

Survey on Existing Science Gateways

Based on DEGREE

Horst Schwichtenberg , Steffen Claus



„A Science Gateway (SG) 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 customized to meet the needs of the targeted community“ (citation from TeraGrid website).

**Portal
fat client**

Characteristics

- SGs should **ease the access** to the Grid (resources such as data, information and computing power)
- SGs are usually restricted to a **specific user community**
- SGs serve as **single point of entry** to community specific services

- SGs may be:
 - Webportal
 - Fat client on a desktop

 - Supports different Grid Infrastructures (bridges)

DEGREE “Surveys” available:

D4.1

- **Earth Science (ES) Portals Survey**

considers:

- ES uptake of Grid technologies
- methods, techniques and grid middleware interfaces
- ES requirements regarding portals

D4.2

- **Generic Portals Survey**

analyses:

- not only ES specific portals
- their approach to inclusion of Grid technology
- ES Portals requirements vs. existing technology

Published on eu-degree.eu

ES covers a wide range of disciplines and activities

No single portal model covers all of the needs

Possible classification:

- → **Data Dissemination oriented**
 - Geonetwork
- → **Collaboration oriented**
 - SSE (Service Support Environment)
- → **Grid-based**
 - Grid-ify (Grid on demand)

• GEONETWORK Portal [www](http://www.geonetwork.org)

- optimized to support spatial data
- allows sharing of geo-referenced thematic datasets in wide community of spatial data users
- enable access to geo-referenced databases, cartographic products and related metadata from a variety of sources
- standard
 - Implements and extends ISO 19115 Geographic Metadata functionalities and OGC services
 - unifying approach is offered to the community, free and opensource
- de-centralized
 - nodes installed in individual organizations
 - single entry point
 - distributed search & data access
- users can:
 - locate and access the data for creating new maps combining various layers of information
 - [processing is done off-line]
 - publish the new maps
- types of users
 - Decision makers, development planners, humanitarian and emergency managers
 - Operational experts, multidisciplinary geographical spatial data analysts and forecasters
 - Researchers and value adders
- Others in this group: MERCATOR, WDC, SPIDR



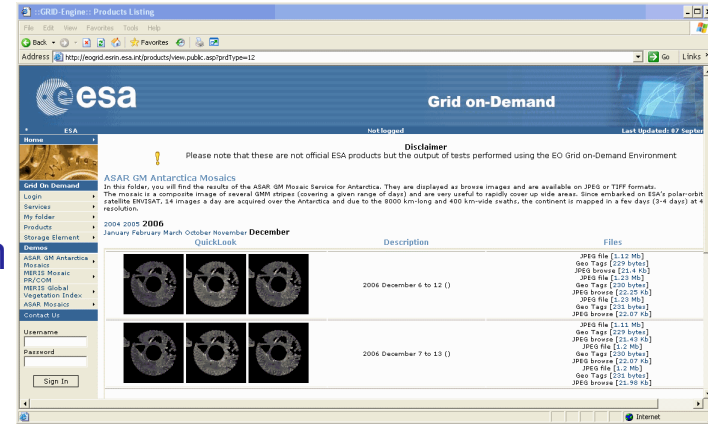
- **SSE (Service Support Environment)** [www](http://www.eoportal.org)
 - Common web-portal based framework for Earth Observation (EO)
 - Allows service providers to easily make their services available
 - Basic and complex ES community services
 - Services integrated directly in the Portal accessed via the Portal
 - New services can be composed using the SSE Workflow
 - **Available Services**
 - Data and Information provision, data conversion and processing, data delivery
 - Thematic mapping, land use, environmental monitoring, etc.
 - Product searches
 - Demonstrations & promotions of new EO environment monitoring services
 - Heterogeneous access to multi-mission satellite data
 - Other new services under development
- **Others in this group: AMI4FOR, UNEP, ETHER, TheVoice**



GRID-IFY (Grid-On-Demand)

(ESA: GPOD/Genesi-DR)

- Spatial data (EO) application and Grid integration Portal framework environment
- Integration of EO Catalogues for product search and retrieval, OGC WMS for displaying product overlays on top of world maps
- Security management
 - User registration, login and automated management of certificate using MyProxy
 - At time of registration the user is assigned privileges to access specific applications, processing algorithms, services and data
- Application porting
 - Implementation and configuration of an application decomposed into a set of processing modules/services using a basic Grid deployment framework model
 - Simplified access to Grid services exploiting the state of the art Grid standards and technology (gLite, GT4)
- Grid-based application deployment
 - Combined scheduling of data and jobs, execution using Grid resources
- Desktop as well as web user interfaces
- Others in this group: IMPECT, VGISC, WEBGRECL, DATACROSSING, IDEAS, KWF-GRID, MEDIGRID



