



Diligent

A **D**igital **L**ibrary **I**nfrastructure
on **G**rid **E**Nabled **T**echnology

Donatella Castelli
ISTI-CNR
Pisa (Italy)

1. DILIGENT overview
2. Planned evaluation framework
3. Evaluation needs

- Duration: 3 years
- Commencement Date: September 2004
- Effort: 1024 p/m
- Cost: 9.8 M Euro
- EU funding: 6.3 M Euro

- Consiglio Nazionale delle Ricerche – ISTI (Italy, Scientific Co-ordinator)
- European Research Consortium for Informatics and Mathematics (France, Administrative Co-ordinator)

- University of Athens (Greece)
- Swiss Federal Institute of Technology Zurich -ETH Zurich (Switzerland)
- Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. – IPSI (Germany)
- University for Health Informatics and Technology Tyrol (Austria)
- University of Strathclyde (United Kingdom)

- Engineering Ingegneria Informatica SpA (Italy)
- Fast Search & Transfer ASA (Norway)
- 4D SOFT Software Development Ltd. (Hungary)

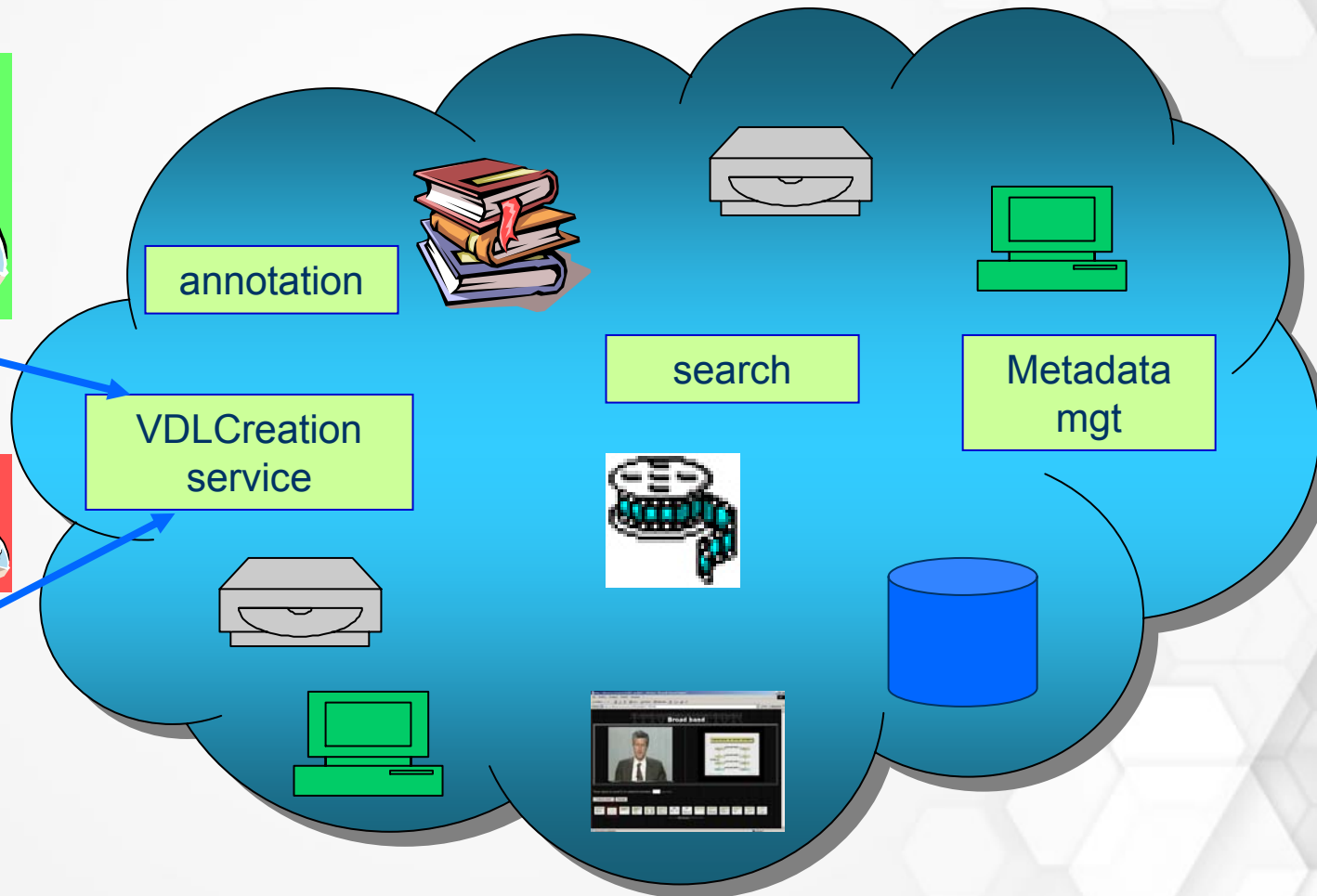
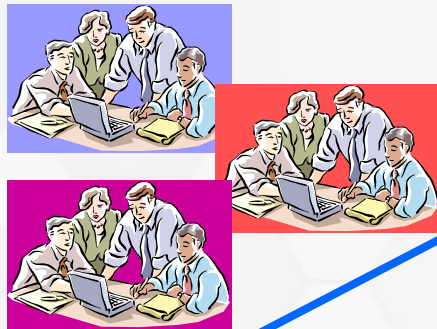
- European Organization for Nuclear Research (Switzerland)

