



### High capacity virtual infrastructures

**Dr. Malcolm Read:**  
e-Infrastructures covers many layers - networks, high performance computing, middleware (e.g. grids), research data and primary and secondary (e.g. libraries) resources. e-Infrastructures is applicable to research but also to education, health, industry, e-government, etc...

**Dr. Vasilis Maglaris:**  
GEANT2 creates the most advanced and wide e-Infrastructure worldwide. It combines IP routed services (to millions of researchers and students) and switched end-to-end Gigabit light-paths (for demanding scientific communities).

**Dr. Andrea Granelli:**  
Digital contents and archiving creates big opportunities that challenge e-Infrastructures, being new competences and design paradigms required.

### Long term preservation of scientific and cultural data

**Dr. Peter Tindemans :**  
Accelerating and coordinating long term access to digital records of science requires a pragmatic action at European level, through a truly public-private partnership. Scientific communities producing "vertical" repositories and "horizontal" digital archives are central.

**Dr. Luigi Fusco:**  
Data preservation is key for scientific communities such as "Earth Observation" and "Earth Sciences". ESA deploy an infrastructure for data acquisition, archiving and dissemination fostering long term Earth Science applications.

**Dr. Neil Beagrie:**  
Current and future activities designed to accelerate the digital preservation agenda in UK are part of the mission of the Digital Preservation Coalition (DPC); the follow up of the Needs Assessment survey will be applicable to a very large global context.

### Data curation

**Dr. Liz Lyon:**  
Adding value mechanisms (such as annotation, linking and knowledge extraction methodology) to large volumes of databases from science, engineering, social sciences and arts/humanities research can provide a maximum benefit to the a wider community.

**Dr. Henk Harmsen:**  
The DANS (Data Archiving & Networked Services) creates a single, transparent data infrastructure and the "thematic development programmes" (TDPs) foster close cooperation with research communities to configure and add value to it for a particular discipline.

**Dr. Manuel Delfino:**  
To function properly and be more competitive, institutions need to derive knowledge from stored and archived data. Under the "utility" scheme, data storage can be "outsourced" and curated by e-Infrastructure schemes generating additional assets.

### Scholarly communication

**Dr. David Nicholas:**  
Examining / profiling behaviour of global virtual scholar, from practical experiences with several digital libraries/millions of users - (characteristics: interested, horizontal searchers, volatile, promiscuous, with short attention span...), helps defining an European strategy.

**Dr. Leo Waaijers:**  
The DARE initiative (and subsequent actions such as "Cream of Science" and "HunDAREd Thousand") is promoting an innovative way of accessing scientific results and seeks its extension at European level.

**Dr. Robert Campbell:**  
Publishers need to understand the context in which researchers operate/factors that promote/inhibit their work. Surveys show factors are common (research areas/continents) and relate to organisational and human issues; productivity "liberators" were identified.



## Third European Conference on Research Infrastructures

**Main conclusion of Session 4**  
**“User needs in IT and  
 Information Infrastructure”**

### High capacity virtual infrastructures

- e-Infrastructures comprise: networks (internet, light paths...); computers (clusters, HPC...); access controls (security, AAA...); middleware (metadata...); finding tools (portals, search engines...); digital libraries (bibliographic, text, images, sound...); research data (national/scientific databases, individual data...)
- Identified challenges: HPC (capability computing at Petascale performance), Authentication and Autorisation mechanisms
- There is a need to move beyond research to reach out new communities: libraries, teaching, administration
- GEANT2, with a global dimension, addresses the needs of individual researchers (millions) but the focus is moving to support e-Science demanding projects.
- The area of digital archives (big opportunities but also new challenges and threats) requires support for the deployment of shared European infrastructures, the adoption of critical technologies (e.g. Digital Right Management Platform, Search engines), new competences and a new design paradigms

### Long term preservation of scientific and cultural data

- The session demonstrated that beyond the 3 classic infrastructures (network, grid, supercomputers) EU needs to address the data preservation as an infrastructural element.
- Consideration of records of science in digital cultural heritage evolved from 'records of history of science' to 'records of science in operation'
- In the data preservation domain, gearing up is best done by focusing on 'Records of Science' (good momentum, broad fields)
- Long term preservation of raw data and data correlation is fundamental for communities such as earth sciences
- DPC is a platform that should not just be understood as a technical infrastructure – social and organisational issues are key.
- National organisations speed-up the way digital preservation capacity of a country can be enhanced and it is vital to acquire a pan-European dimension.

### Data curation

- Data curation is a key element in a reinforced e-Infrastructure strategy.
- Data curation requirements within an e-research cycle need to be addressed (including data creation/capture/collection, data management/storage/validation, data disclosure/publication/scholarly communication/citation, adding value/data linking/annotation/visualisation/simulation, knowledge extraction/data mining/ modelling/analysis/synthesis)
- Quality of the data, quality of access and data usage and quality of preservation are key aspects
- When addressing curation issues, added value to stored data should be generated (enrichment virtuous cycle - maintaining and adding value to a trusted body of digital information for current and future use)
- DANS and PIC are relevant examples – because the challenges are global, European solutions should be found

### Scholarly communication

- The definition of a strategy for scholarly communications needs to be based upon evidence – evaluating the behaviour of millions of users in accessing digital libraries provides excellent insight for future actions
- It is important not to create new infrastructures based upon old paradigms - the digital consumers should be the driving force
- The example of DARE shows that it is possible to deploy digital academic repositories in practice through an innovative approach, with the support of researchers
- Funds are important – but a political push through e.g. a “European Charter on open access to publicly funded knowledge” stating the principle and defining the responsibility of universities and research institutes (complementing the charters of Bologna and Lisbon on higher education) would be very relevant
- Surveys on conditions/liberators of researchers effectiveness show the importance of organisational and human resource issues (bureaucracy, facilities to recruit and retain support staff).

### Generic

- The advent of powerful e-Infrastructures (infrastructures based on high performance networking and grid) consolidated in FP6, opens new avenues that need to be further exploited in FP7
  - Provide Europe with world class high-end computing capability
  - Deploy pan-European reinforced AAA mechanisms
  - Deploy storage facilities able to generate value for researchers
  - Support the deployment of digital archives (digital repositories of scientific information could be frontrunners) and facilitate the open access to scientific results
- A new set of communities (humanities, social sciences) are showing a greater interest for e-Infrastructures raising more sophisticated requirements (namely related to data)
- There are significant national and organisation initiatives guaranteeing the success of these endeavours at European level
- Best results will be achieved by addressing in a coherent way the deployment of new infrastructures (e.g. HPC) and the incremental and continuous nursing of “virtualised” infrastructures.