

External TT Network Meeting
CERN – November 7th, 2006



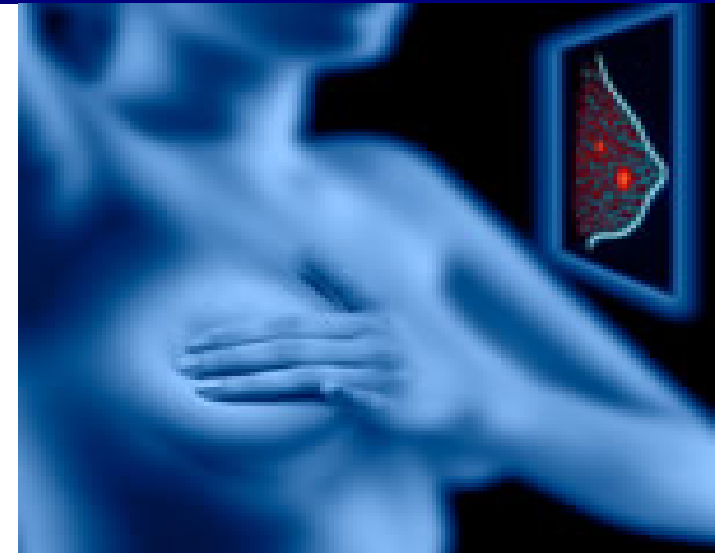
Information Society
Technologies

Mammogrid Project

J. Galvez



Overview



CERN Technology Transfer

What is Mammogrid?

- A project to build a: **pan-European distributed Database** of mammography images using GRID Technologies.
- Aim: To provide a **demonstrator** for use in epidemiological studies, quality control and validation of computer aided detection algorithms.

Time & Euros

- The project is funded by the European Commission (Framework Programme 5)
- Start: Sep. 2002 - Duration: 36 months – August 2005
- Cost: 2 M€



Mammogrid Consortium

- **CERN** (Project Management - Technical Coordination)
 - Vitamib (France) - subcontractor Finance/Admin
 - **Mirada Solutions (UK)** – now part of Siemens – Medical Image Analysis S/W
 - **University of Oxford (UK)** – Medical Vision Laboratory
 - **University of Pisa (I)** – Medical Physics section
 - **University of Sassari (I)** – Maths & Physics Dept
 - **University West of England (UK)** – Computing Research
 - **University of Cambridge (UK)** – Addenbrookes Hospital
 - **University Hospital of Udine (I)** – Institut of Diagnostic Imaging
-
- Ospedale Valdesse Torino (I) - Breast Screening Unit - subcontr.
 - Zybert Computing Ltd. (UK) - Subcontractor for GRIDserver

Thanks to Prof. Salvator Roberto Amendolia – who has motivated this initiative.

Why a Mammography Database?

- Breast cancer is a huge problem:
 - 1 in 8 women will develop breast cancer during lifetime,
 - 19% of cancer deaths are due to breast cancer,
 - 24% of all cancer cases are breast cancers,
 - there are 348,000 cases in EU & USA, 50,000 die every year

“A way to achieve the above is through **repositories** of mammography data for research and training that contain sufficiently large statistical samples”

The MammoGrid Challenge

Building this repository is not trivial because:

- Large numbers of exemplars are required to include sufficient numbers of abnormal cases to make significant statistical inferences.
- Data itself is large: 2 breasts x 2 views x 4K x 4K pix x 2 bytes = 128 Mbyte per patient per visit, 3M women per year UK, ~ 400 Terabytes in UK alone, (26M women in US), >100M mammograms on file .
- Acquisition is highly variable same image may look different depending on machine and parameters. How do you compare?
- Patient privacy and data security is key.
- Many relevant items of metadata: data need to be sorted efficiently according to demographics, acquisition information and image quantification parameters.