- **Key Characteristics:**
 - Spatial Data Access and Processing
 - OpenGIS Services
 - OPeNDAP
 - Metadata access and search
 - Generic Java Web Services using
 - Tomcat
 - Axis
 - Gridtechnologies
 - GT4
 - Gridftp
 - glite

only few with Ontologies

	1. IMPECT	2. GRID-FY	3. GeoNetwork	4. AMI4FOR	5. VGIS	6. WebGReIC	7. Datacrossing	8. UNEP	9. IDEAS	10. Mercator	11. World Data Cntr	12. K-WF Grid	13. MedIgnid	14. SPIDR	15. ETHER	16. The Voice	17. SSE
Technologies Used																	
Grid																	
GTK	x	x				x			x					x		x	
WSRF				x	x				x			x	x	x			
EGEE, LCG, gLite			x						x					x			
Workflow																	
Resource Brokering / Metascheduled				x	x							x				x	x
GridFTP		x			x				x								x
Storage management, SRB, etc.		x			x												
MyProxy		x			x												
SRB						x											
OGSA-DAI					x				x						x		
Replica manager													x				
OMII									x					x			
e-Collaboration																	
Wiki								x						x			
Content Management System		x						x									
Custom-built content management				x	x	x		x		x						x	x
Webservices, SOA																	
Apache Tomcat			x		x	x	x	x						x	x	x	
Axis		x				x		x						x		x	
SOAP	x	x	x			x		x						x		x	x
XML over HTTP			x			x		x									
WS-I, UDDI																	
Ontology, Semantic Web																	
OWL					x							x					
Portal																	
Gridsphere		x										x	x				
Portlet		x												x	x		
JSR168, WSRP		x										x	x				
Metadata																	
Metadata search					x	x			x		x			x		x	
Register new metadata	x	x	x														x
ISO 19115 Geographic information			x					x									
Spatial data, OGC, OpenGIS	x	x	x				x	x	x					x			x
Metadata sharing			x														
Catalogue Services			x		x												x
Lucene indexing engine			x								x						
Data access																	
FTP										x	x			x			x
DODS / OPENDAP / LAS										x							
NETCDF									x	x				x			
FGDC									x					x			
Z39.50																	

- **Generic requirements**
 - Interoperability between different Grid MW & infrastructure
 - Reliability & QoS
 - Standard "off-the-shelf" tools for integrated Grid Security and User management
 - Dynamic content authoring, addition of customized services, registration of available resources
 - User support, how-to, tutorials
- **ES specific requirements**
 - Strong emphasis on Metadata and Data, its Discovery and Access
 - Working with very large datasets and number of files
 - Integration of heterogeneous distributed services (Grid & Geo-services, OGC)
 - Support "Gridification" in Geo-services and Spatial Data standards
 - Tools & interfaces readily useable by ES Scientist
 - as application assembler as well as end user
 - automated tools to assist deployment ES applications and libraries on the grid (workflows)
 - Facilitate integration with ES web services
 - Interoperability with ES data catalogues
 - Support for Earth Science sensors and thematic data

ES specific Requirements (towards Portals)

- Data Dissemination / management / interoperability
 - Discover, identify and work with ES data
 - OGC-compliant Web Services integration
- Collaboration
 - For ES virtual communities
 - Sharing of knowledge, data, results
- Grid-based Services
 - ES data intensive processing
 - Orchestration & coordination of tools & services of different infrastructures

Technology	EnginFr ame	Clarens	GEON	Sakai	GRB	GEODE	Media Grid	P-Grade	GeneFrid	Telescien ce	LEAD	UserSmar ts GX
Grid												
GTK	x		x		x	x	x	x				
WSRF				x		x	x				x	
GridFTP					x		x	x		x		
RDBMS	x	x		x		x						
myProxy	x	x			x	x		x		x		
Workflow, WFDL					x			x	x		x	
OGSA			x					x	x			
EGEE, LCG, glite	x		x			x		x				
e-Collaboration												
Wiki				x								
RSS	x			x								
Webservices, SOA												
SOA	x		x	x			x				x	x
SOAP	x	x	x	x	x	x	x		x		x	x
WSDL		x										
Apache tomcat	x	x									x	
Portal												
GridSphere			x			x		x	x	x	x	
JSR-168, portlet			x	x		x		x	x	x	x	
WSRP			x	x								
	x	x										x
Java Script	x	x										
State machine workflow engine				x		x						
Metadata												
Metadata catalog access		x	x			x						x
OGC, GIS			x									x
Google maps, GWT	x		x								x	

- Area:
 - Weather/Atmosphere
- Standards/Interoperability:
 - On top of Teragrid, WS-*
- Services:
 - Experiment builder (workflow)
 - Personal workspaces
 - Data search, analysis and visualization

