedures, execution scripts, job description languages, command line based interfaces and lack of standards.
This makes the learning curve very steep and keeps non IT-experts away.

GSI

Request for the certificate

User's identity is confirmed by the RA

Certificate is emitted by the CA

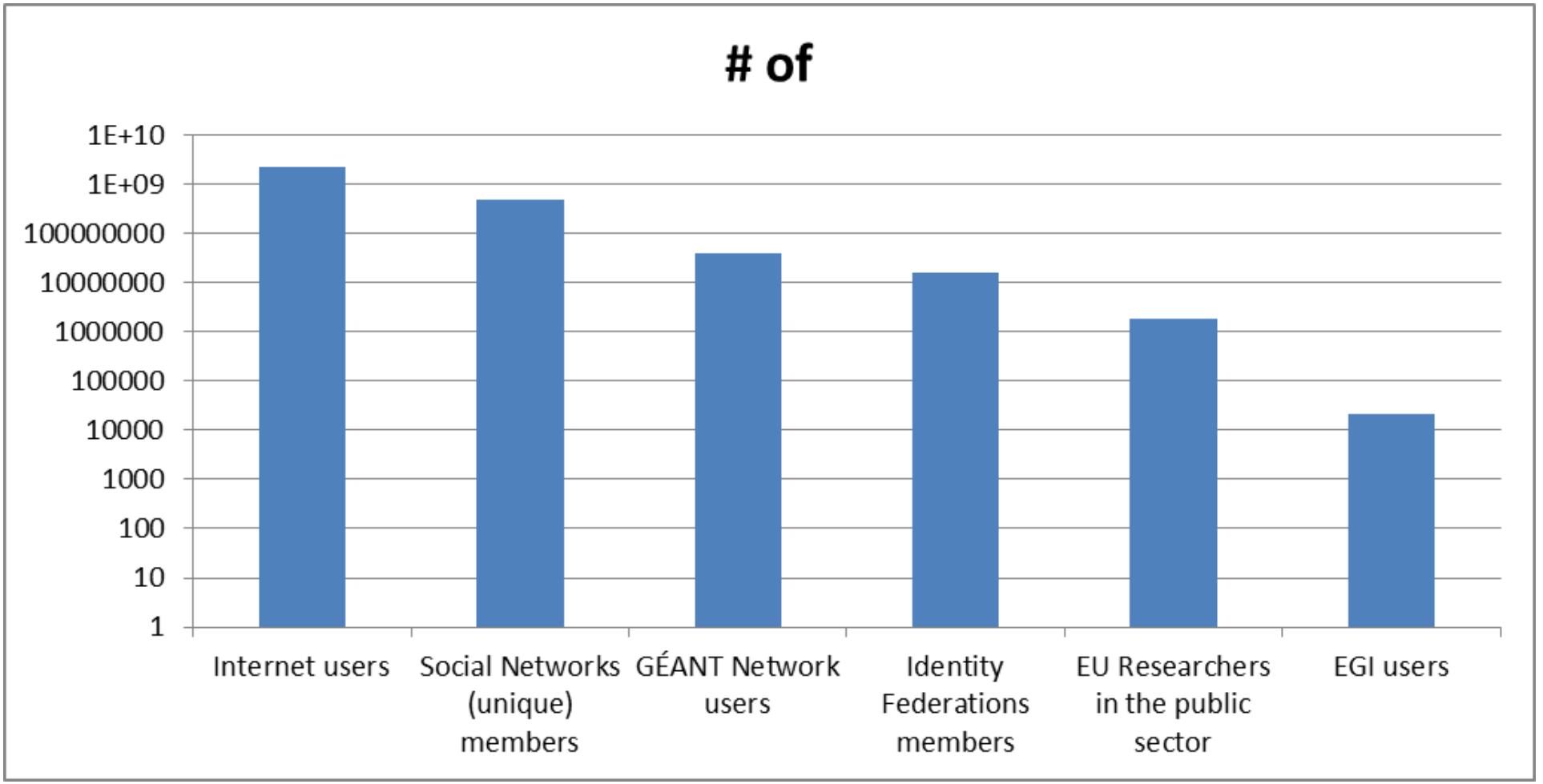


Certificate is used as an access key to the grid

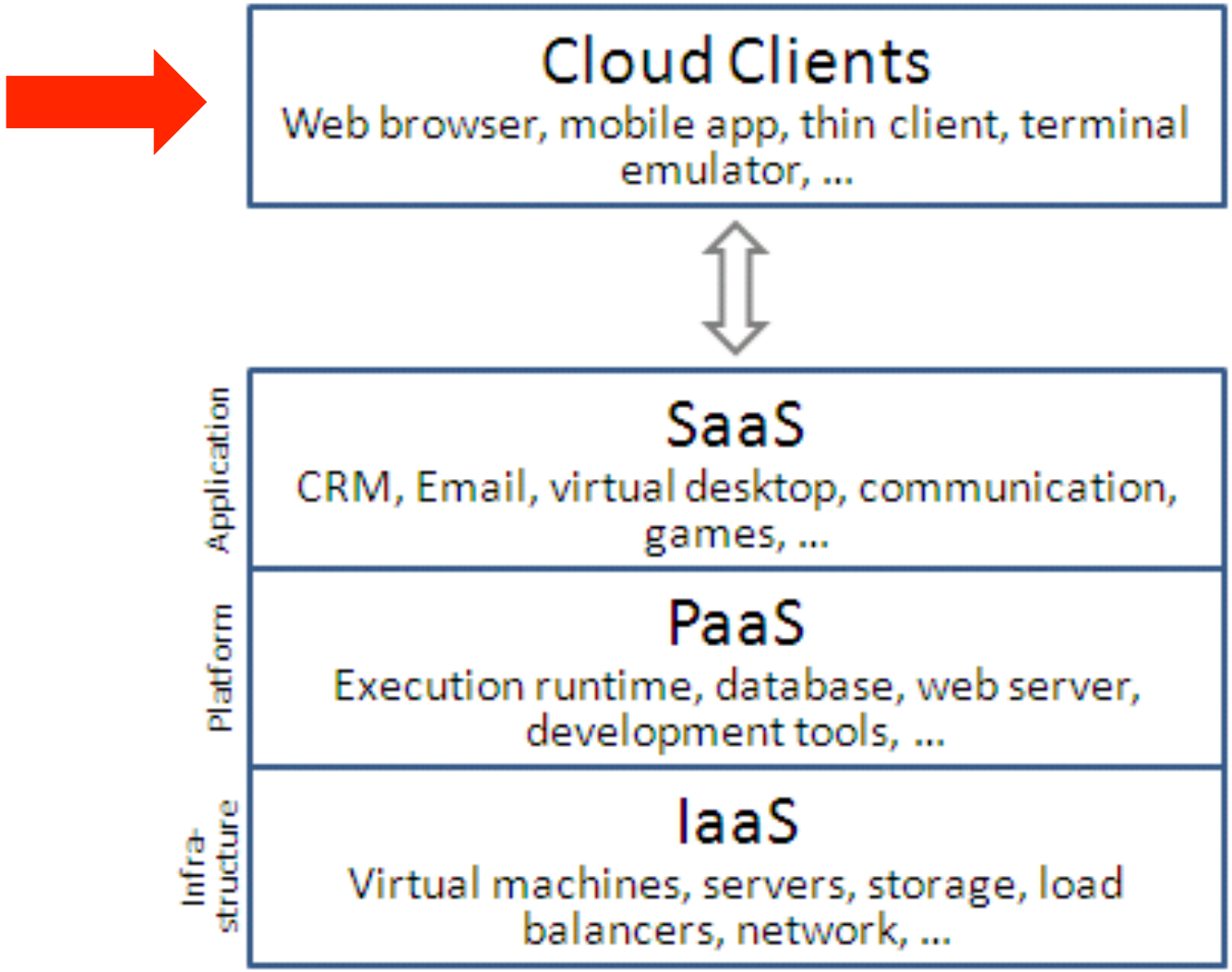
The eResearch2020 report
(http://www.eresearch2020.eu/eResearch%20Brochure%20EN.pdf)

- Some barriers in the adoption of Grids:
  - **Changes on Grids means changes on applications**
  - Time required to adapt usual workflows
  - Lack of structure to support anonymous access
  - Download and installation of applications
  - **Interface**
  - Slow to get to compared to other resources
  - **Difficult to use in the beginning**
  - **Time spent to get the application compiled and running**

