The Open Access initiative in the context of the European ICT Research Infrastructure Service

Open Access Conference

Athens, 15 December 2008

Kyriakos Baxevanidis
European Commission - DG INFSO
Acting Head of Unit, GÉANT & e-Infrastructure Unit

e-infrastructure

European Commission
Information Society and Media

# *Political context and role of Science*

Research & Innovation defines the future of the nations – role of ICT

- Lisbon strategy: Innovation is the most important factor in determining Europe's success through next decades

- Innovation is the societal and economic manifestation of hope *("Innovate America", Dec 2004)*

- Innovation is fuelled by scientific research

- ICT: EU's most innovative and research intensive sector (standing for 25% of total EU research effort, 5.6% of GDP, and 45% of EU productivity growth in 2000-2004) *(EU i2010 Annual Report 2006)*

e-infrastructure
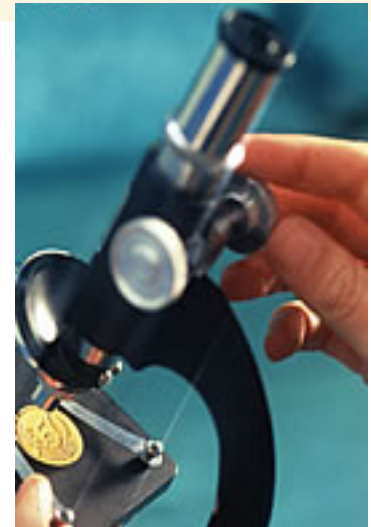
Science in the 21ˢᵗ century

- Global challenges with high societal impact (e.g. climate, environment, energy...)

- Big challenges... need of multi/inter-disciplinary research... global collaborations (spread of skills and knowledge)... sharing of resources

- Data deluge... born digital material

- Mediating/enabling role of ICT... addressing complexity

e-Science

computational science

theoretical

experimental

empirical
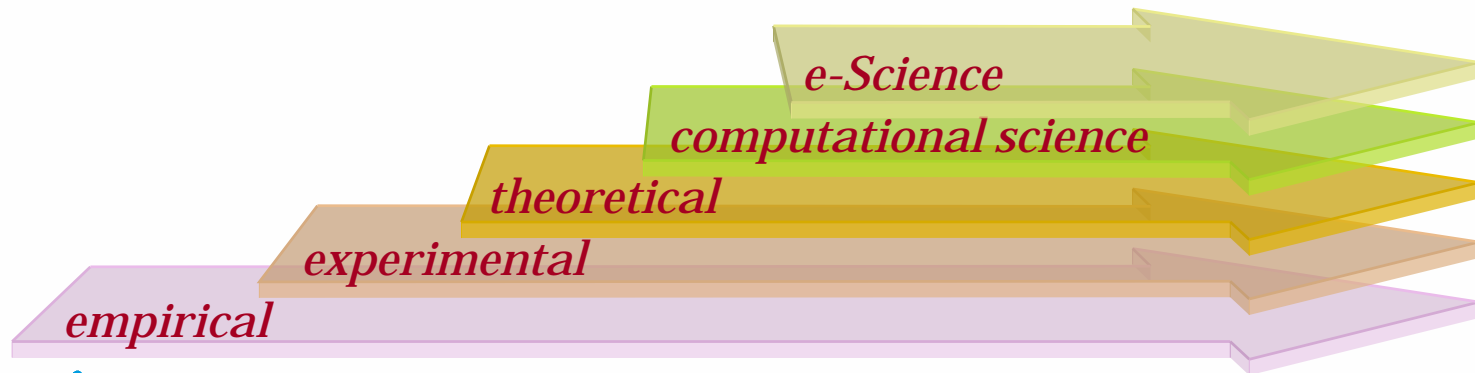
e-infrastructure

e-Infrastructure:
- ICT facilities and services enabling scientific research
## *- supporting Open Access implementation*

e-Infrastructure - main service categories

- **Connectivity**
- **Access to mash commodity computing & storage resources (grids...)**
- **Access to High End Computing (HEC)**
- **Access to data across organisational and scientific-discipline boundaries**
- **International links**

e-infrastructure

Connectivity service (GÉANT)