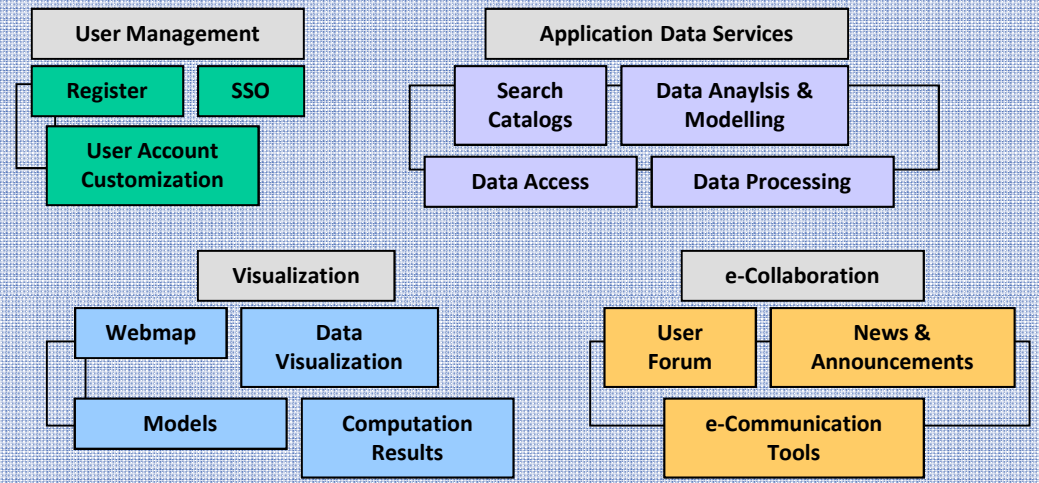
- Area:
 - Mechanics, (Compute- and data-intensive applications)
- Standards/Interoperability:
 - Globus, gLite, SGE, OpenPBS, LSF
- Services:
 - Presentation (?)
 - Data/Metadata
- (entfernen?)

- Generalized component-based framework

- Requirements & design objectives serving two different domains
- End-user: ease-of-use
- Application-developer: ease-of-assembly

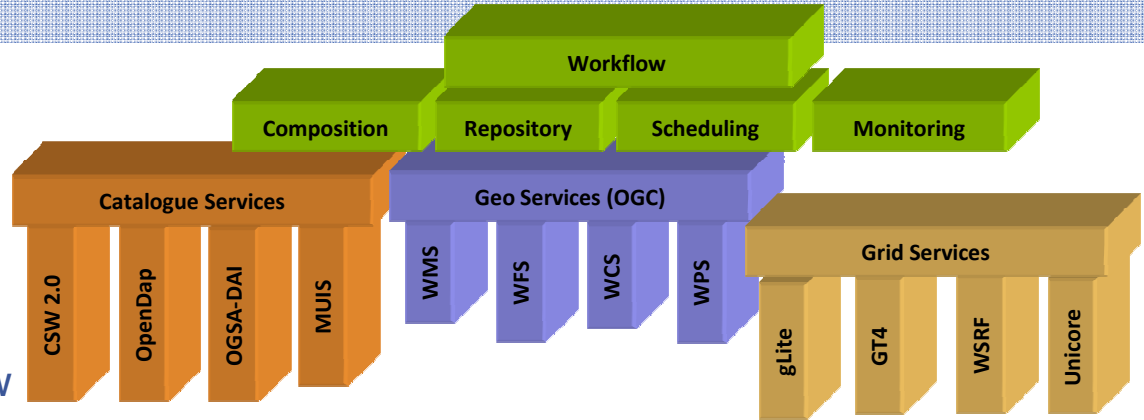
- Front-end: User interface

- Domain of the end-user
- Reusable services stored in workflow repository
- Users can invoke available workflows and compose new ones
- May be browser, fat client, desktop applications...