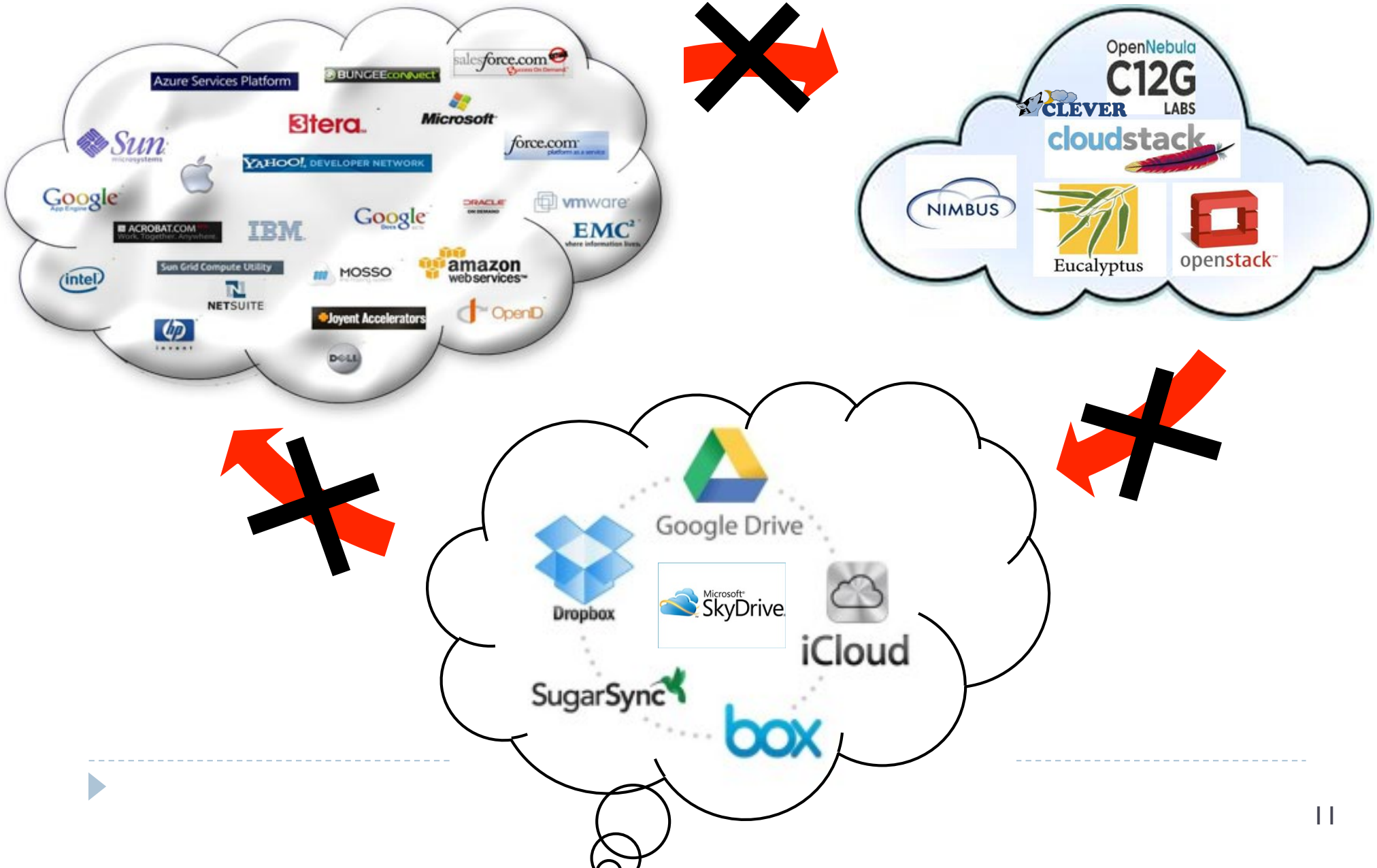
Some figures…



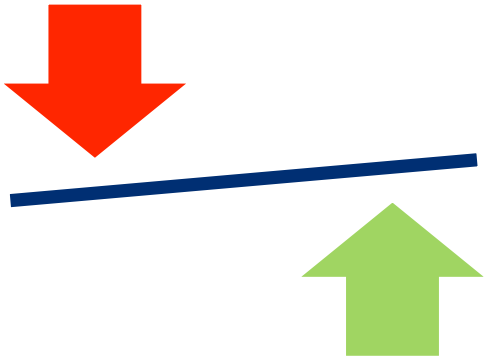
The cloud stack



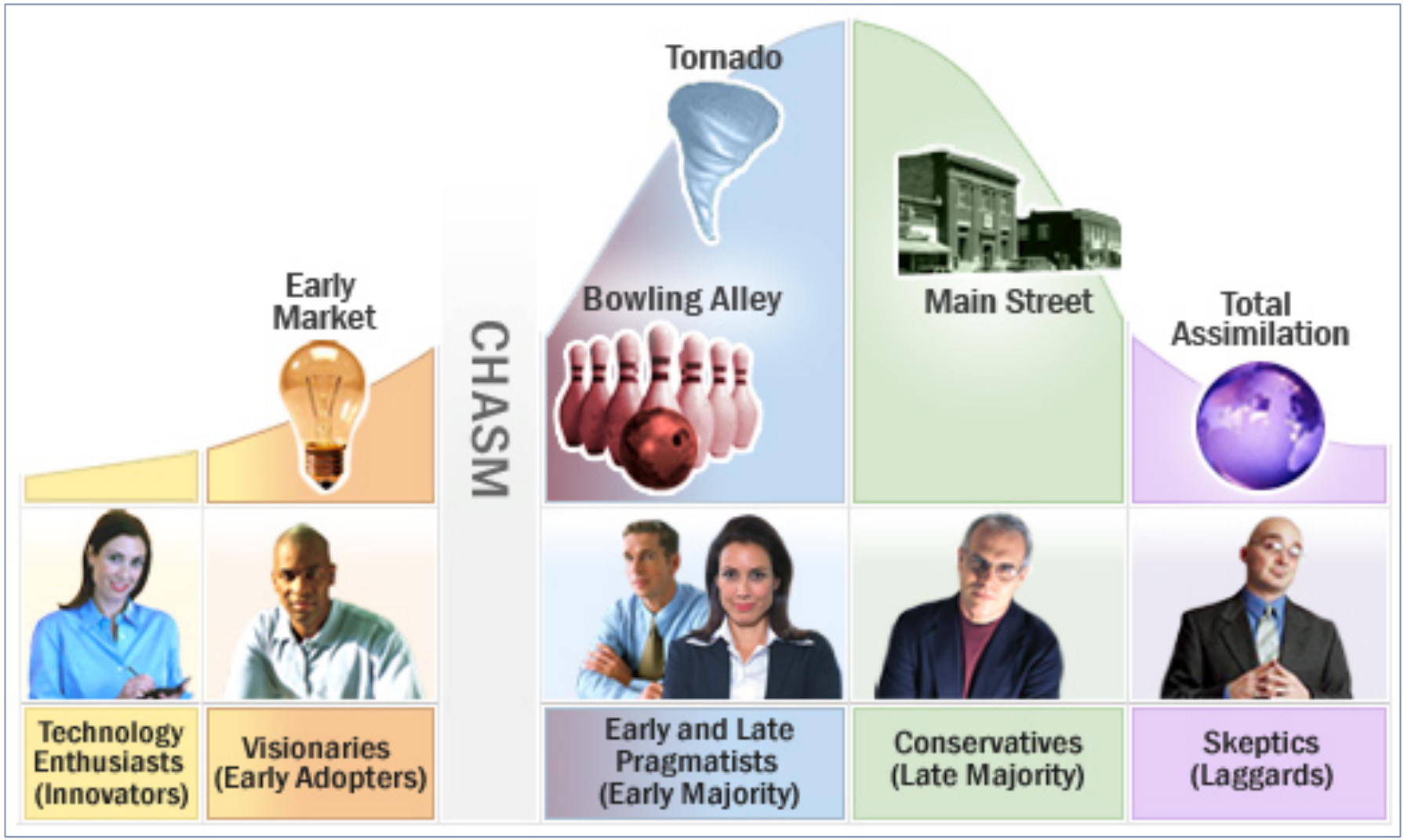
A very «cloudy» sky…



… and difficult choices to make



The path to technology uptake – Where are we with e-Infrastructures ?



