



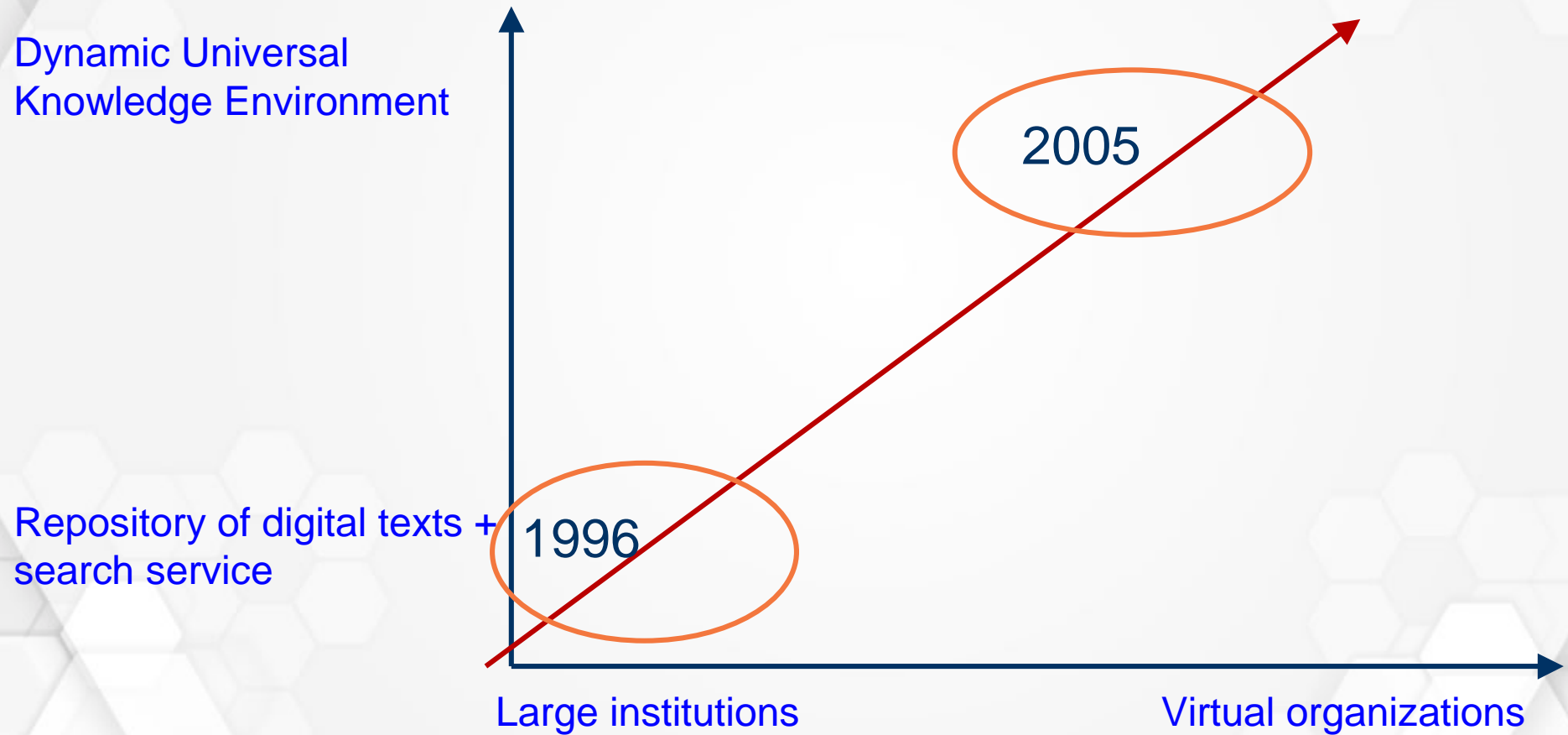
**Diligent**

**A Digital Library Infrastructure  
on Grid ENabled Technology**

# **Virtual digital libraries: The DILIGENT Project**

Donatella Castelli  
ISTI-CNR, Italy

# The DLs evolution



# New information objects

## Live documents

- a fixed text
- a pollution map
- a table summarizing data from millions of observed satellite measures
- a graph reporting an analytical trend of certain information extracted from a great amount of observed data

### International Report on Mediterranean Sea Chlorophyll Distribution during year 2003

#### 1. Scientific and Societal Concerns

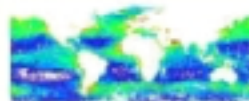
Any scheme to monitor the ocean biota and their environment must strive to address the major scientific and societal concerns of the day pertaining to marine life. This section summarises some major concerns that emerged during discussions at the meeting. Many other concerns could have been included, but space precludes a complete listing of concerns.