- European Space Agency – ESRIN (Italy)
- Scuola Normale Superiore (Italy)
- RAI Radio Televisione Italiana (Italy)

Objective

Create an advanced Grid-based **Digital Library Infrastructure** that will allow members of dynamic virtual organizations to create on-demand transient virtual digital libraries based on shared computational, storage, multimedia, multi-type content and application resources

DILIGENT Framework



Today the demand for DLs is high

- DLs are perceived as a necessary instrument to support communication and collaboration among the members of communities of interest
- Many application domains require DL services, e.g. e-Health, e-Learning, e-Government, for managing and accessing their not-standard information objects
- The organizations that demand a DL are often small, distributed, and dynamic; they want to use DLs to support temporary activities such as courses, exhibitions, projects, etc.

The current DL development model

- The construction and management of a DL requires high investments and specialized personnel
- Years are spent in designing and setting up a DL
- Multimedia handling is very limited since it requires high computational resources
- The systems lack interoperability and the services provided are difficult to reuse

Current trends

- **Sharing of Content and Service Resources**
 - + Customised views of the underlying shared resources
- **Digital Library Management Systems**

DLMSs implement all the key services and management features required to support the entire spectrum of DL functionality
- **DLs as components of more generic shared infrastructures**

These infrastructures serve a multitude of application areas

Architectural framework

The realisation of this new vision requires a suitable architectural framework

DILIGENT will build an architectural framework based on **Grid technologies**

Why Grid technologies?