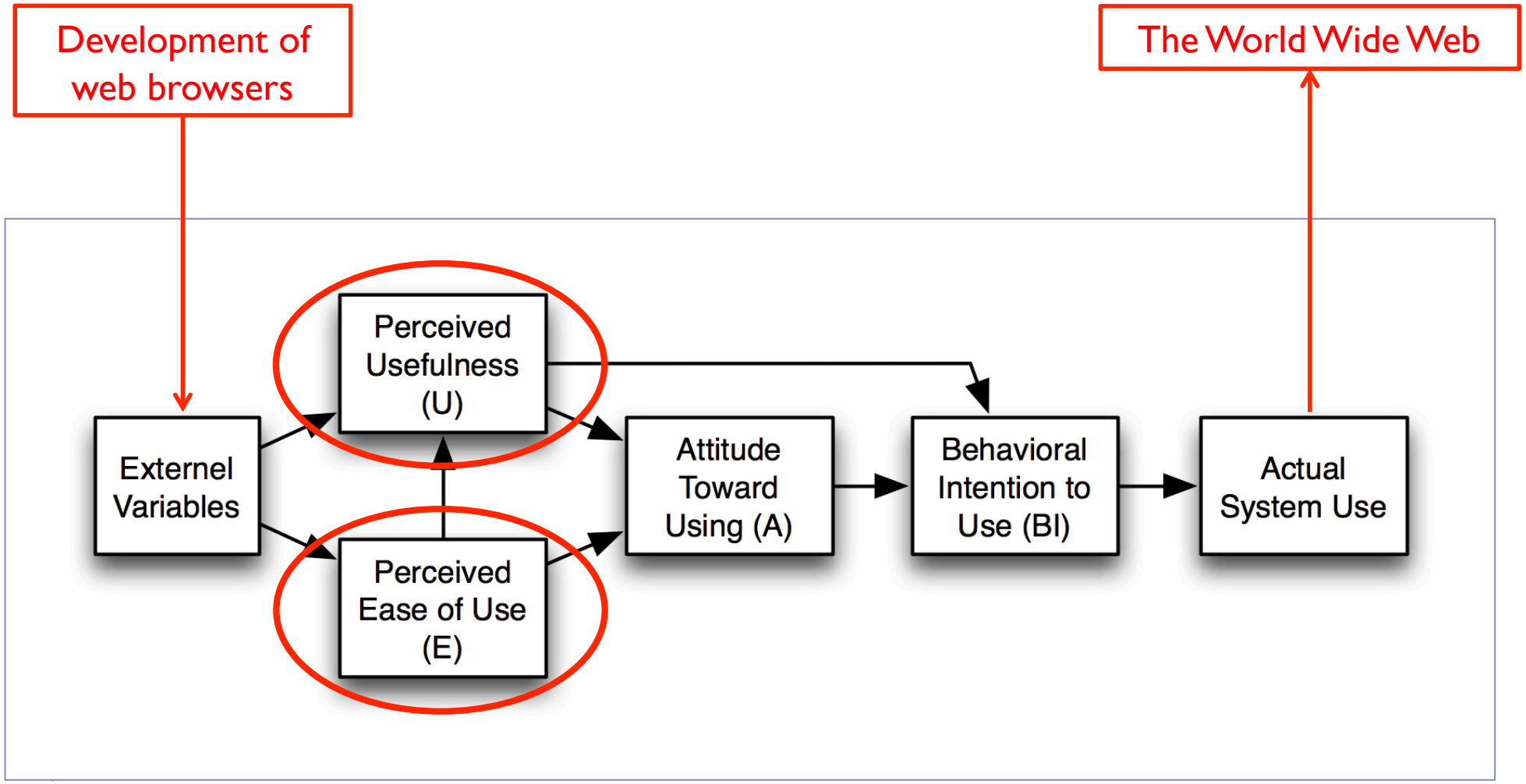
The Rogers "bell-shape" curve - Rogers, E. M. (1962), "Diffusion of Innovations", Glencoe: Free Press.

Interoperability – Does this definition apply to e-Infrastructures ?

- **Interoperability** is a property referring to *the ability of diverse systems and organizations to work together* (inter-operate). The term is often used in a technical systems engineering sense, or alternatively in a broad sense, taking into account social, political, and organizational factors that impact system to system performance;

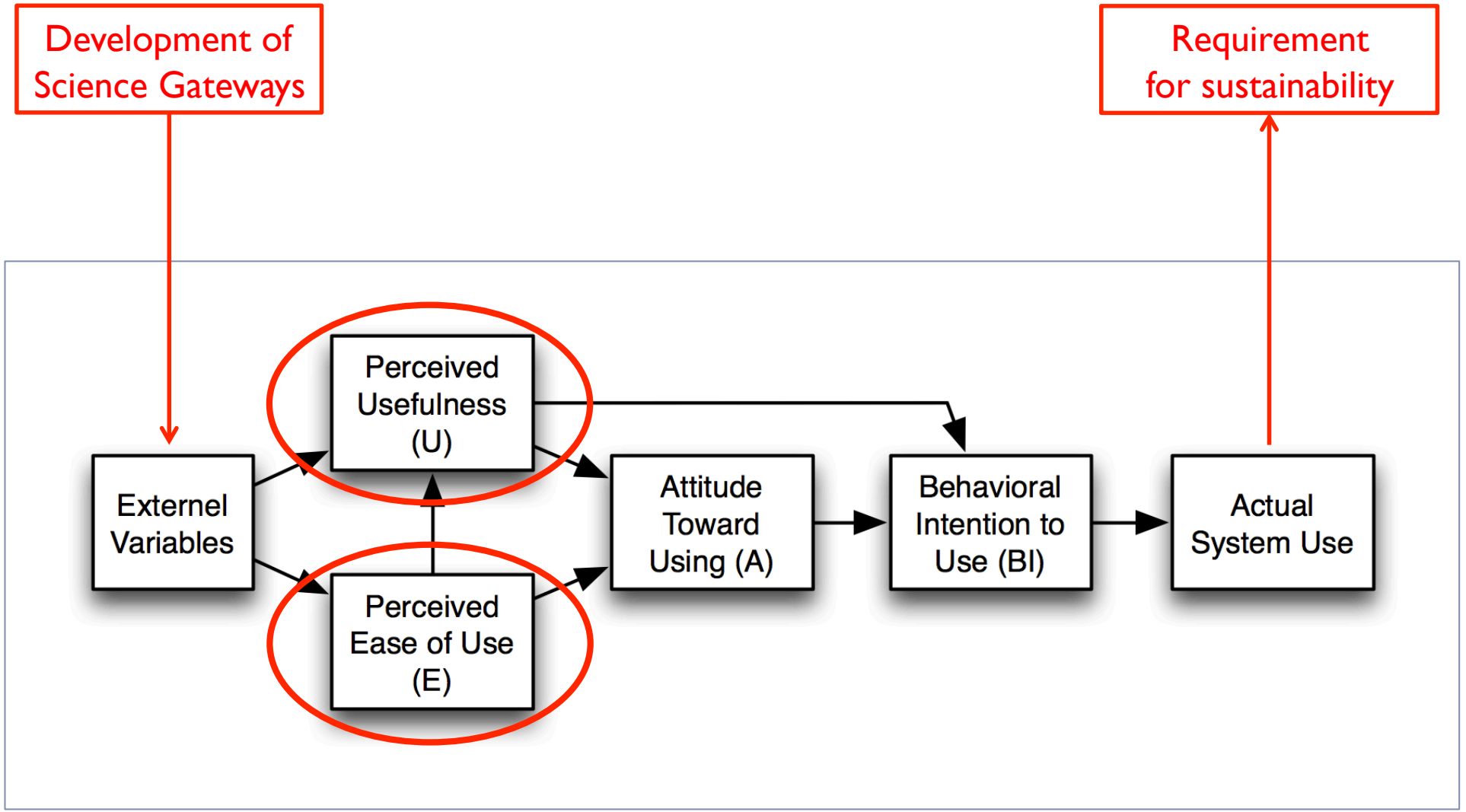
- According to ISO/IEC 2382-01 (*Information Technology Vocabulary, Fundamental Terms*), **interoperability** is "*The capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units*".