Connectivity service model

- Full-fledged operational service to all research institutes in Europe

- "One-stop-shop" service provision on National (NREN) & EU-level (DANTE)

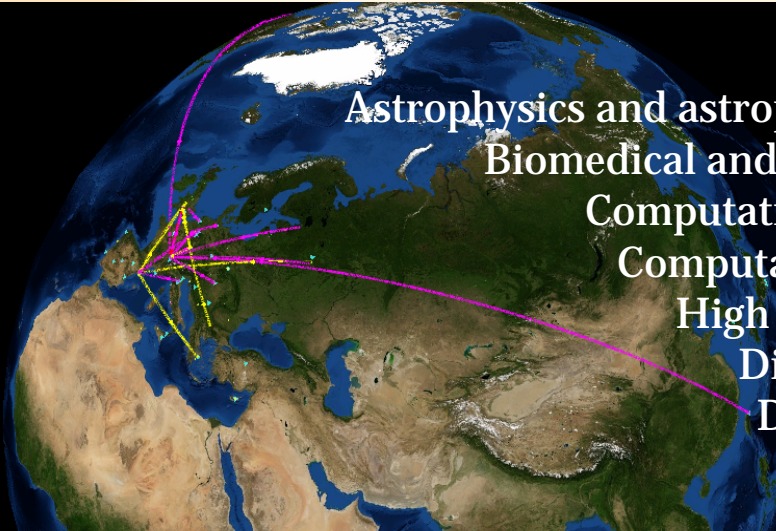- Model: agreed sharing of connectivity-costs among partner-countries; stable funding scheme
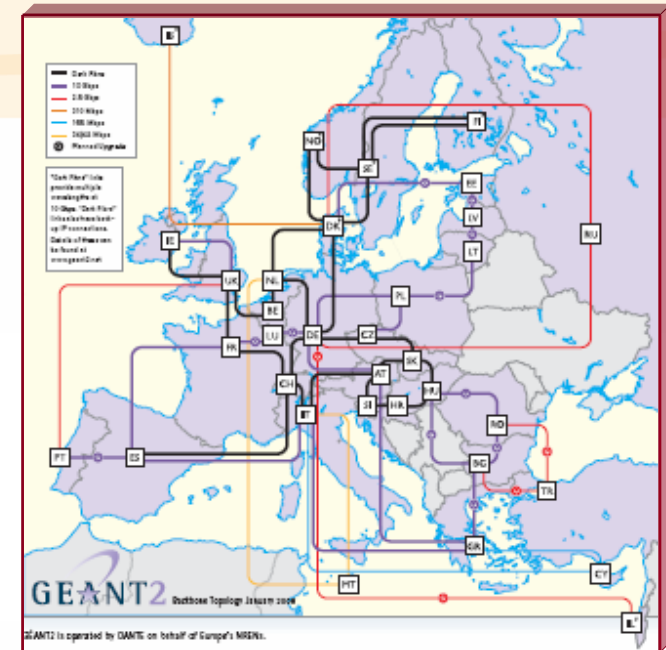
- Pan-European coverage
  (40+ countries /3900 universities / 30+ million students)

- Hybrid architecture:
  - connectivity at 10 Gb/s (aggregated traffic)
  - dark fiber wavelengths (demanding communities)

e-infrastructure

Access to mass commodity computing & storage (EGEE III)

Scheduled = 17356
Running = 18359

Astrophysics and astroparticle physics
Biomedical and bioinformatics
Computational chemistry
Computational sciences
High Energy Physics
Disaster recovery
Digital Libraries
Earth sciences
Infrastructure
Geophysics
Finance
Fusion