A GRID Infrastructure is ideal

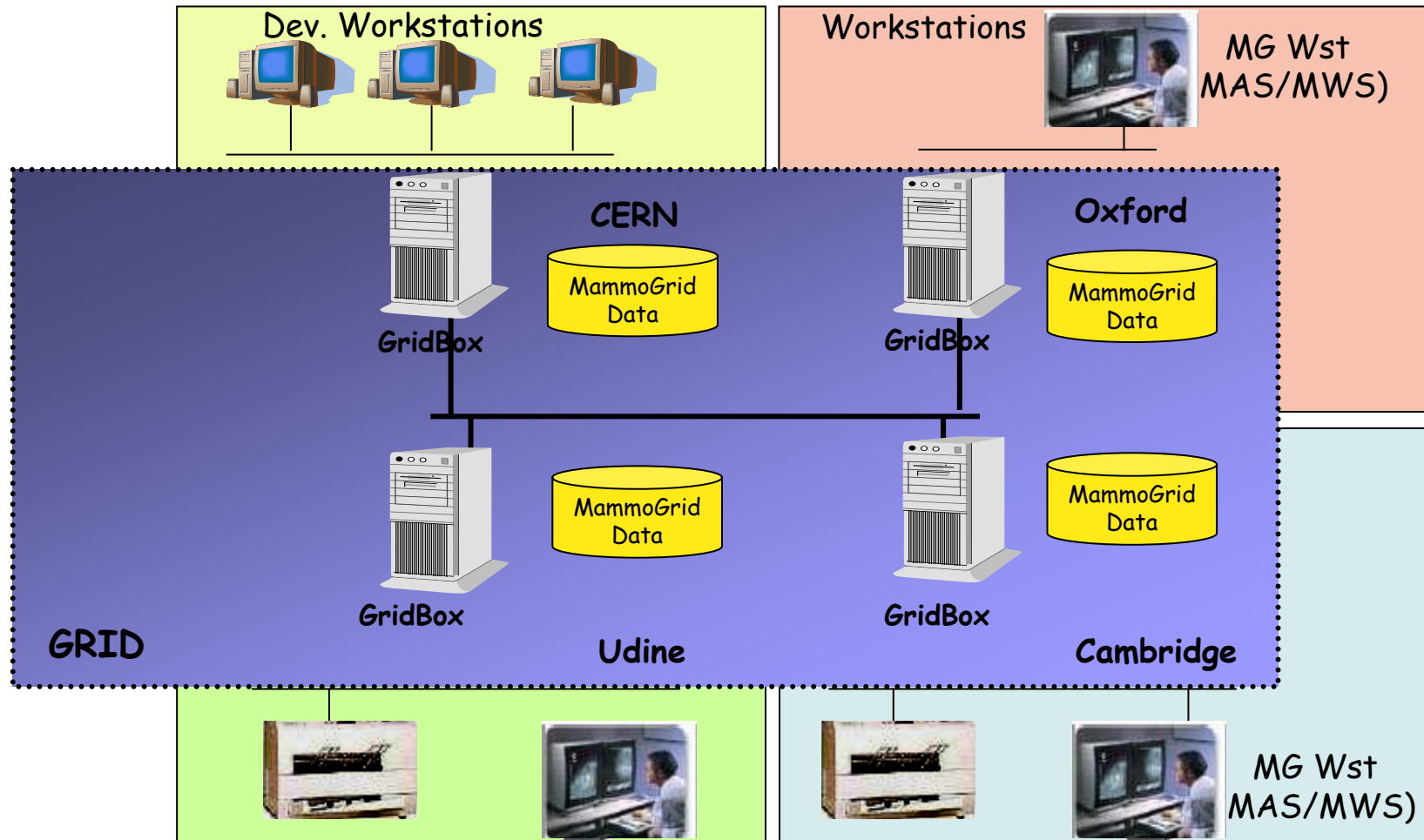
- The Databases to statistically validate image based clinical hypothesis are:
 - Populated by large number of cases
 - Contain large files (1 mammogram 10Mb+)
 - Geographically distributed repositories
 - Heterogeneous database formats
 - Need to be accessible to co-workers
- Development and validation of medical image analysis solutions demands:
 - Computationally expensive simulations.
 - Repeated runs for optimal parameter tuning.
 - Remote execution and maintenance
- Services (e.g. security) must be system-resident, invisible, generic