IT acceptance model – **the Web**



Davis, F. D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS Quarterly* 13(3): 319–340

IT acceptance model – **the Grid**



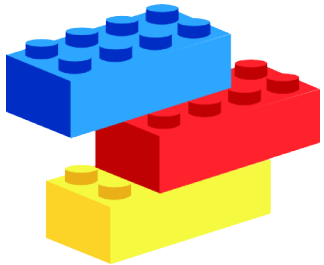
Davis, F. D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS Quarterly* 13(3): 319–340

Community-driven web portals have started to integrate Grid Tools and Applications

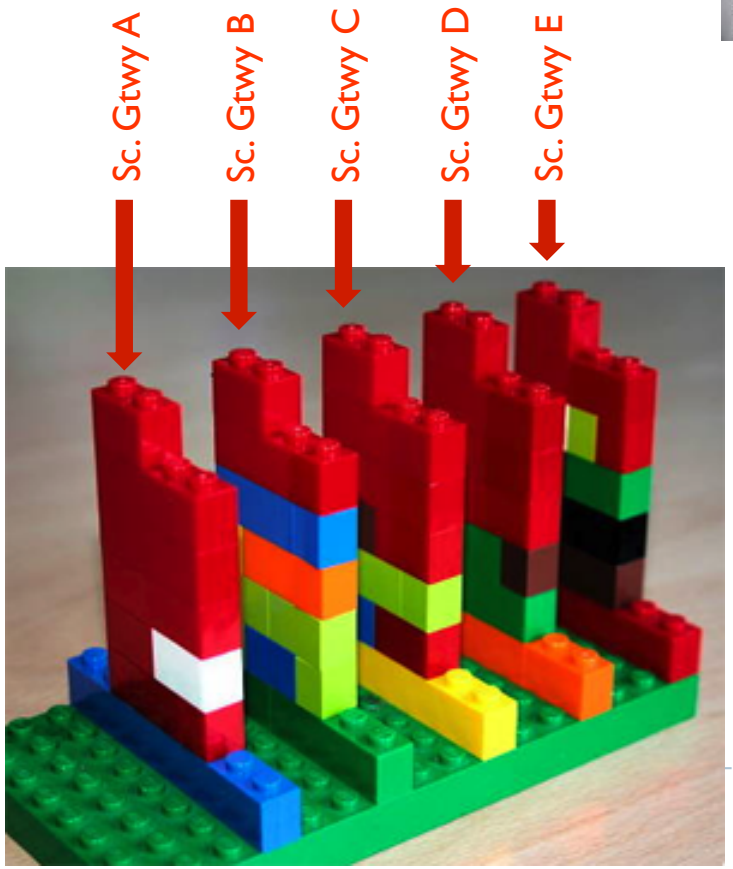
"A Science Gateway is a community-developed set of tools, applications, and data that is integrated via a portal or a suite of applications, usually in a graphical user interface, that is further customized to meet the needs of a specific community."

Teragrid/XSEDE

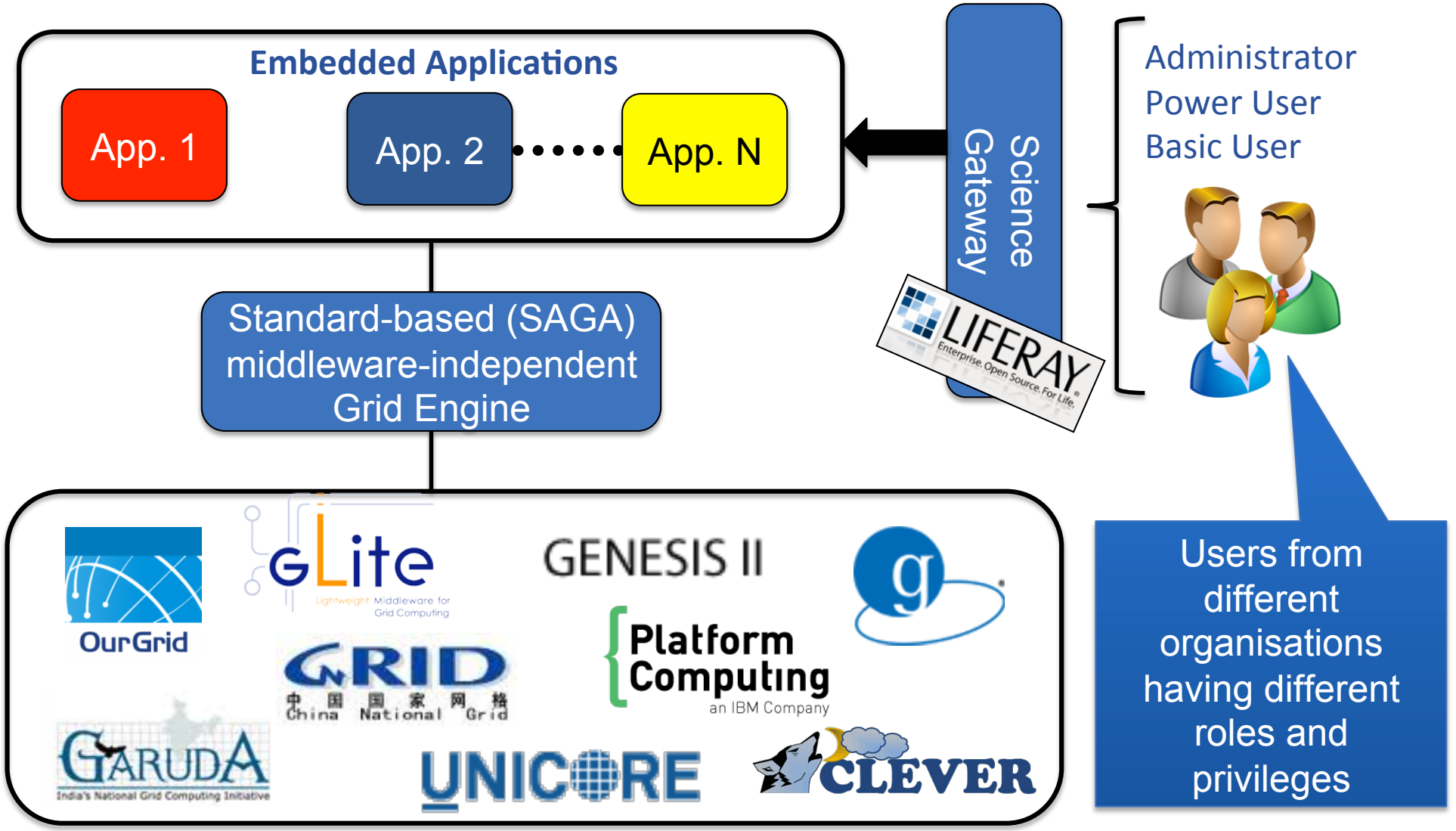
Primary requirement: building Science Gateways should be like playing with LEGO