- >250 sites
- >60 000 CPUs, 25 Pbyte of storage
- ~150 000 jobs successfully completed per day
- 200 Virtual Organisations
- >8000 registered users, representing 1000s of scientists

- *Model based on pooling of computing resources using common or interoperable middleware platforms*
- *Middleware developments heavily based on Open Source model*

e-infrastructure

EGEE/National service provisioning model

● Regional Operations Centre

○ GGUS (Grid User Support Centre)

□ Operations Coordination Centre

**National Grid Service, Grid Operations Service Centre**

- **Regional Operation Centres (ROCs) and Regional Support Teams** for basic service operation and first level support on the national or regional level; those are based on national or regional grid service centres

- **Operations Coordination Centres (OCCs)** as more central entities that manage the operation of the entire EGEE Grid from a single centralised location

- **Grid User Support centres (GGUSs)** coordination and oversight for user support in EGEE; they provide a well-known place where all user queries may be directed to ensure that they are addressed

- **VO Support** registration, treatment of VO issues
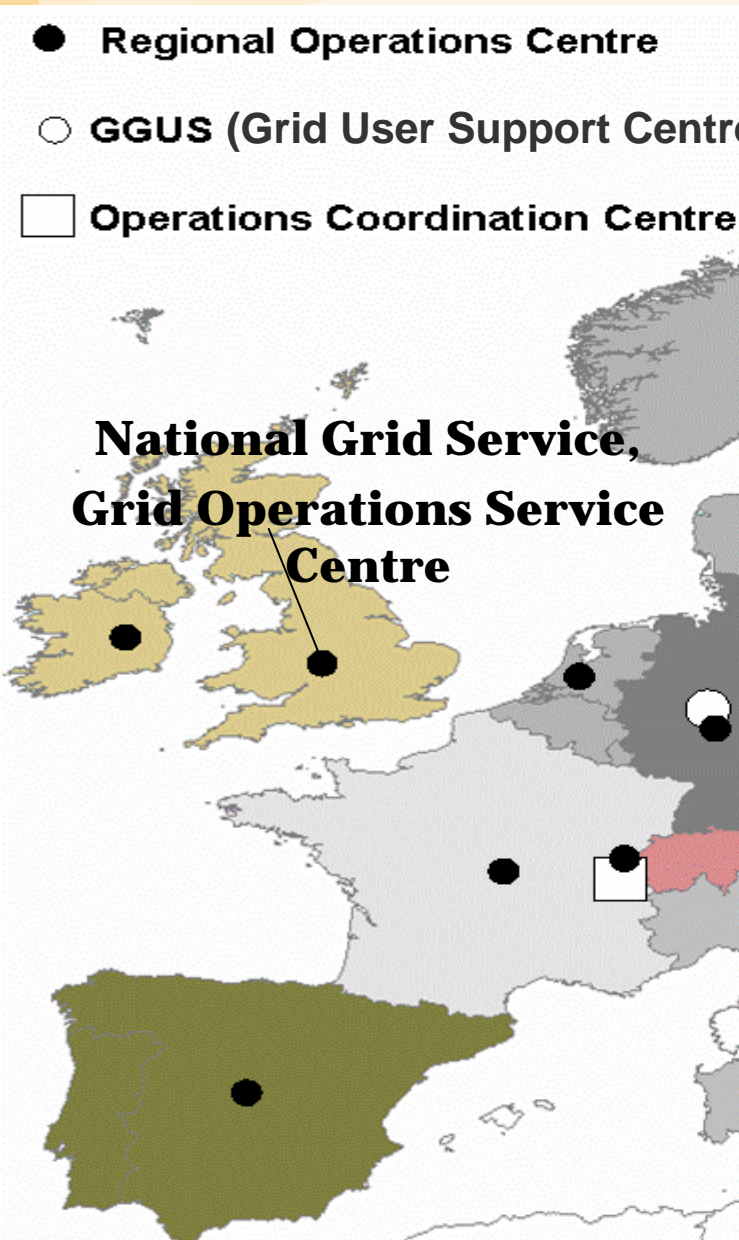
Example press article on EGEE