MG aims

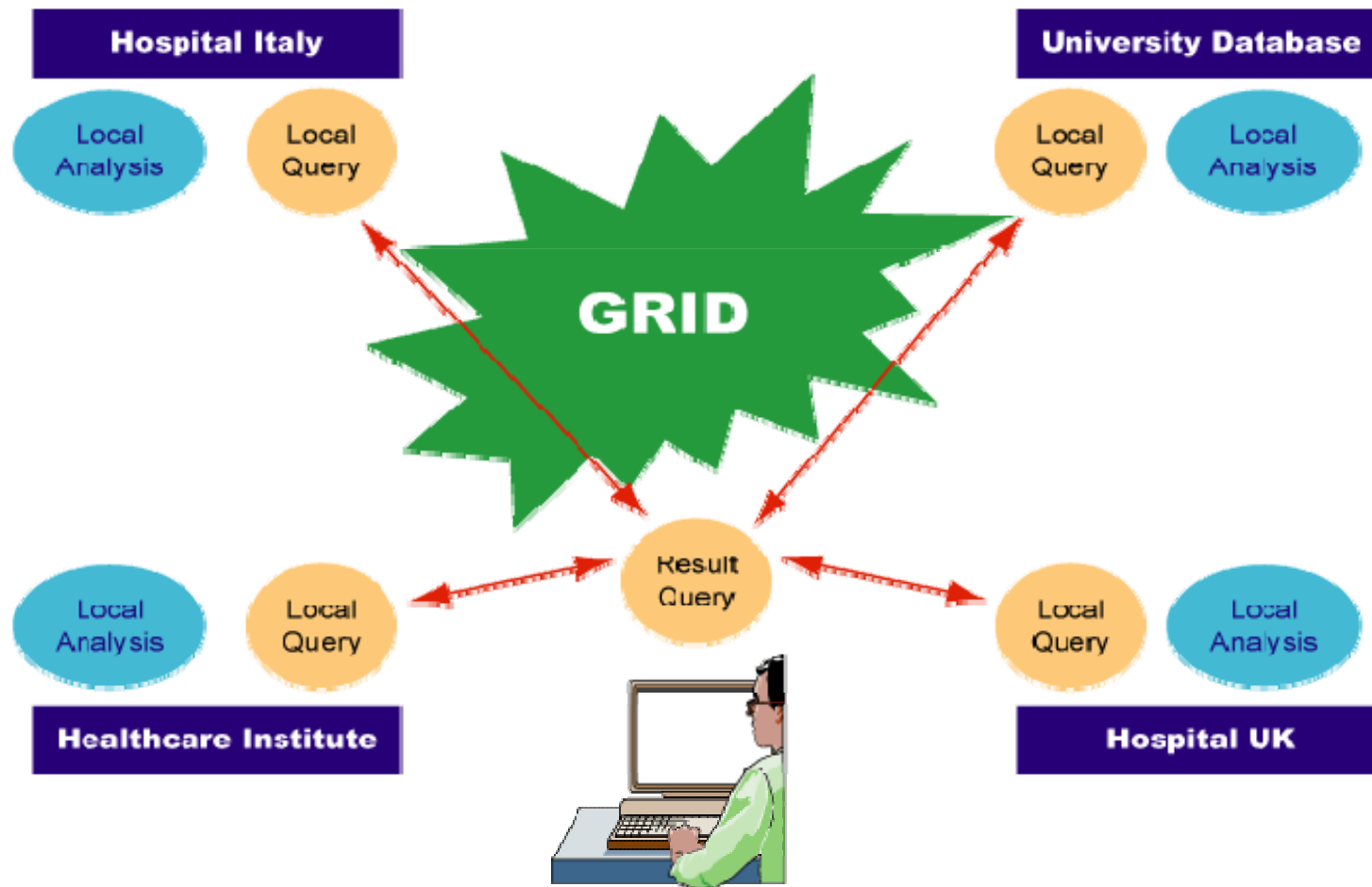
- In summary:

- investigate the feasibility of developing a **European database of mammograms**, accessed using emerging Grids software.
- a set of important **healthcare applications** using this database can be enabled so that **Grids** can be harnessed to support co-working between healthcare professionals across the EU.