- **Standards**
- **Simplicity**
- **Easiness of use**
- **Re-usability**

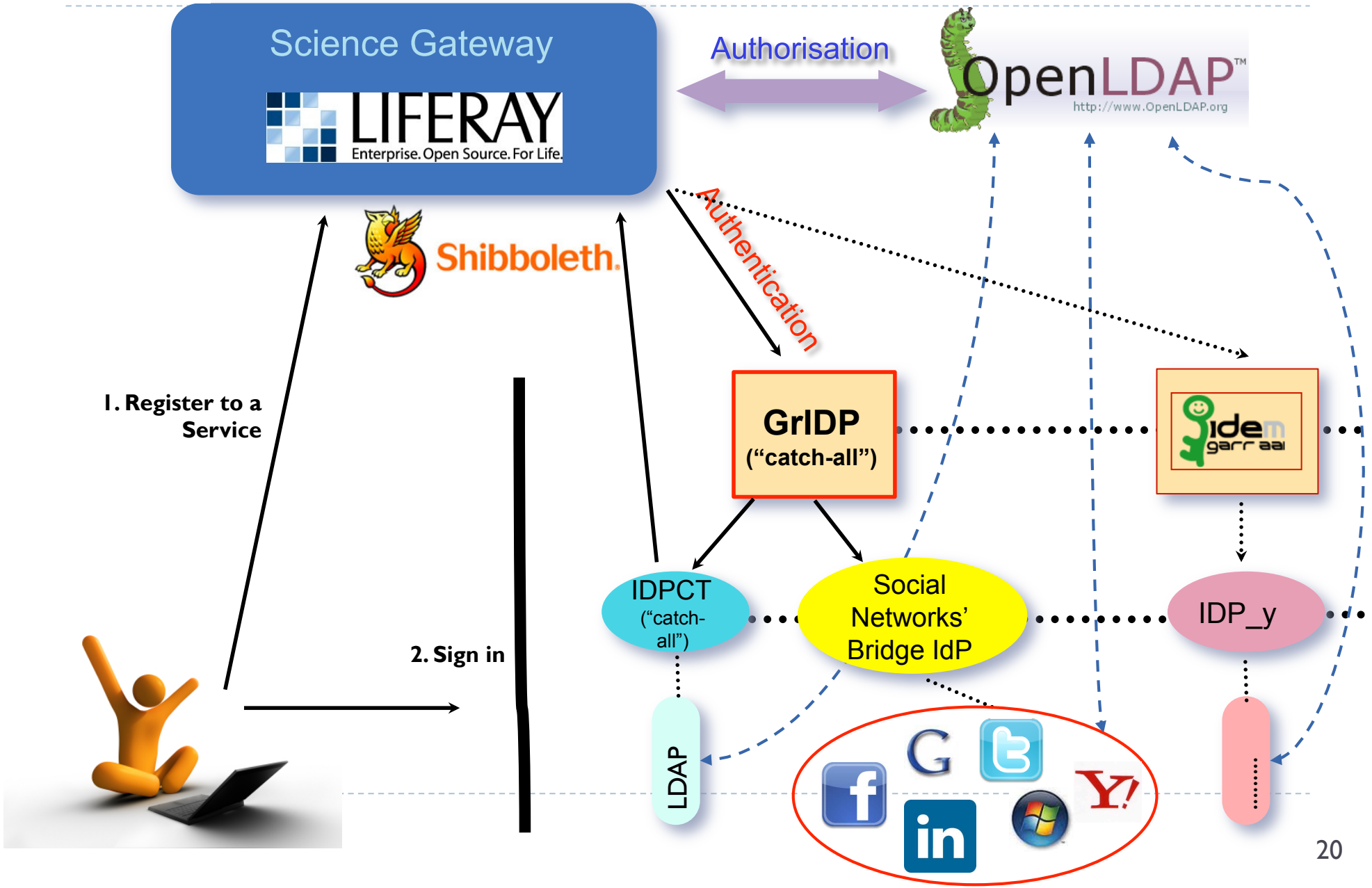


The Catania Science Gateway framework



Middleware supported so far

AuthN & AuthZ Schema

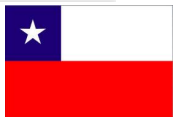


Official Identity Federations currently supported by Catania Science Gateways

cafe

CARSI

COFRe
Comunidad Federada REUNA



grnet

idem garr aai

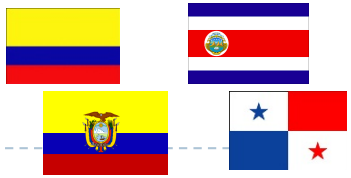
RCTSaai
REDE CIÊNCIA, TECNOLOGIA E SOCIEDADE

SIR

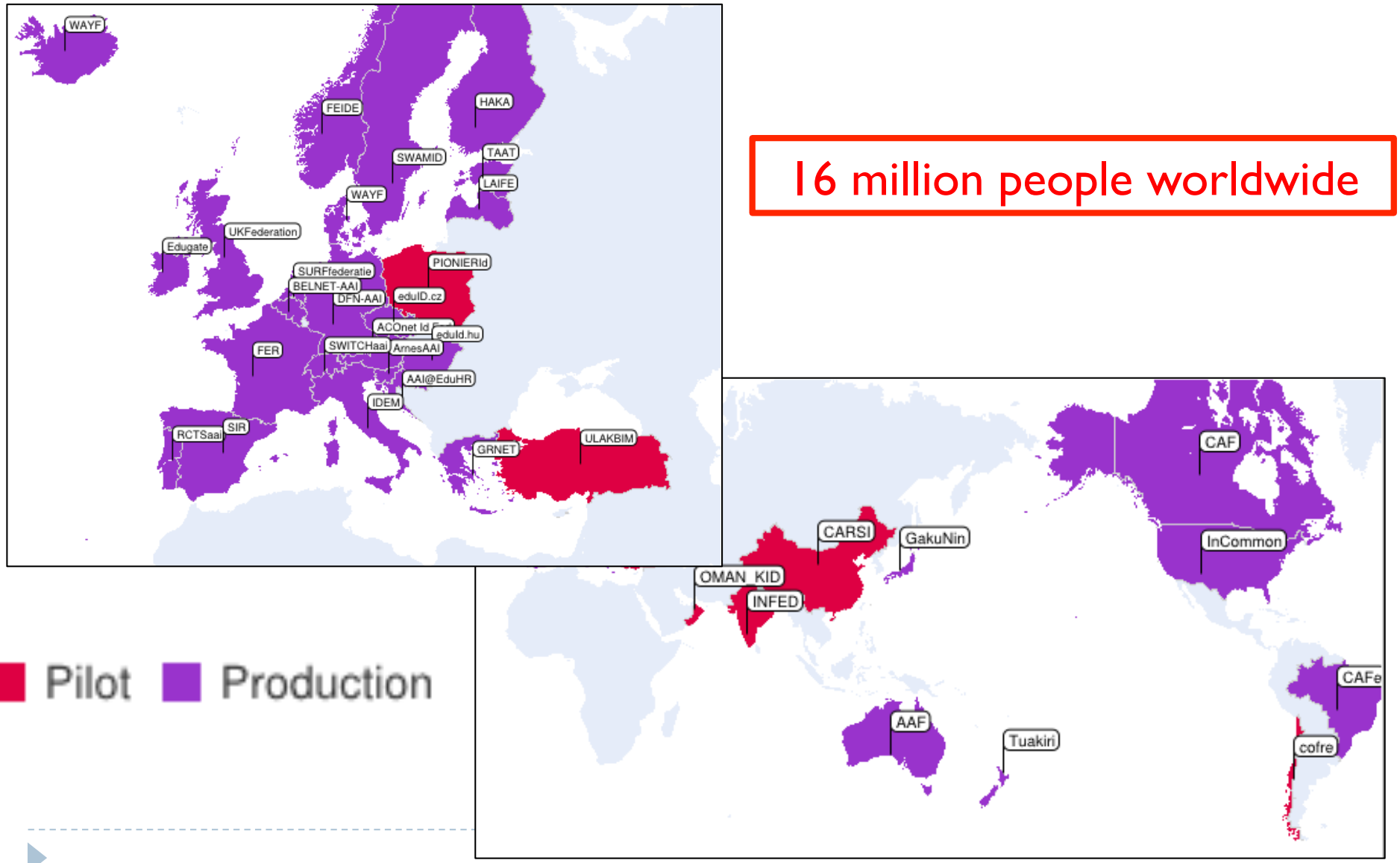
IdFMex

Universidad Nacional Autónoma de México
UNAM

To be created
26-30 Nov



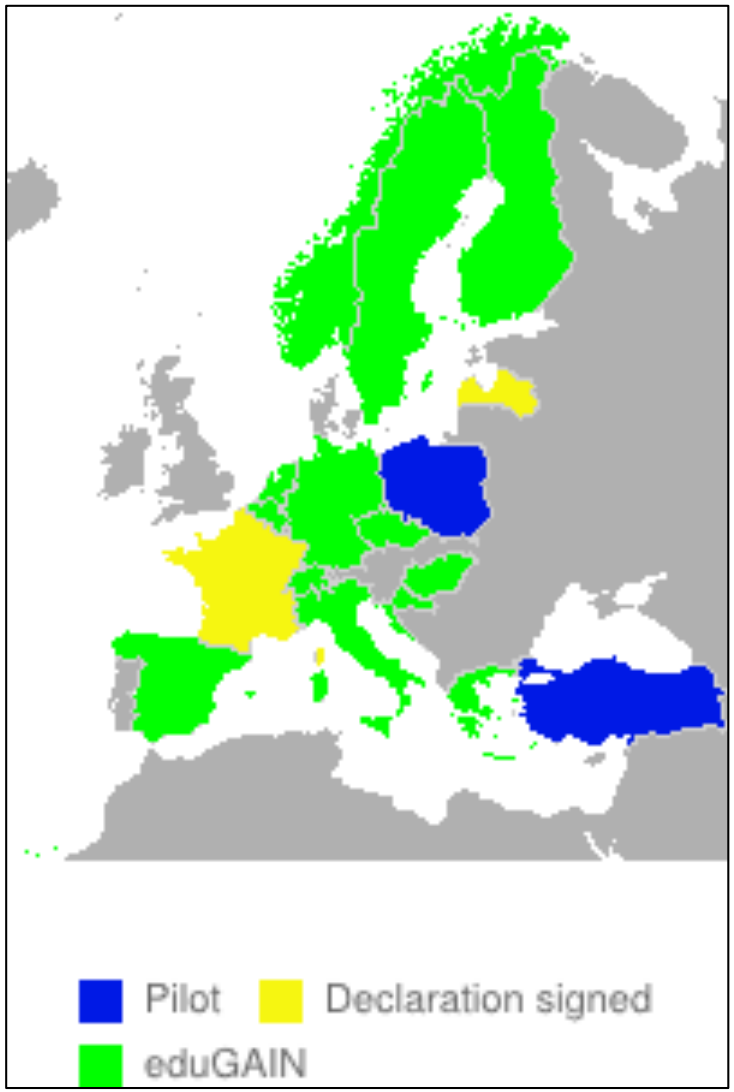
Identity Federations in the world
(https://refeds.org)

