



Enabling Grids for E-science

# Data management

**H. Schwichtenberg – FhG/SCAI**

**[horst.schwichtenberg@scai.fraunhofer.de](mailto:horst.schwichtenberg@scai.fraunhofer.de)**

[www.eu-egee.org](http://www.eu-egee.org)



Information Society  
and Media



- Access to distributed remote data has to be provided with IT-solution/Infrastructure to open the possibility to run complex experiments (coupled models e.g. Climate) and to share data in a large community
- Experiences of recent years show that available capabilities to access to data in Grid Infrastructures/middleware have to be improved



- Survey on requirements by DEGREE SSA:  
<http://eu-degree.eu>  
For Earth Science Applications we need:
- Support for complex workflows, robust & fast replication of data is indispensable
- Support for Metadata intensive applications in distributed environments
- Webservice (WS-\* Standards) based interfaces/tools (ex. with Open GIS services)
- Data access and management solutions with commercial OS's (e.g. Microsoft Windows)
- Fast transfer of large files and of a large number of different files
- Different and extended Data policies have to be processed
- The middleware has to be extendable with semantic technologies for data
- Access to “ remote data “ – outside of grid infrastructures --

## *Data Services – Beyond Storage to Use*

### *What services do users want?*

How do I make sure that my data will be there when I want it?

How can I combine my data with my colleague's data?

How should I organize my data?

How should I display my data?

What are the trends and what is the noise in my data?

My data is confidential; how do I make sure that it is seen/used only by the right people?

How can I make my data accessible to my collaborators?



Enabling Grids for E-scienceE

## Progress since EGEE'06

[www.eu-egee.org](http://www.eu-egee.org)



Information Society  
and Media