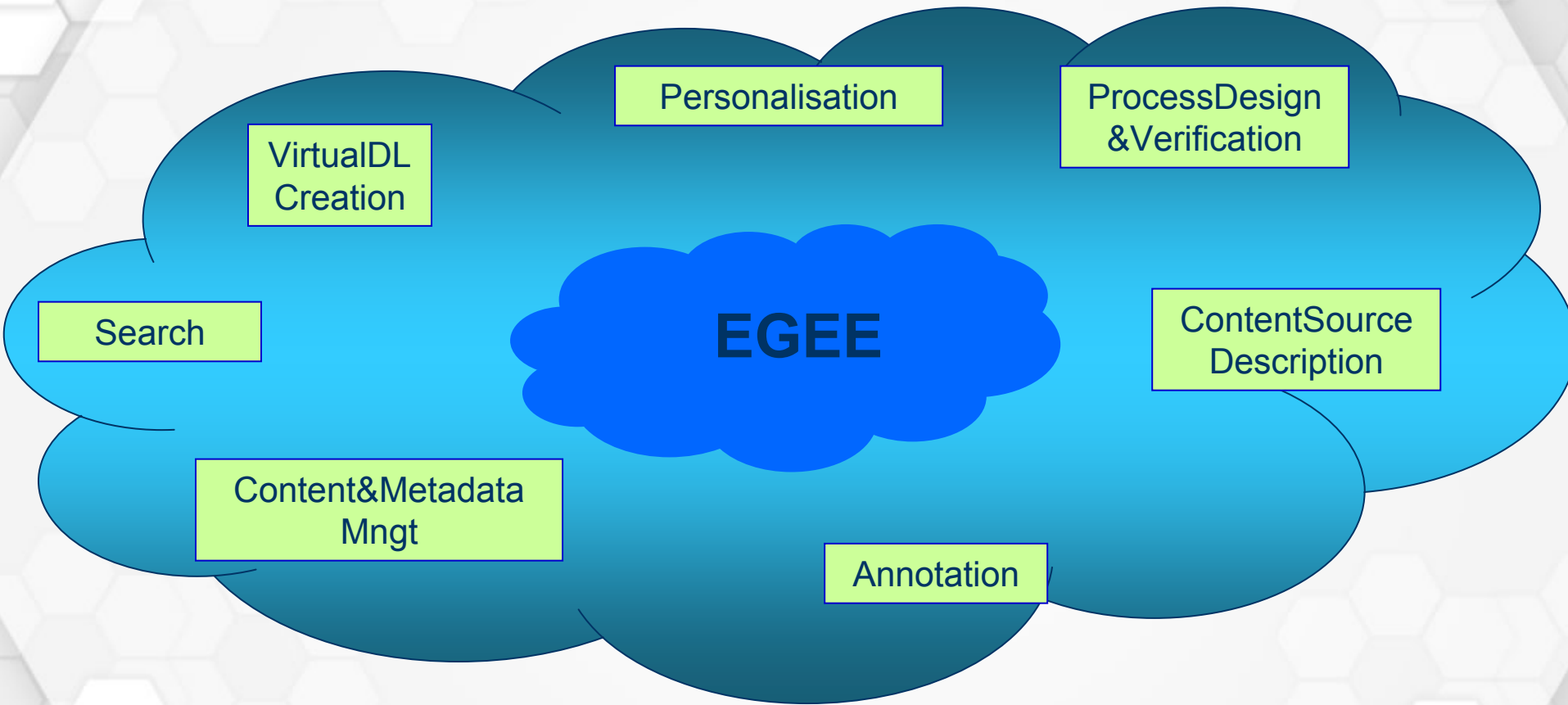
- Controlled sharing of resources
- High computing capabilities for the handling of a wide variety of information objects