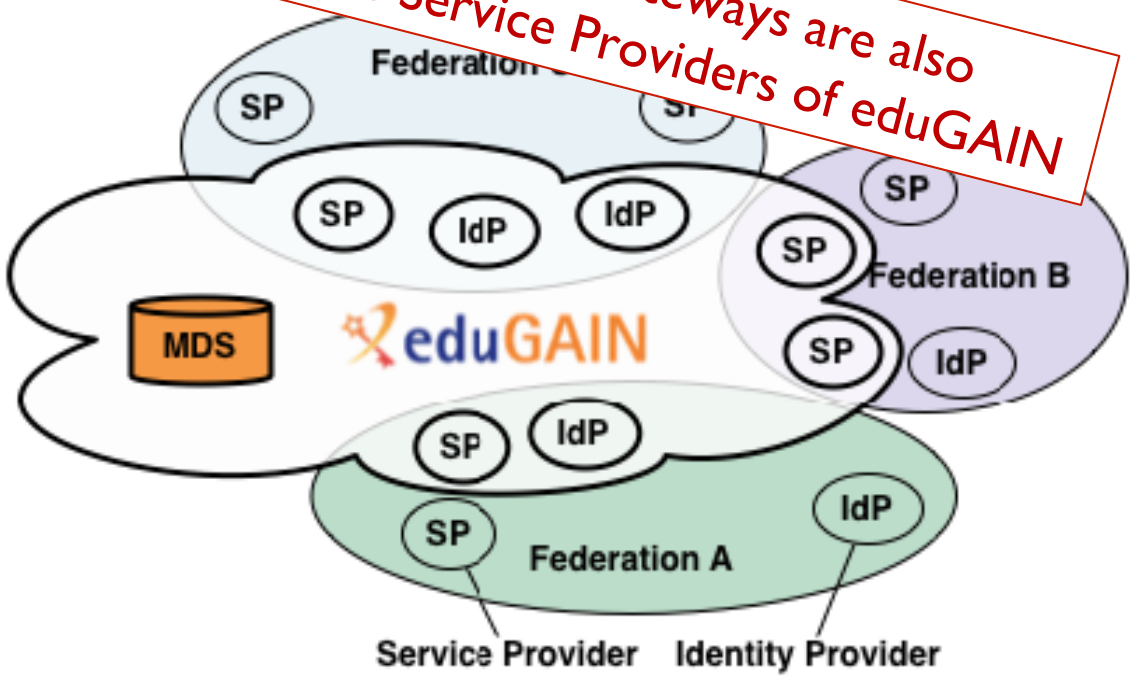


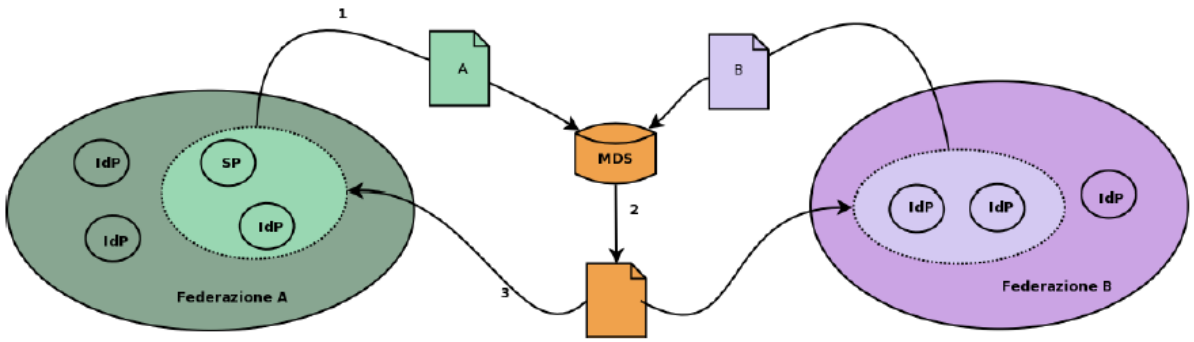
16 million people worldwide

eduGAIN
(www.edugain.org)

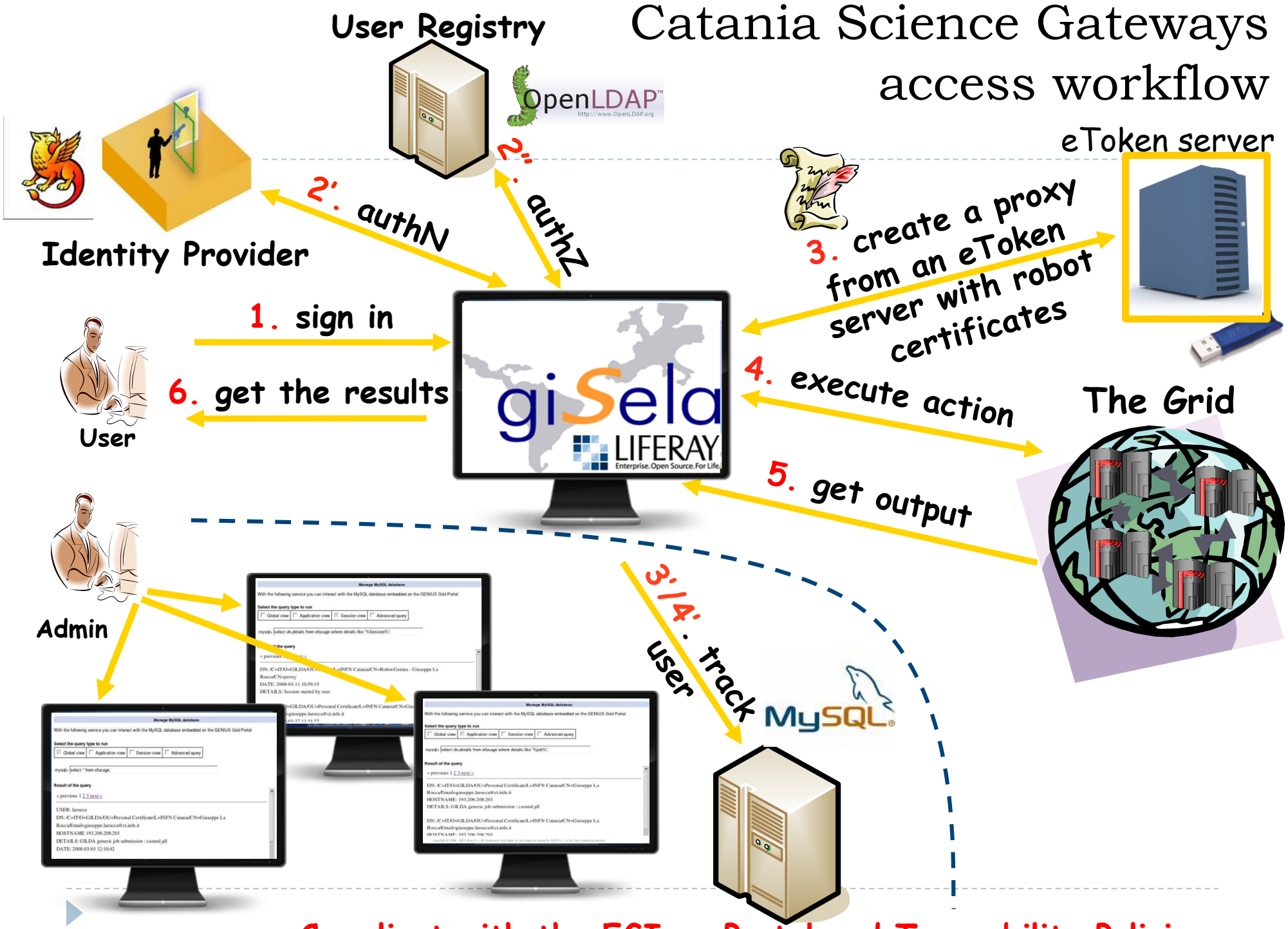
Catania Science Gateways are also registered as Service Providers of eduGAIN