#### 1.1. Biodiversity and Conservation

Marine biodiversity is not easy to assess and is generally poorly known. There are many complicating factors, including a three-dimensional, fluid, mobile environment, its vastness, and its challenging depths. Away from shore, primary producers and primary grazers are usually small, drifting forms that undergo spatial variability and seasonal changes.

The larger invertebrate grazers have a range of life history stages, often with planktonic and benthic phases. Many large animals are migratory. Ocean habitats can be linked by the dispersal of planktonic larvae, and in this way, the systems can be interconnected even at a distance.

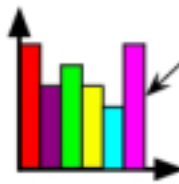
Finally, the higher-order diversity of life is much greater in the oceans than in terrestrial systems—there are 13 unique phyla in the oceans and only one on land. Marine biodiversity is essentially the evolutionary history of life. In general, long-term environmental stability seems to increase biodiversity and, conversely, global climate change can be expected to decrease it.



Jan - Apr 2003

	X1	X2	X3	X4	X5	X6	X7	X8	X9
Y1	12	13	15	26	11	34	45	45	54
Y2	32	12	46	67	21	22	44	12	44
Y3	23	33	56	77	32	44	12	55	33
Y4	44	34	12	55	34	45	12	22	44

Measures of yyy



Values of xxx

Automatically  
updated with  
the most recent  
data

# Dynamic Universal Knowledge Environments

- In order to satisfy this demand we need:
  - ◆ New DL systems able to provide innovative services, especially capable of supporting multimedia and multi-type information objects
  - ◆ Fast and unexpensive DLs development models, based on sharing and reuse of resources

Which technology?

## A solution

### • The Grid technology

- ◆ Large processing and storage capabilities for handling the wide variety of multimedia and multi-type information objects
- ◆ Controlled sharing of resources

**DILIGENT**

**A Digital Library Infrastructure on Grid Enabled Technology**

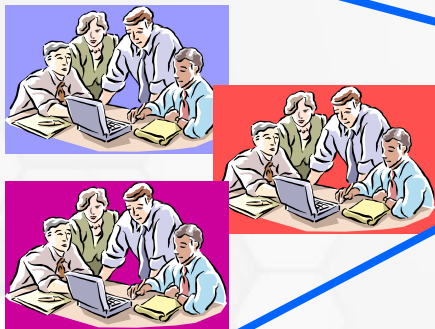
# Participants

- Italian National Research Council – 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)

## DILIGENT objective

Create a test-bed **Digital Library Infrastructure** that will allow members of dynamic virtual research organizations to create on-demand transient digital libraries based on shared computing, storage, multimedia, multi-type content and application resources

