Sharing and Processing Medical Images on the Grid

Ignacio Blanquer
Vicente Hernández
Valencia University of Technology
Institute for the Implementation of Advanced Information and Communication Technologies (ITACA)
Objectives

• To Present the Current Challenges of Medical Imaging and Propose the Grid Technologies as a Potential Source For Solutions.
• To Discuss How Different Projects are Facing Those Challenges.
• To Present a Development Implemented in the UPV for the Problem of Sharing and Processing Medical Images.
• Medical Imaging Concepts.
• Advantages of Medical Imaging and Current Barriers.
• Different Projects in the Area.
• The TRENCADIS System.
• Conclusions.
Medical Imaging Concepts

- Medical Imaging Deals with the Bitmap Representation of Anatomical, Morphological or Functional Information Relevant to the Management of a Patient’s Health.
- Medical Images can be Still or Dynamic, Deep Grayscale or Full Coloured, 2D, 3D or 4D and Typically are Very Large.
- Medical Image Processing Involves Tissue Identification (Segmentation), Projection (Rendering), Registration, Fusion, etc.

Medical Imaging Concepts · Advantages and Barriers · Current Status · TRENCADIS · Conclusions
• Medical Images are Generally Stored and Processed in DICOM (Digital Imaging and Communication in Medicine) Format.

• There are Many Different Modalities of Medical Imaging, Related with Different Physical Principles
  – X-Ray.
  – Computer Tomography Imaging.
  – Magnetic Resonance Imaging.
  – Positron Emission Tomography.
  – Ultrasound.

• Or Different Medical Procedures
  – Functional Imaging.
  – Spectrometry.
  – Angiography.
Advantages of Medical Imaging

• Medical Images Constitutes a Main Information Source for Diagnosis and Therapy.
• Medical Images are Used to Identify Trauma, Organ Malfunction, Tumours, Surgery Planning, etc.
• They are also Used for the Quantitative Evaluation of Masses, Flows, Injures,...
• Medical Images are Present in all Medical Disciplines.
Challenges in Medical Imaging

• Medical Images are Large and Thus Post-Processing is Computationally Intensive, Exceeding in Many Cases the Resources of Hospitals.

• Key Information in Medical Images can be Difficult to Observe, Even for Trained Specialists.

• Training is Mainly Based on Evidence.

• Privacy is a Key Issue Dealing with Patient Data, and Even More with Medical Images.

• The Data Produced Yearly in a Medium-Sized Hospital, is on the Order of Terabytes, So Organisation of the Data is Difficult.

• Data is Stored Distributed, but Consolidated Access is Difficult or Inexistent.
Projects Facing Those Problems

• **Biomedical Informatics Research Network (BIRN).**
  – Oriented to Neuroscience and Leaded by the University of San Diego.

• **CaBIG**
  – Oriented to Cancer Studies and Supported by the NIC of the USA.

• **MAMMOGRID**
  – European Project Oriented to Mammograms.

• **Information eXtraction from Images (IXI) - NeSC**
  – Oriented to Post-Processing and Supported by the National e-Science Program of the UK.

• **Medical Data Manager**
  – Developed by the CNRS and the CERN in the Frame of EGEE and AGIR Projects and Focused on Data Storage and Exchange.

• **TRENCADIS**
  – Developed by the GRyCAP of the Technical University of Valencia and Focused on Semantic Integration.
• BIRN is a Collaborative Environment for Sharing Data and Processing Tools in the Frame of Neuro-Sciences.

• 39 Research Groups of the National Institutes of Health are Participating in Four Areas: Mouse, Non-Human Primate, Brain Morphometry and Functional.

• The Project has an Strong Aim on Support and Reliability.
  – Data Organisation. BIRN Virtual Data Grid Based on SRB.
  – Security and Privacy. GSI-Like Authentication and Authorisation Mechanism.
  – Processing. Services Stored on Processing Sites.
• CaBIG is a Network Devoted to the Advance in the Study of Cancer.

• CaBIG has Different User Communities, One of them Dedicated to Medical Imaging.