Catania Science Gateways access workflow

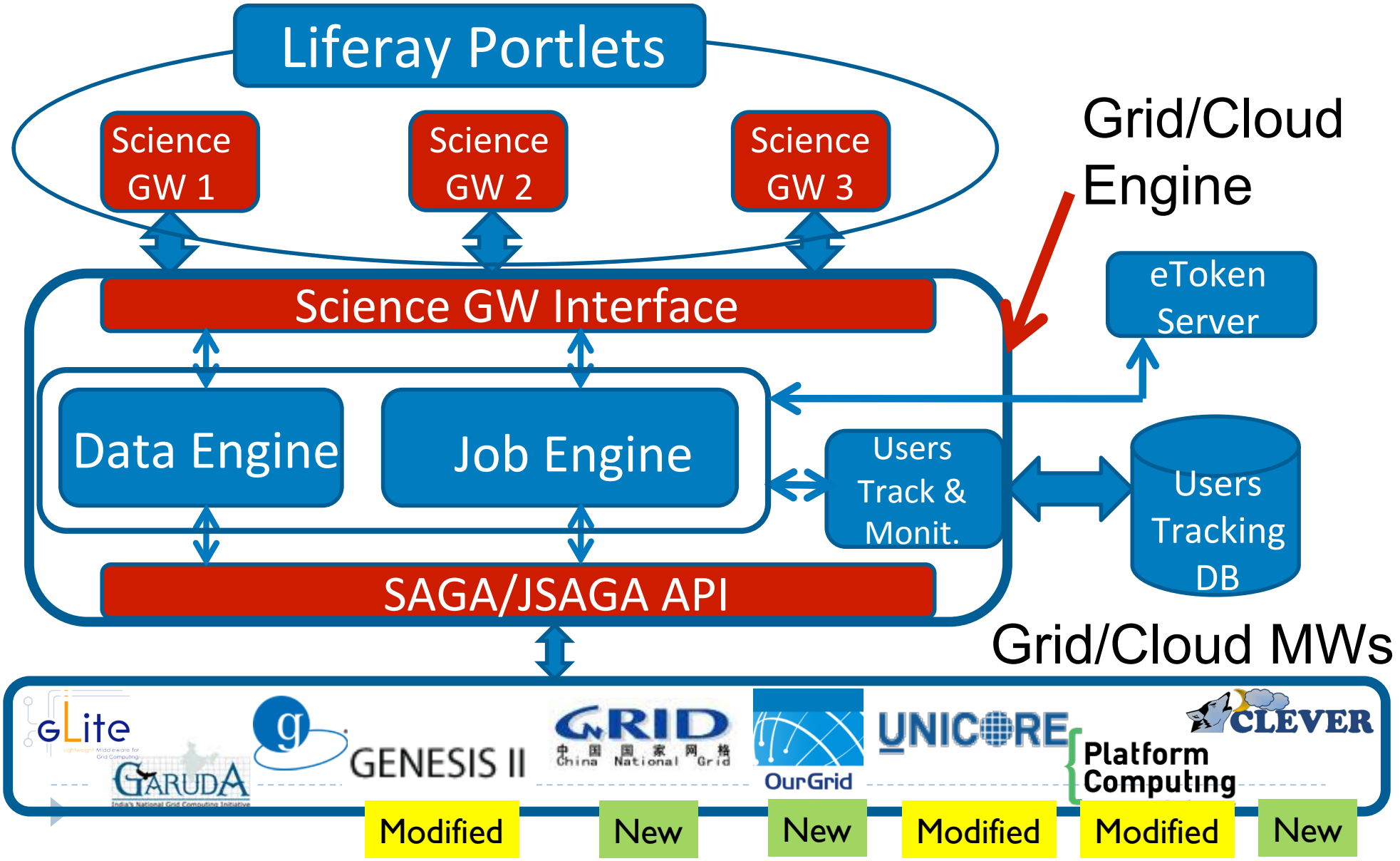


Compliant with the EGI.eu Portal and Traceability Policies

The CHAIN Worldwide Interoperability Demo
(http://science-gateway.chain-project.eu)

- To demonstrate that:
  - e-Infrastructures can be made interoperable to each other at user application level using standards
    - with the meaning of interoperability given in slide 14;
  - VRC-specific applications can be submitted from anywhere and run everywhere