EGEE Grid attacks Avian Flu

During April 2006, a collaboration of Asian and EU laboratories has analysed 300,000 possible drug components against the avian flu virus H5N1 using the EGEE Grid infrastructure (for the docking of 300,000 compounds against 8 different target structures of Influenza A neuraminidases, 2000 computers were used during 4 weeks – the equivalent of 100 years on a single computer)

H5

N1

Credit:
Y-T Wu

e-infrastructure

European Commission
Information Society and Media
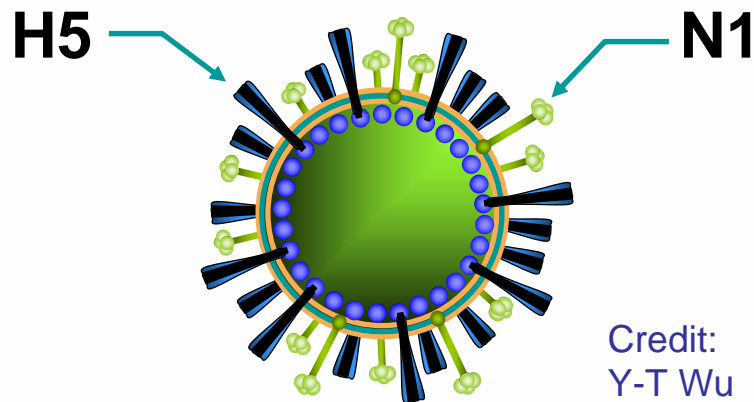
Current grid-based resource-provisioning model under evolution

EGI_DS (Design study)

NGI-
Nordic States

Coordinating entity (EGI)

Key elements of new scheme:

- Coordination/pooling of National and European funding for sustainable service provision and for more efficient planning of investments

- Service provisioning beyond project cycles

- One-stop-shop service provisioning (including training) to users who want to access grid-based resources

NGI-GR

e-infrastructure

European Commission
Information Society and Media

DEISA2: access to virtual HEC resources

- 11 sites/7 countries connected at 10 Gb/s

- Model: *pooling* of HEC-time for grand scientific applications (EU level)

- *Aggregated* peak performance close to 1 Petaflop

- Global data management service
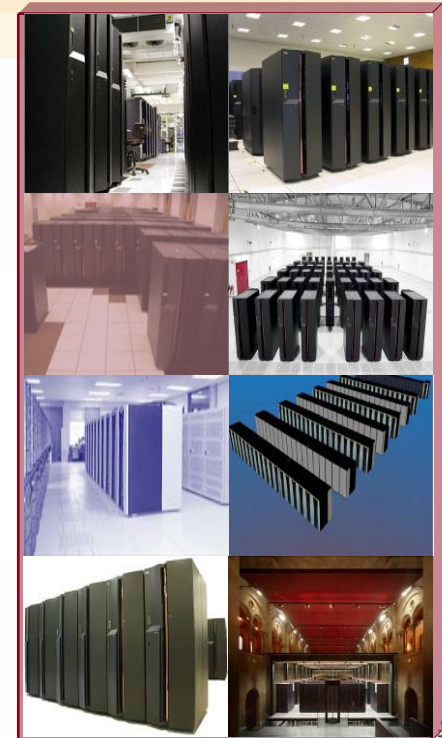
Distributed
European
Infrastructure for
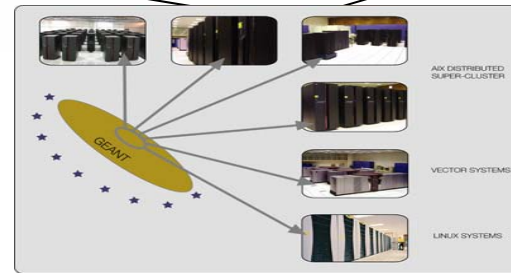Supercomputing
Applications