The Grid infrastructure: EGEE

Enabling Grids for E-science in Europe - EGEE

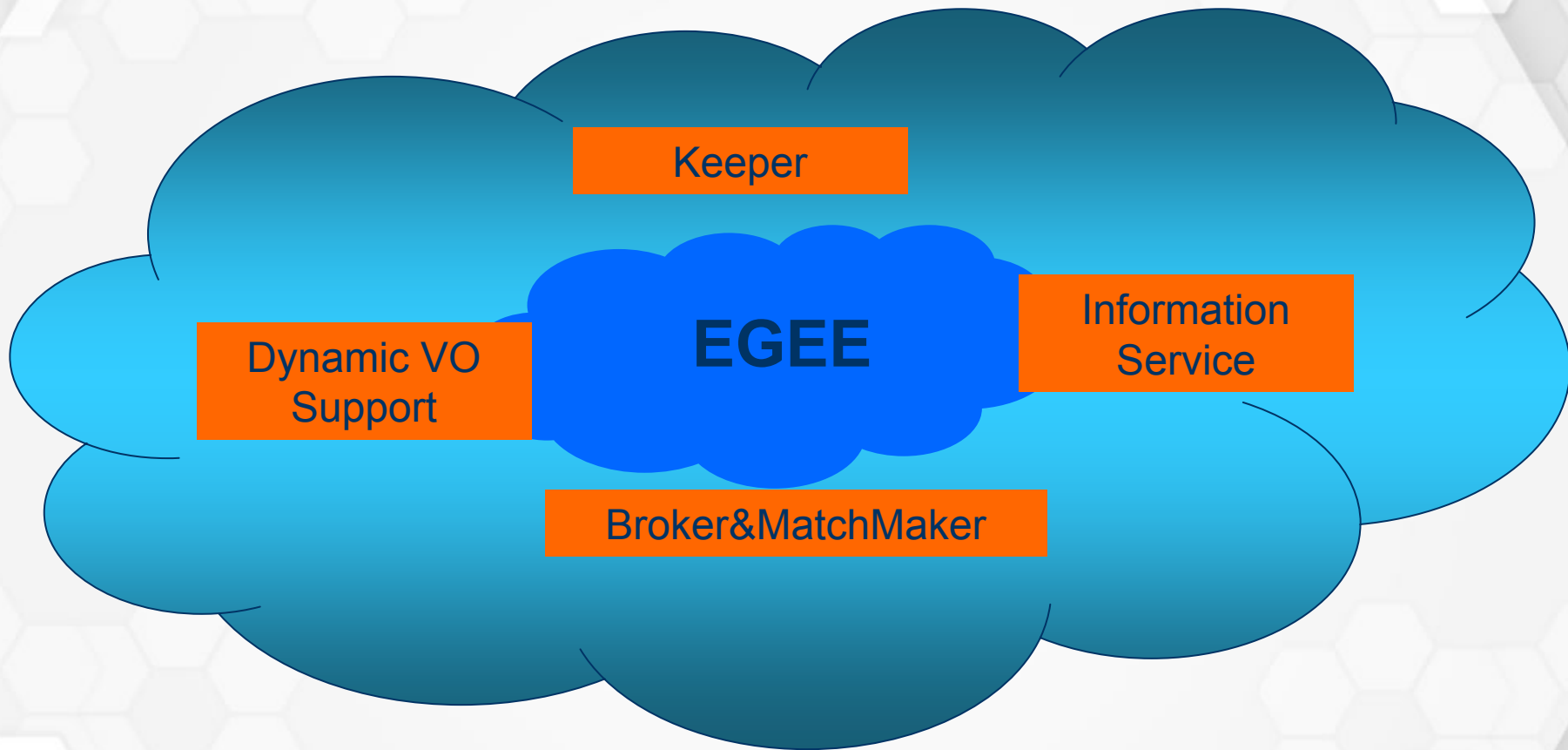
- Two years project funded by the European Commission with funding of over 30 Million Euros
- Part of a four-year programme, where the results of the first two years will provide the basis for assessing subsequent objectives and funding needs
- Experts from 70 organisations and 27 countries with the common aim of:
 - ◆ providing researchers in academia and industry with a 24 hours-a-day access to major computing resources, independent of their geographic location
 - ◆ developing a service grid infrastructure built on European recent advances in grid technology