The Catania Grid & Cloud Engine

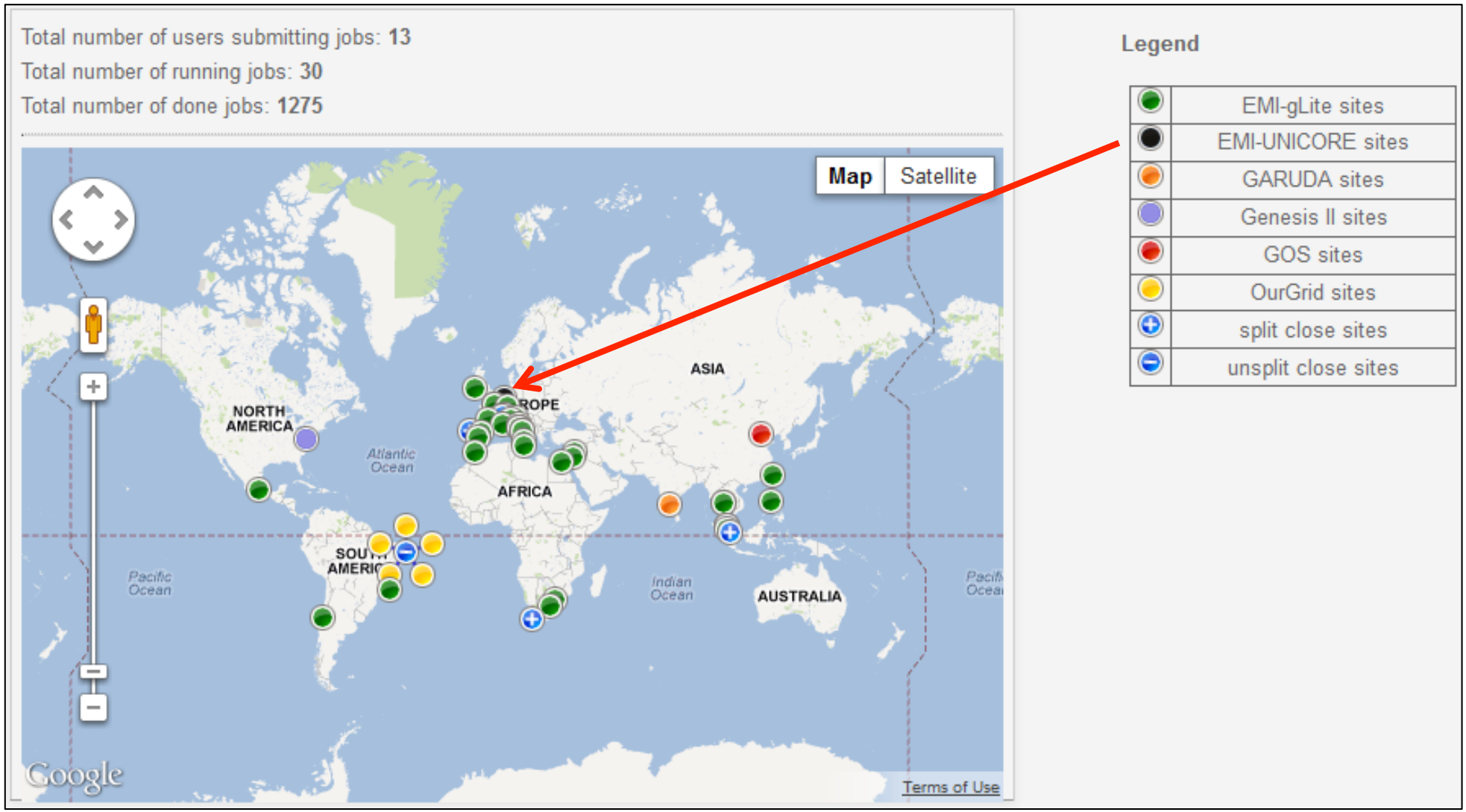


CHAIN Demo Contributors

GENESIS II gLite GOS JSAGA Platform Computing OurGrid UNICORE

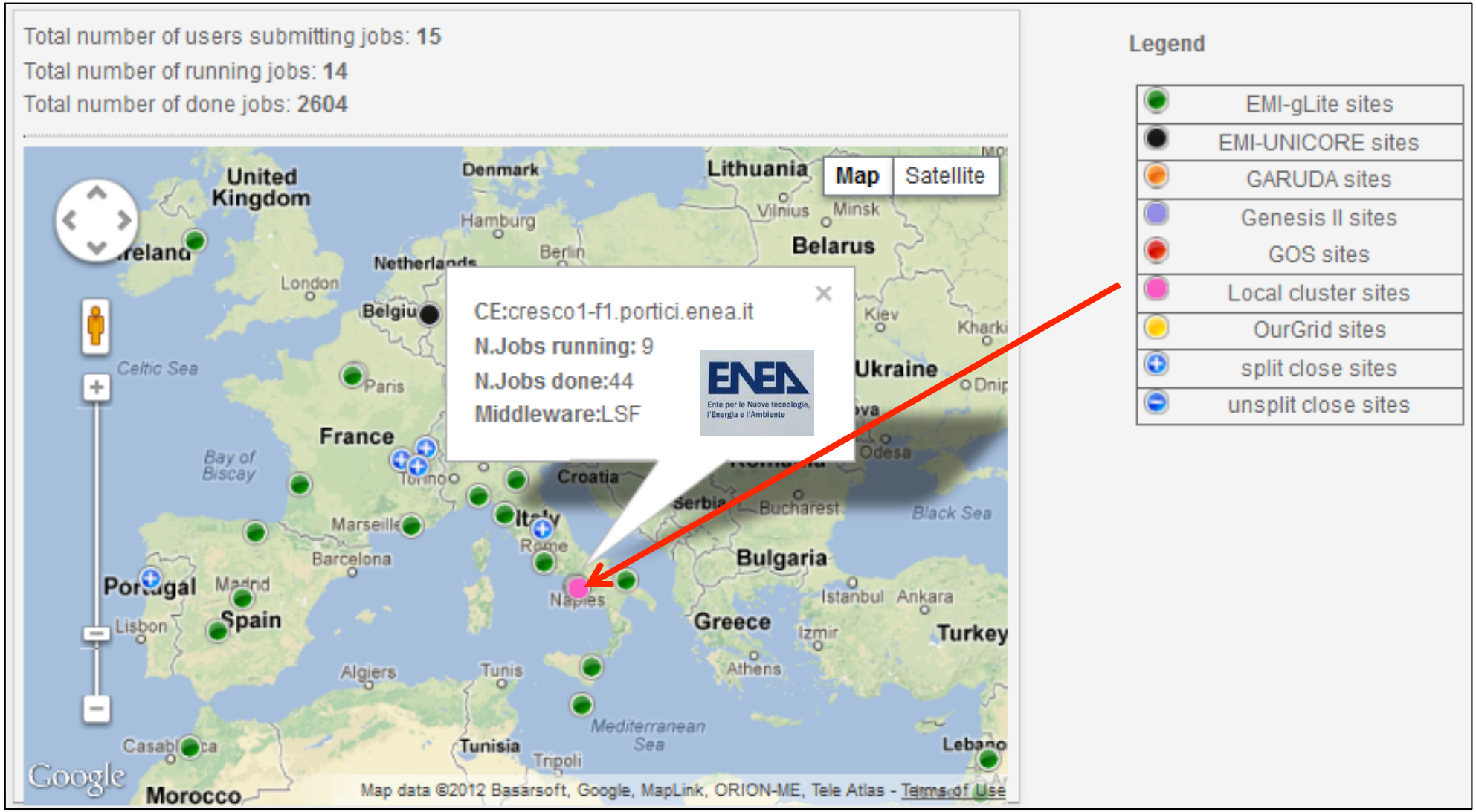
CHAIN Demo Status

(http://science-gateway.chain-project.eu/demo-status)



CHAIN Demo Status

(http://science-gateway.chain-project.eu/demo-status)



Science Gateways and clouds – The MyCloud service

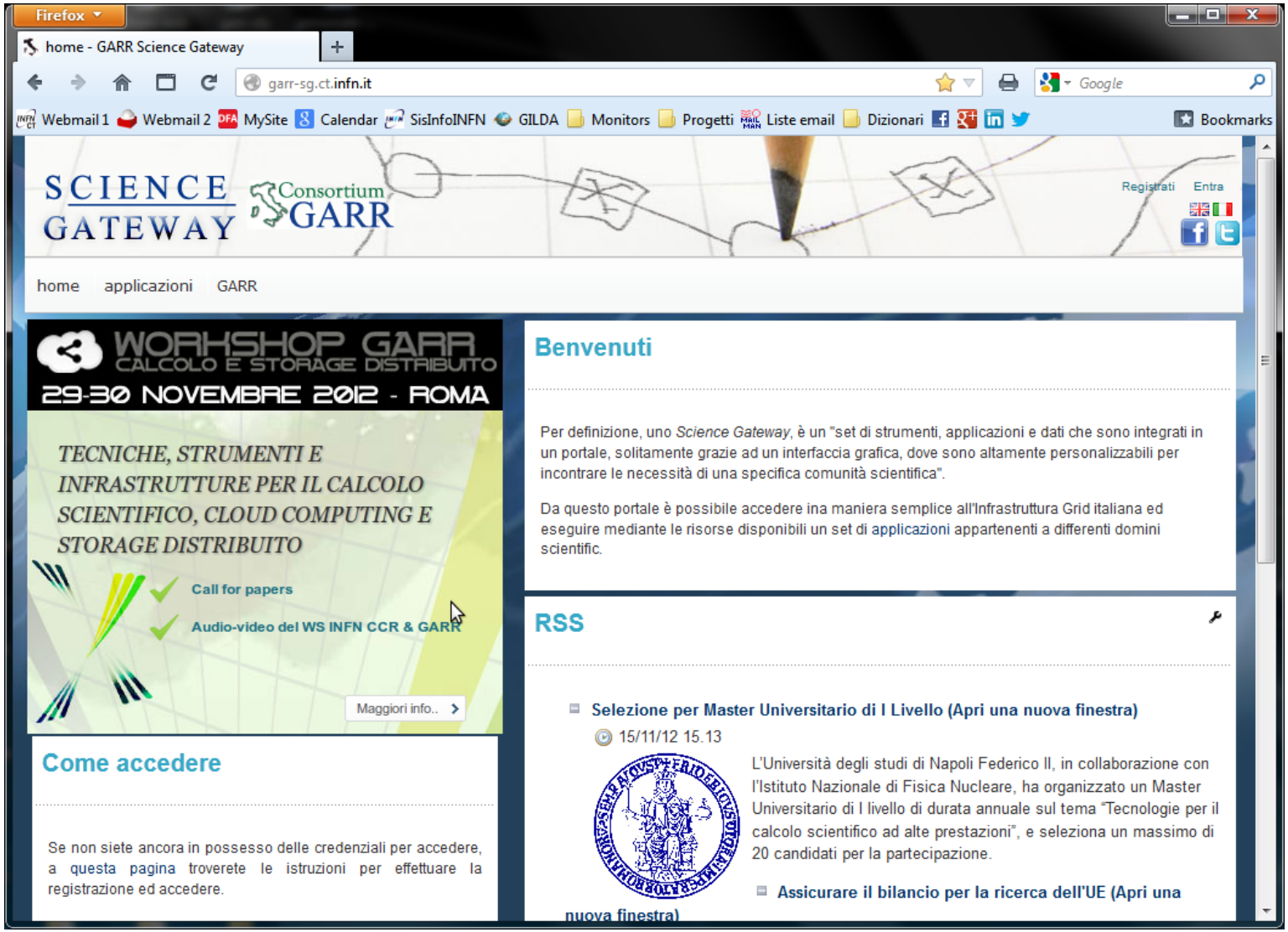
Powered by:

CLEVER

Summary and conclusions

- e-Infrastructures can be very beneficial platforms for many users, provided they are really «easy to use» and users are at their centre
- The Catania Science Gateway framework, with support for Identity Federations changes the way Grid infrastructures are used, hugely widening their potential user base across continents and organisations, especially non-IT experts and the "citizen scientist"
- The adoption of standards (JSR 286, SAGA, SAML, etc.) represents a concrete investment towards sustainability
- The CHAIN worldwide interoperability program demonstrates that, through Science Gateways based on standards, users can really access global e-Infrastructures in a seamless and ubiquitous way independently of the underlying middleware (local, volunteering, grid, cloud)
- **We propose the use the same approach to gather distributed resources from all over Italy and build a truly Italian e-Infrastructure, yet respecting local specificities and exploiting competences of all organisations interested in participating in this endeavour**

The GARR Science Gateway



Thank you !