European Commission
Information Society and Media

e-infrastructure

DEISA2 service provisioning model

Service layer on top of National supercomputing service schemes using grids

Enabling new applications through re-engineering national production quality codes to operate on DEISA (EU-level) grid

DEISA Extreme Computing Initiative (DECI) enables use of DEISA by "grand challenge" applications in science and engineering
- Calls for proposals to select new applications (many in operation, e.g. Cosmology, Materials Science, Fluid Dynamics, Biophysics)
- Support by Applications Task Force (ATASKF) (consultancy...)
- Service differentiation for different classes of applications

e-infrastructure
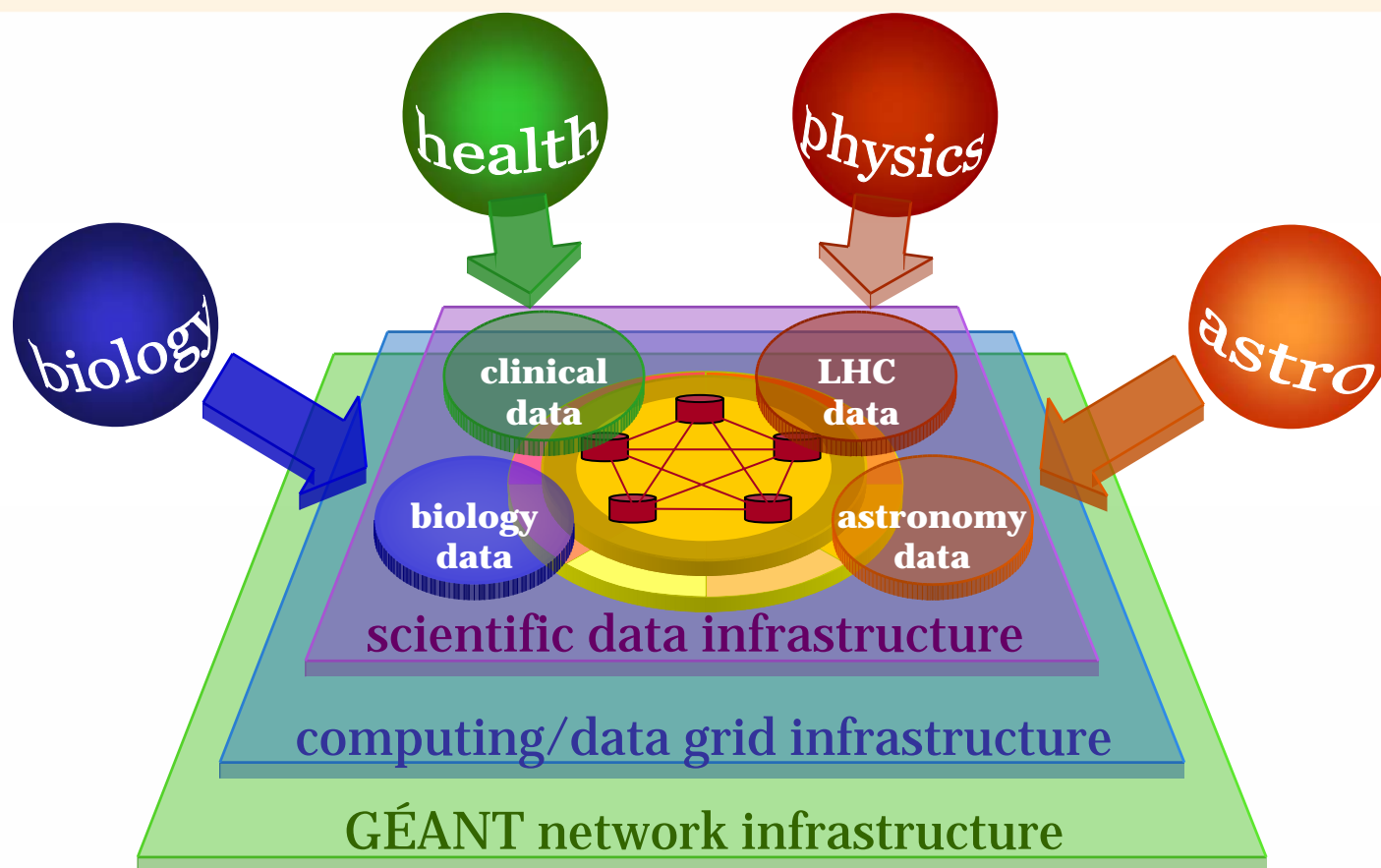
PRACE: new generation HEC resource (petaflop scale)
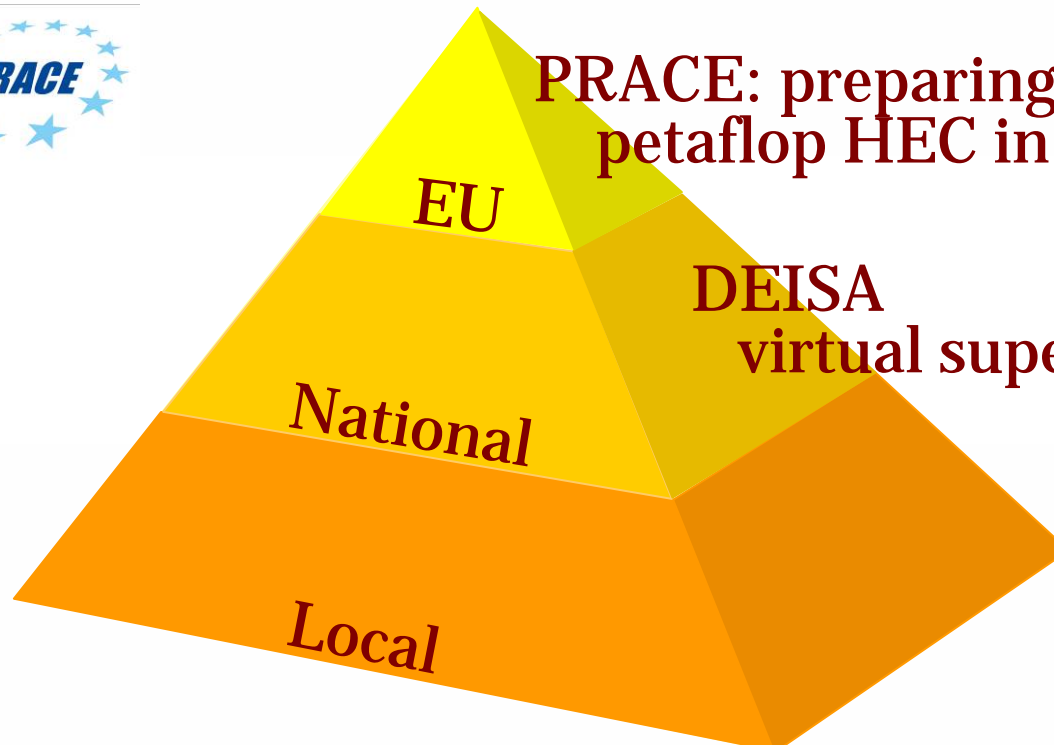
PRACE: preparing framework for petaflop HEC in Europe