DILIGENT technical solution



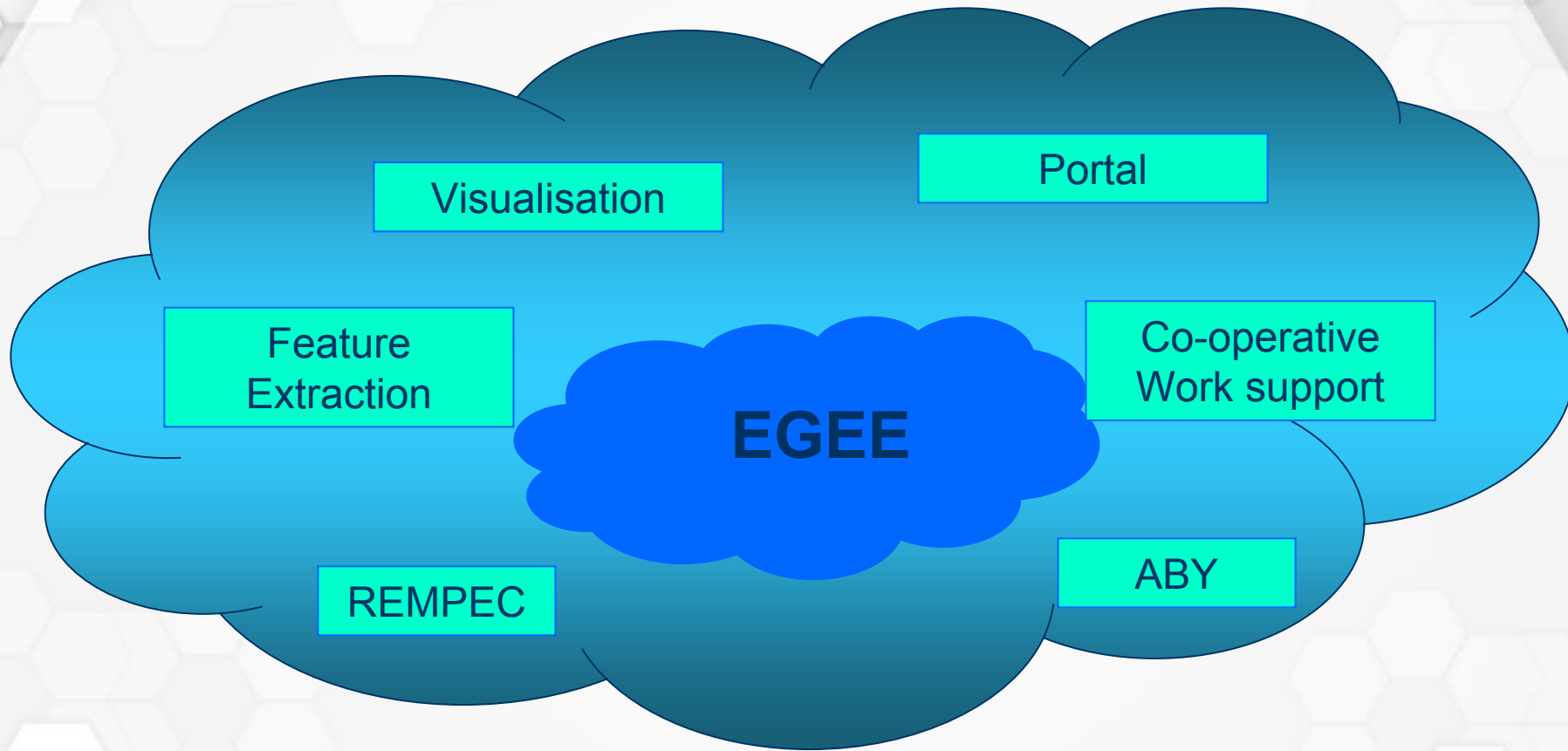
DILIGENT will be built by integrating Digital Library services on the infrastructure developed by EGEE

Technical Solution (cont.)



DILIGENT will enhance EGEE with the functionality needed to support the complex services interactions required to build, operate and maintain transient virtual digital libraries

Technical Solution (cont.)



Application services that provide user communities specific functionality will be added

- What has to be evaluated
 - ◆ Basic DL services
 - ◆ VDLs
 - ◆ Infrastructure
- Experimentation and evaluation in two different, complementary application domains
- The user communities involved in the evaluation will provide feedback since the early design phases and for the whole duration of the project

The ImpECt Scenario (1)

Implementation of Environmental Conventions

- Earth Science Domain
 - well-established tradition in exploitation of new technologies
 - wide variety of content types (maps, satellite images, etc.)
 - very large, dynamic data sets
- Community represented by:
 - **European Space Agency**
 - The Italian Ministry of Environment and selected European coast-guard offices
 - REMPEC the Regional Marine Pollution Emergency Response Centre (Malta)
 - UNESCO IOC (Intergovernmental Ocean Committee) (Paris)
 - ITOPF, International Tanker Owners Pollution Fed. Ltd. and MOIG, Mediterranean Oil Industry Group
 - ICRAM, Italian Central Institute for Marine Applied Research