## Consumers



## DILIGENT DL infrastructure

Service A

Service B

Service C

DLCreation  
service

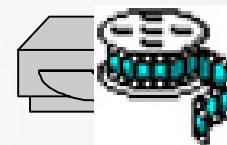
Service D

Service E

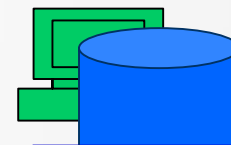


## Producers

3D processing



simulation



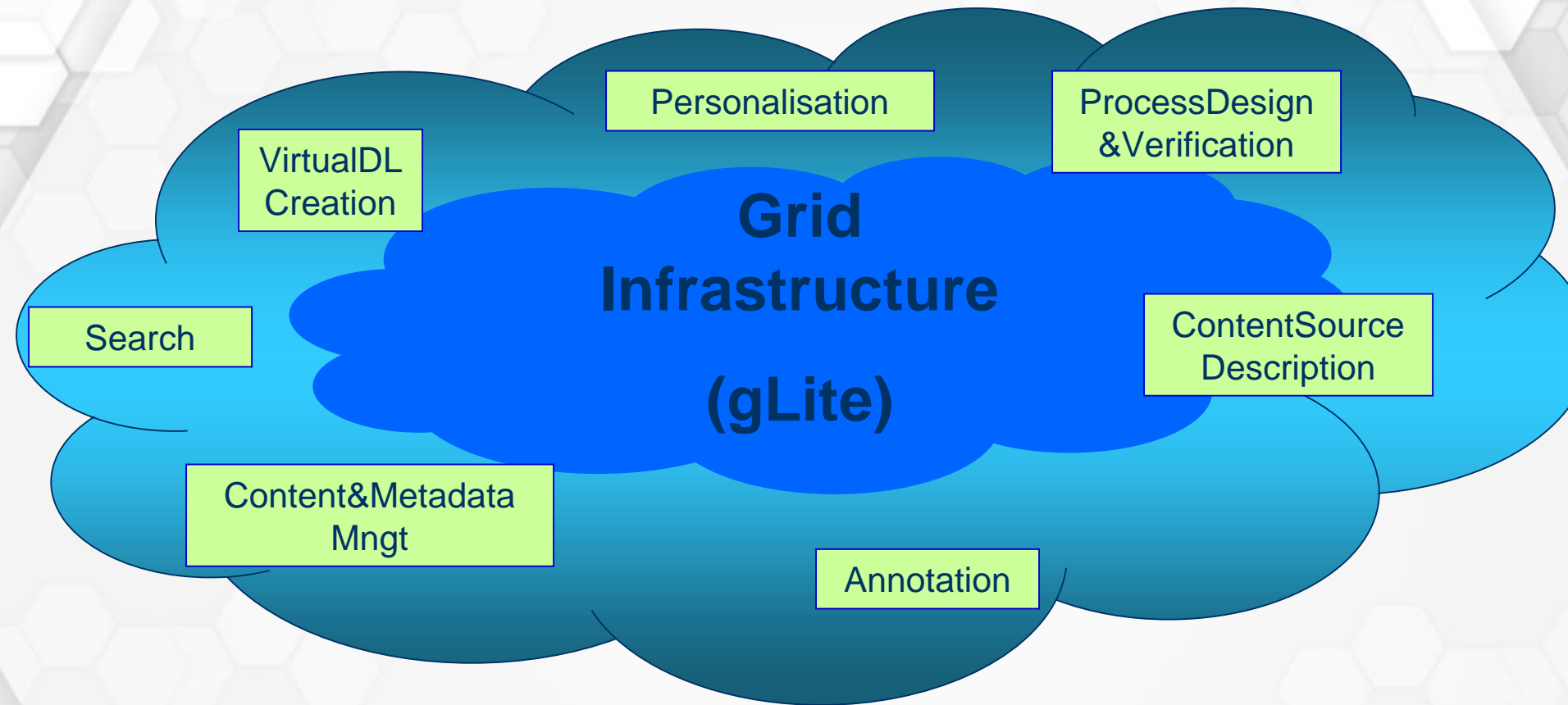
Feature  
extraction



Speech  
recognition



# DILIGENT DL Infrastructure



- The high computing and storage capabilities will be obtained by relying on the Grid infrastructure
- The DILIGENT project will develop the knowledge management services and the services that are needed to handle them

# DILIGENT functional view

## Application specific functionality

Portal Generator

Feature  
Extraction

Visualization

## Process Management

Process  
Execution &  
Reliability

Process  
Design &  
Verification

Query Process  
Optimization

## DL Creation&Management

Keeper

Broker&MatchMaker

Information  
Service

DL Generator

Dynamic VO  
Support

## Index&Search Management

Search

Index

DataFusion

Content Source  
Description & Selection

Personalization

## Content&Metadata Management

Annotation

Metadata Broker

Metadata  
Management

Content Security

Content  
Management

Content  
Wrapper

Storage  
Monitor

## Architectural View

- **DILIGENT will**
  - ◆ Adopt gLite as middleware
  - ◆ Exploit gLite as software
  - ◆ Join the EGEE infrastructure

## DILIGENT- gLite relationship (1)

- DILIGENT **adopts** gLite as Grid Middleware

The DILIGENT application will be composed by

- ◆ services provided by the DILIGENT project
- ◆ services provided by the gLite distribution

Both DILIGENT and gLite services will be deployed on the DILIGENT test-bed infrastructure