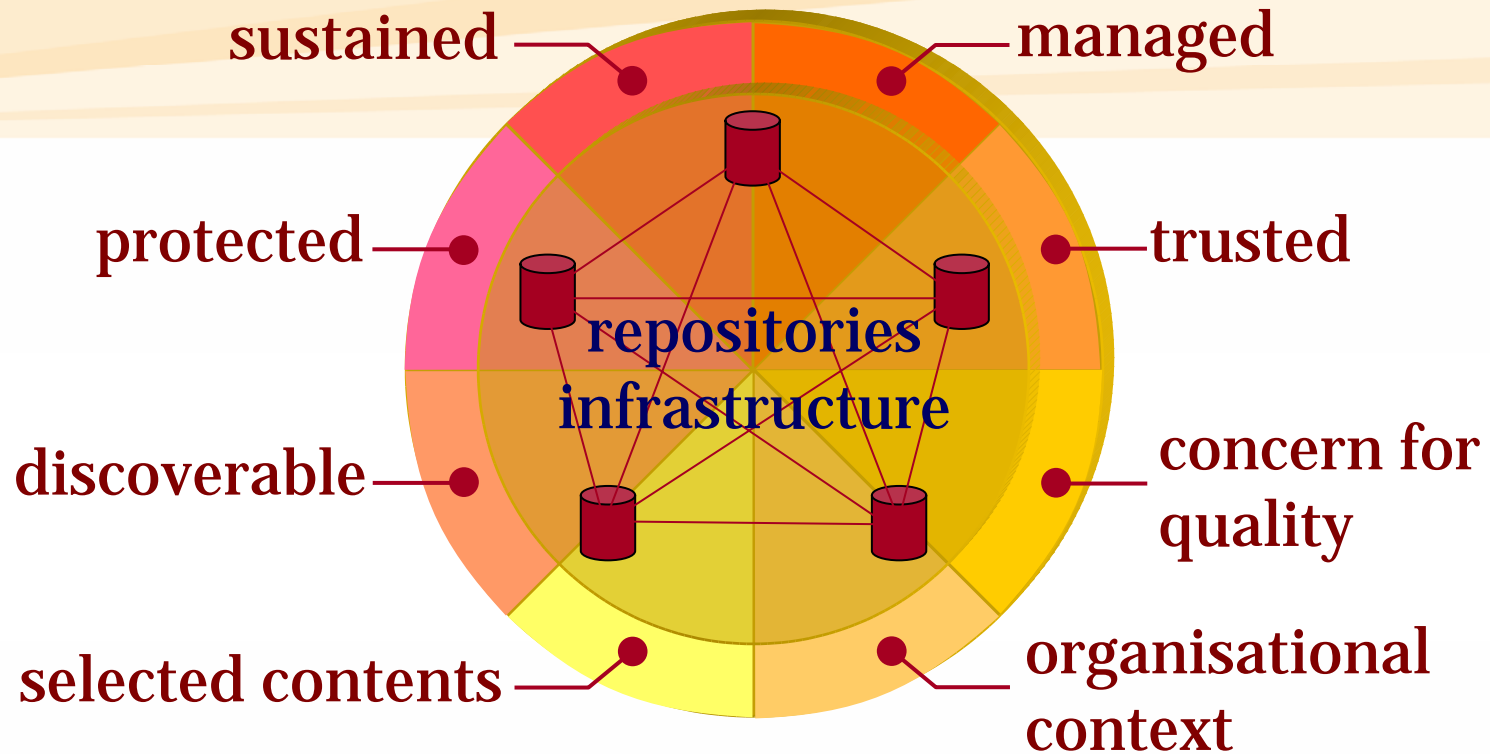**EU**

DEISA
virtual supercomputer access

National

Local

- *For the first time agreement among Europeans to co-fund such an infrastructure*
  - *Shared use model*

e-infrastructure

European Commission
Information Society and Media

Access to data (data as an infrastructure)

Focus: repositories, enhanced use, sharing

sustained · managed · trusted · concern for quality · organisational context · selected contents · discoverable · protected

repositories infrastructure

Address key requirements on data infrastructures:

- Data preservation, curation
- Enhance extraction of meaning from & use of data within scientific disciplines and across them
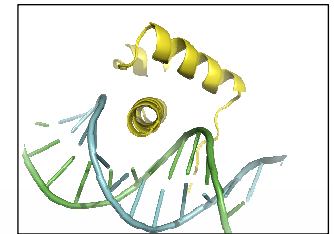- Incentives for sharing

e-infrastructure

source: eSciDR study (adapted)