- candidate information collections:
 - ESA ESRIN catalogues (<http://odisseo.esrin.esa.it>, <http://cat.envisat.esa.int>)
 - ESA EO Data products portal (<http://www.eoportal.org>)
 - Oil spill repository:
<http://intelligence.jrc.cec.eu.int/marine/oceanides/poster/screenshot.jpg>
 - European Maps: <http://eu-geoportal.jrc.it/>
 - The CEOS International Directory Network (<http://idn.ceos.org>)
 - International Charter for Disaster Management (Space Agencies)

- use cases for DILIGENT VDLs:
 - preparation of periodical reports
 - creation of decision supporting mechanisms in case of environmental accidents

The ARTE Scenario (1)

- Cultural Heritage Domain
 - ◆ IT technology exploitation still in infancy
 - ◆ multidisciplinary collaborative research
 - ◆ image based retrieval/semantic analysis of images

- Community represented by:
 - ◆ **Scuola Normale Superiore**
 - ◆ **Rai Radiotelevisione Italiana**
 - ◆ Brown University – Department of Italian studies
 - ◆ Centre de Recherche en Histoire des Sciences et des Techniques
 - ◆ Universidade da Coruña – Research Team on Hispanic Emblematic Literature
 - ◆ University of Glasgow – HATII
 - ◆ Università di Pisa – Facoltà di Lettere e Filosofia – Corso di Laurea Cinema Musica e Teatro
 - ◆ Studio Azzurro Produzioni

The ARTE Scenario (2)

- candidate information collections:
 - ◆ ABY - An Atlas of Memory Images, <http://www.ctl.sns.it>
 - ◆ An archive of the images contained in treatises on the art of memory of the XVI Century and memory images employed in different sectors of medieval and Renaissance culture (the visual arts, literature, mystical and devotional practices, etc.).
 - ◆ The Dream in Renaissance Literature and Visual Arts, http://www.cribecu.sns.it/analisi_testuale/settore_informatico/progetti/sgno/_en_index.html
 - ◆ An archive of literary texts and visual images that narrate or describe dreams
 - ◆ The Italian National Broadcasting RAI Educational, <http://www.educational.rai.it>
- Use cases for DILIGENT VDLs
 - ◆ Support multidisciplinary research revolving around images
 - ◆ organization of courses

New framework evaluation parameters(1)

New DL framework: Is it good or is it bad?
Under which conditions?

- Creation/Maintenance of Digital Libraries
 - ◆ DL development and maintenance costs
 - ◆ Technical complexity
 - ◆ DL set up delay
 - ◆ DL adaptability

- Sharing
 - ◆ Sharable resources
 - ◆ Complexity of the sharing process
 - ▶ From the provider point of view
 - ▶ From the infrastructure point of view
 - ◆ Level of resource sharing policies enforcement

New framework evaluation parameters(2)

- Quality of the DL infrastructure
 - ◆ Expandability
 - ◆ Availability
 - ◆ Robustness

- Integration capability/interoperability

VDL definition criteria

- VDLs definition criteria can refer to service/collection evaluation parameters
 - The best repository service for maintaining environmental reports (doc. and collection model, preservation, policy enforcement, ...)
 - The best indexing service for my collection of emblems images (feature extraction, performance, ...)
 - The best annotation service for my terabyte collection of satellite images (ann. model, ann. handling, performance, ...)
 - The best co-operative work service for the students of the “Italian Study” course (usability, functionality, cost, ...)
 - The best collection of emblems (metadata formats, completeness, largest number of annotations, ...)

Enhancing existing VDLs

- VDLs can be dynamically changed to provide a better service
 - The selected index does not work well with my B/W collection of images
 - The selected visualization service takes too long to show the structure of the data in a graphic form

- A DL conceptual evaluation framework is extremely important for a project like DILIGENT
- This framework must also take into account:
 - ◆ the system/infrastructure evaluation criteria
 - ◆ the creation, maintenance and usage costs