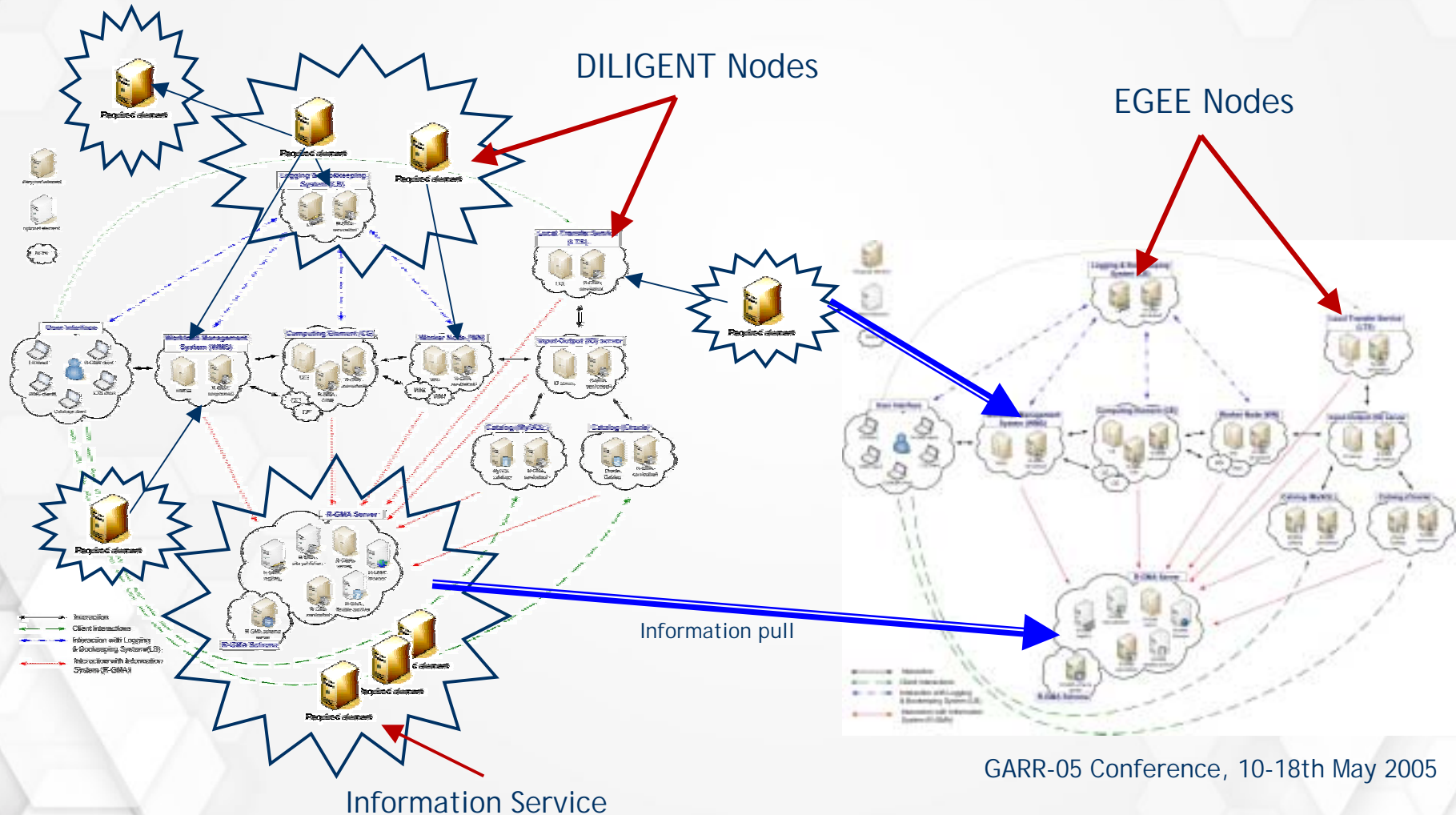
## DILIGENT-gLite relationship (2)

- DILIGENT **exploits** gLite services, components, or modules

DILIGENT services can be designed to include or wrap  
“pieces of gLite software”

## DILIGENT- gLite relationship (3)

The DILIGENT infrastructure **joins** the EGEE infrastructure



# Our users (1): Environmental researchers

## Implementation of Environmental Conventions

**Objective:** DLs for supporting the scientists in  
protecting the environment against pollution

**Focus:** the control of the marine environment of the  
Mediterranean Sea

**Activities:**

- ◆ producing reports
- ◆ prepare conferences
- ◆ analyze data about the environment in case of accidents

## Our users (1)

- Large variety of content types (e.g. maps, satellite images, reports)
- Large amount of data
- High processing required to produce useful outcomes
- **Participants**
  - ◆ **European Space Agency**
  - ◆ Ministero Italiano dell'Ambiente e altri uffici di Guardie Costiere europei
  - ◆ 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, Istituto Centrale per la Ricerca scientifica e tecnologica applicata al mare



## Our users (2): Culture heritage researchers

### ARTE Project

**Objective:** DLs for supporting the work of teams of researchers working on the humanities domain

**Focus:** to collaboratively investigate the usage of images and texts in ancient books and to establish semantic relationships among them

#### Activities:

- ◆ Organization of courses
- ◆ Exhibitions
- ◆ Conferences

- Many multidisciplinary archives
- Strong need of texts, images and videos semantic analysis and search across this heterogeneous documents
- Few resources
- Participants
  - ◆ **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

## Conclusions

- The eInfrastructure is the means that will allow to construct Dynamic Universal Knowledge Environments serving a large number of research communities
- The experience and the services developed by DILIGENT can be exploited by other knowledge-based applications
- The DILIGENT system can serve as a basis for many other applications: e-learning, e-health, e-government

# Contacts

[www.diligentproject.org](http://www.diligentproject.org)

- Donatella Castelli (CNR-ISTI, scientific co-ordinator)  
[donatella.castelli@isti.cnr.it](mailto:donatella.castelli@isti.cnr.it)
- Jessica Michael (ERCIM, administrative co-ordinator)  
[jessica.michel@ercim.org](mailto:jessica.michel@ercim.org)