Examples of data infrastructure projects

Flexible, robust, scalable and cohesive pan-European infra-structure of Digital Repositories

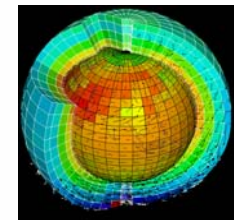
Improving protein annotation through coordination and integration of databases

Common Information Model and tools for using climate data and models

e-Infrastructure Prog. supporting further the involvement of user communities

- **Funding provided for linking of new user communities to existing infrastructures ("<u>User Communities</u>" calls for proposals of e-Infrastructure programme)**

e-infrastructure

European Commission
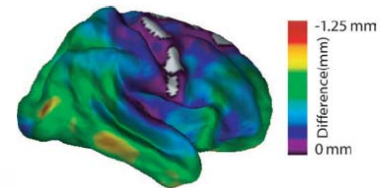Information Society and Media

Examples: involvement of new user communities using targeted funding
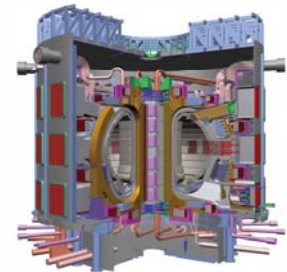
VOs on seismology, meteorology and environment, supported by south-eastern Europe grids
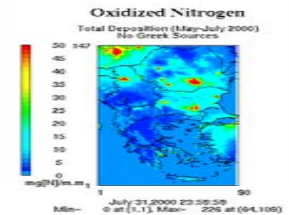
Grid for European neuroscientists working in the field of imaging of Alzheimer's disease

Modelling capabilities for ITER and future fusion devices using parallel Grid computing and HPC

New EU initiative: Open Access (OA)

http://ec.europa.eu/research/science-society/open_access

e-infrastructure

What is Open Access?

- **Alternative mechanism (e.g. to that of peer reviewed publications in scientific journals) for scientific information to be openly accessible at no cost over the internet**
  - o **Scientific information can be of any form (articles, data etc)**

e-infrastructure

European Commission
Information Society and Media

Open Access and e-Infrastructure

In implementing the Open Access policy there is a need to address questions like:

- **Where to store my scientific information if no on-line facility is available?**

- **How to ensure common EU-level standards in accessing such scientific information?**

- **Where to get help/training?**

e-infrastructure

European Commission
Information Society and Media

Call for Proposals to support technical implementation of OA-pilot

Call: INFRA-2009-1.2.3 - Scientific Information Repository supporting FP7

- **Call published on 9 Dec 2008; deadline 17 March 2008; budget: €4 m**
- **Seeks projects to support the establishment & operation of a technical infrastructure of digital repositories to deposit & access scientific articles & data produced under FP7 (incl. those of European Research Council)**

Call for Proposals to support technical implementation of OA-pilot

Facility should among other services provide:

- **Helpdesk and 24/7 operational service**
- **Post print authoring tool to allow researchers to convert articles into more readable & user friendly formats**
- **Monitoring data and statistics on use**
- **Information and guidance on use (training...)**
- **Preparatory action towards extension of e-Infrastructure ability to cope with storage, management & access of scientific data**

Further information: http://cordis.europa.eu/fp7/ict/e-infrastructure/
Relevant Work Programme: Research Infrastructures 2009
ftp://ftp.cordis.europa.eu/pub/fp7/docs/wp/capacities/infrastructures/
n_wp_200901_en.pdf

Summary

- **It is important to further support the Research and Innovation process in Europe**

- **New *Open Access* EU-policy aims at accelerating and enhancing innovation**

- **e-Infrastructure in the front line of supporting pilot implementation of Open Access policy**

European Commission
Information Society and Media

e-infrastructure

The real voyage is not about discovering new lands but about looking things with new eyes...

Thank you