System Architecture

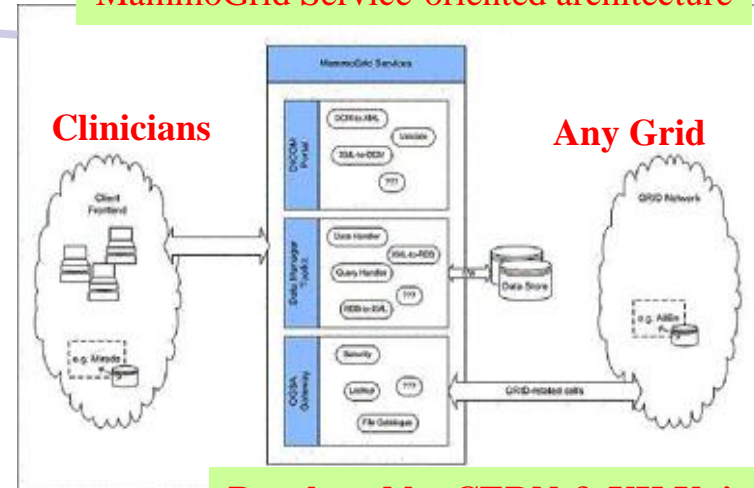
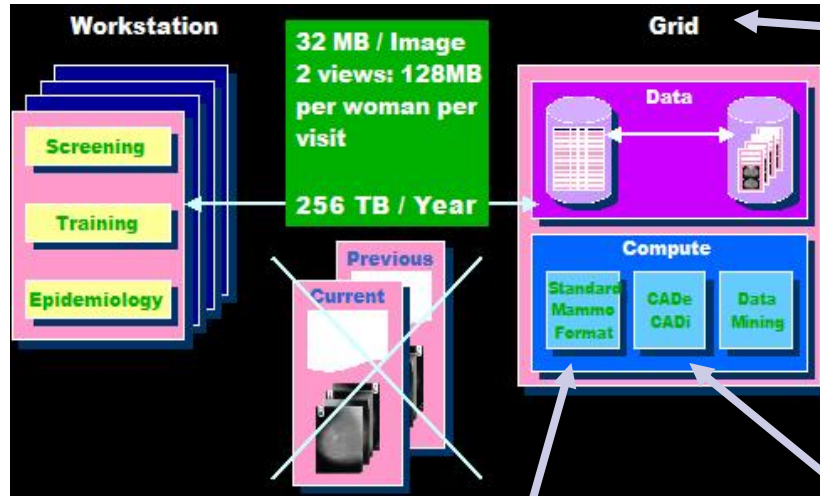