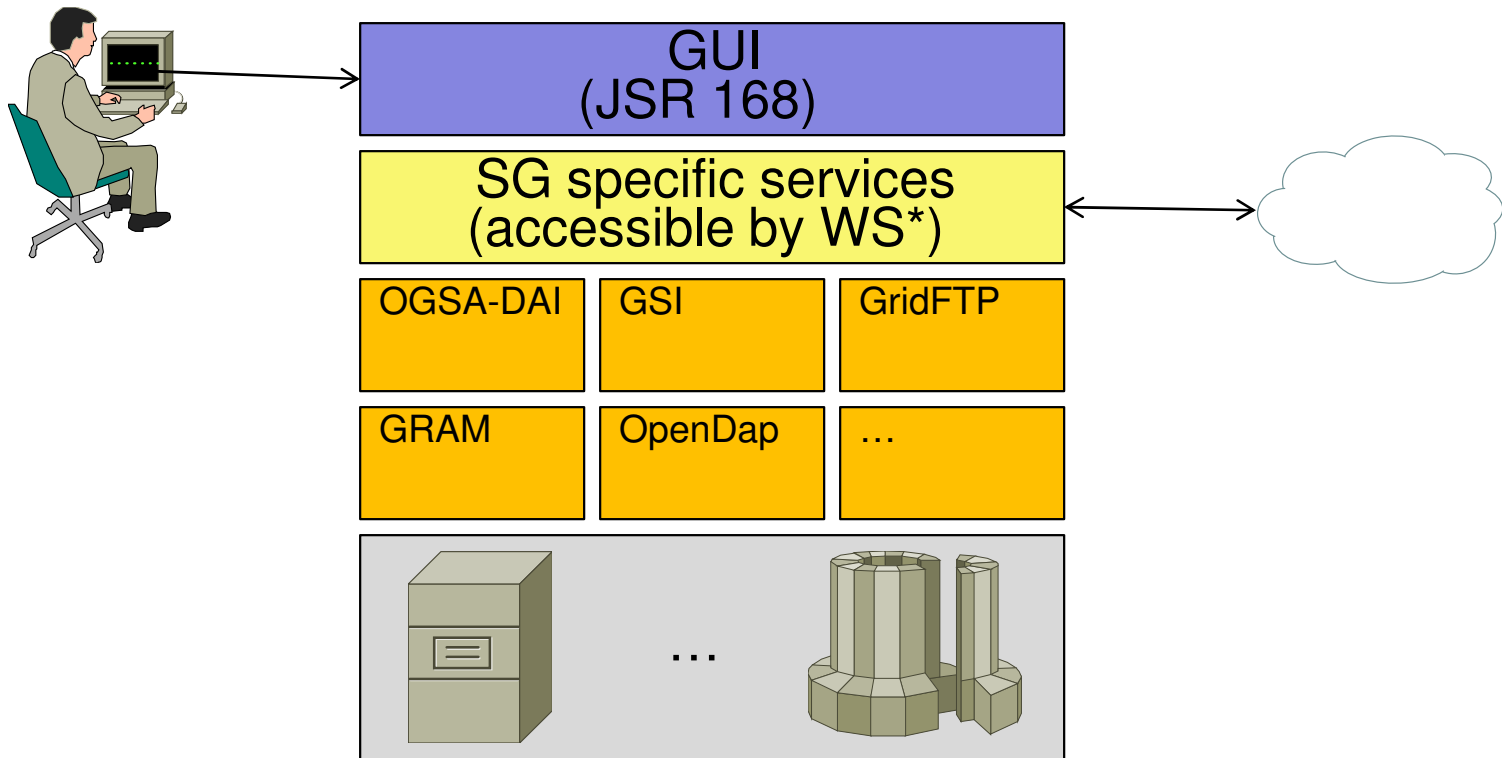


- Back-end: Services interface

- Domain of the application-developer
- Assemble new service components
- Publish services in workflow repository



- Science Gateways provide (mostly) a **user-centric, graphical** web based (JAR-168, GridSphere) access to an (rather) **application-centric** set of tools on top of the Grid (Globus, gLite, ...).



- Degree Survey
 - [http://www.eu-degree.eu/DEGREE/internal-section/wp4/DEGREE-WP4-D4.1 ES Portals Survey v1_0.pdf/view](http://www.eu-degree.eu/DEGREE/internal-section/wp4/DEGREE-WP4-D4.1%20ES%20Portals%20Survey%20v1%200.pdf/view)
 - [http://www.eu-degree.eu/DEGREE/internal-section/wp4/DEGREE-D4.2 v1_0.pdf/view](http://www.eu-degree.eu/DEGREE/internal-section/wp4/DEGREE-D4.2%20v1%200.pdf/view)
- Contact
 - Horst.schwichtenberg@scai.fraunhofer.de

- **Data dissemination**
 - Registration and publication of new sources of data
 - Search, locate and discover details of registered data collections
 - Access to data
- **Collaborative**
 - Structured, customized organization of the portal pages according to dedicated application themes, activities and functions
 - Facilitate customizations of the portal information and content by the realtime integration of contributions from individual users
 - User identity management, access permissions control, account settings and customization of the individual user's environment
 - Customized domain-specific tools for e-Collaboration
- **Grid-based**
 - Front-end user interface for largescale dynamic processing ES specialist datasets
 - Orchestration & coordination of low-level tools & services
 - Ability to interface to different infrastructures
 - Generic framework model to facilitate addition and easy “gridification” of new ES applications, independent of middleware implementation specifics
 - Provide ready access to large Grid-based ES data collections and to support the easy integration of new data
 - Collections for use in the Grid-based ES applications data processing.