MammoGrid : A federated medical image analysis system

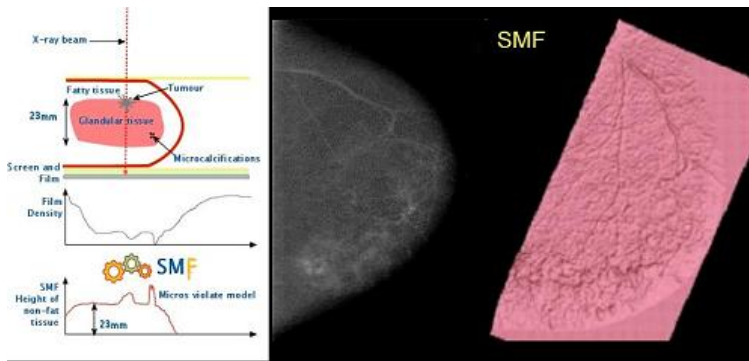


MammoGrid at a glance

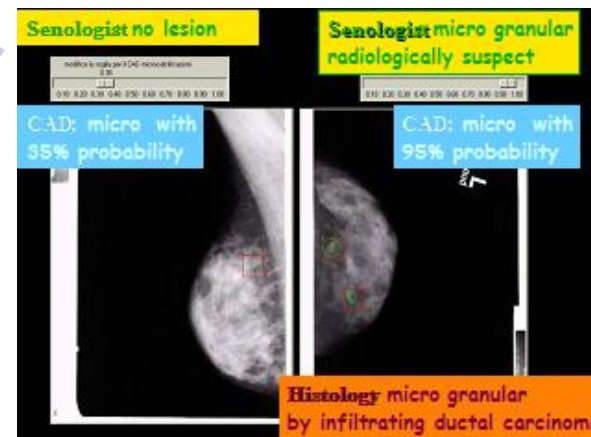
MammoGrid Service-oriented architecture



Developed by CERN & UK Univ



Developed by UK Univ & SME



Developed by Italian Univ

MammoGrid database

■ Cambridge

- 1423 patients
- 9716 images total (4815 - SMF)
- Annotation reports (DICOM SR) : 391
- Associated image data size : ~14 Mb
- File storage size : 260 Gb

■ Udine:

- 1479 patients
- 17285 images total (8634 - SMF)
- Annotation reports (DICOM SR) : 429
- Associated image data size: ~23.5 Mb
- File storage size : 220 Gb

- Since then >30000 images uploaded from acquired data (i.e. > 400Gbyte file storage)
- Batch upload (and batch SMF) is continuing together with live acquisition

Comments

- Effective and ongoing collaboration between Clinicians & researchers
- Great achievements in the Project
- Project was on-time and on-budget
- Deployed system is state-of-the-art, and fully working
- Challenging technical issues addressed
- A path to marketing the product is in sight
- A remarkable organizational and economical effort by the Consortium to cope with the difficulties

Transferring technology

- MammoGrid @ CERN is funded/supported by the [Technology Transfer Group](#)
- The success of the project has led to interest from outside companies and hospitals:
 - the health-care authorities of Extremadura, CIEMAT and [MAAT GKnowledge \(MAAT\)](#) propose to deploy a re-engineered version of the Mammogrid system (Mammogrid Plus) in collaboration with the [Udine and Cambridge Hospitals](#),
 - that will enhance the ability of doctors:
 - to verify test results,
 - to obtain a second opinion and to make use of the clinical experience acquired by the hospitals involved in the project.
- MAAT Knowledge is requesting a **non-exclusive License** of a subset of the Mammogrid results in view of commercial exploitation,
- A License Agreement was prepared and distributed to all members of the Consortium (TT Group, MAAT, Siemens)
- A Partnership Agreement is in preparation to setup the MGPlus Collaboration (TT Group + MAAT)

Conclusions

- EU is putting a large emphasis on Health Care Grid applications,
- MammoGrid is one of the IST FP5 projects specifically targeted at deploying Health Grids
- New biomedical data integration projects such as **Health-e-Child** are underway
- Need to set up mechanisms to achieve technology transfer from e-infrastructures and grid middleware development projects to practical life sciences and healthcare grids:
 - A very successful Technology Transfer of the MG Project to the Region of Extremadura in Spain

EU and Health Grids

- EU is putting a large emphasis on Health Care Grid applications,
- MammoGrid is one of the IST FP5 projects specifically targeted at deploying Health Grids,
- e-Health deals with using information and communication technologies to develop an intelligent environment that:
 - enables ubiquitous management of citizens' health status,
 - assists health professionals in coping with some major challenges,
 - Or integrates the advances in health knowledge into clinical practice.

High Energy Physics vs. MammoGrid

- MammoGrid heavily relies on technologies developed primarily in the field of High Energy Physics.
 - Similarities
 - Large number of big files
 - Files can be sensibly organized in directory tree
 - Need to replicate and move file copies between sites
 - Need to execute commands on the node which hosts data locally
 - Difficulties
 - Complexity of co-working in medical environment
 - Lack of trained IT personnel
 - Confidentiality !!!