• CaBIG Relies on the CaGRID Technology, Which Provides:
  – Data Organisation. It Uses OGSA-DAI to Federate Different Existing Resources.
  – Security and Privacy. It Implements a Centralised Authentication and Authorisation Mechanism Based on PERMIS and Two Own Components (GUMS and CAMS).
  – Processing. It Implements and API For Grid Services, and Dynamic Execution of User Code is Forecasted.
MAMMOGRID is Focused on Creating a Health Knowledge Infrastructure Oriented to Mammograms.

It Pretended Using Grids for Federating a European Distributed Database Sharing Data and Processing Services for Computer Assisted Diagnosis.

MAMMOGRID Technical Issues are:

- Data Organisation. Based on the Alien Catalogue System. (Centralised Catalogue on Distributed Data). Data is Normalised to Increase Homogeneity.
- Security and Permission. Data is Anonymised and Users are Managed with the GSI Interface.
- Processing. A Combination of Local and Remote Processing.
• **Information eXtraction from Images (IXI) - NeSC**
  – IXI is Oriented to the Development of Grid Services for Segmentation and Registration of Medical Imaging.
  – It was an Important Focus on Workflows, Being Developed a Specification Language: MICL (Medical Imaging Component Language).

• **IXI Provides**
  – Security and Privacy. Relies on GSI.
  – Processing. Based on GT GRAM. Grid Interfaces to Advanced Processing Services.
• MDM is a General-Purpose System for Storing and Sharing DICOM Data Using Grid Standard Protocols.

• MDM Faces the Medical Imaging Challenges Considering:
  – Data Organisation. Data is Pseudoanonymised and Made Available to the Grid Through SRM Interfaces and gLite 1.5 Catalogue System.
  – Privacy and Security. Data is Encrypted and Decrypted on the Storage Resources. Keys are Stored on Hydra Servers and Metadata in AMGA.
• **Towards a Grid Environment for Processing and Sharing DICOM Objects**
  – TRENCADIS Aims at the Development of a Middleware to Create Virtual Repositories of DICOM Images and Reports.
  – It Uses a Semantic Model for Organising the Data.
  – Data is Encrypted and Decrypted to Ensure Privacy Protection.
  – High-Performance Services are Included with the System.
  – Architecture Totally Based on WSRF.
• Objective: Creation of Virtual Shared Repositories of Medical Images.
  – Complementary to PACS.
  – Intended Mainly for Research and Training.
  – Multicentric and Multiuser.
  – Data to be Shared is Explicitly Selected.
  – Data is Pseudoanonymised Before Entering in the System.
TRENCADIS: Data Indexation

- Semantic Organisation
  - Users Organise Themselves on Virtual Communities.
  - From all the Images and Reports Available, Only Those Matching the Selection Criteria of the Virtual Community Profile are Accessible.
  - Further Filters are at the Experiment and the View Levels.
TRENCADIS: Data Access

- Data Is Stored Distributed.
- A Service Virtualises each Local Resource and Keeps Track of the Data Stored.
- The Global Index Only Registers the Sites that Are Relevant for Each Experiment, Making the Effective Query in Parallel on the Distributed Repositories.
- Local Indexes are Updated When New Data is Available.
TRENCADIS: Security

- **Authentication and Authorisation**
  - Users are Authenticated Through X.509 Certificates in an “Single Sign-on” Procedure (Using Proxies).
  - Roles of the Users (And Though the Virtual Community and the Access Permissions) are Managed Through VOMS proxies.

- **Privacy**
  - All Transactions are Based on Secure Protocols.
  - Data is Encrypted on the Grid Storage to Avoid The Access of Users with Privileges.
  - Keys are Split and Shared Through the VO Group.
Grid Processing Services.

- TRENCADIS Provides a Link To a WSRF Processing Service.
- Currently Volume Rendering and Co-Registration are Supported.
- Volume Rendering is Computing Intensive and Implemented Using MPI.
- Co-Registration is Very Time Consuming and is Implemented in a High-Throughput Model.
- Workflow Support and Interactive Scheduling are Under Study.
  - Potential Choices are MOTEUR and SDJ Support.
• The Convergence of Grid Technologies with Web-Based Technologies is Fostering the Take-off of Grid Production Services.

• Biomedical Community, and Specially Medical Imaging is Really a Community that Could Benefit Strongly from Grid Technologies.

• Although Challenges Exist, There are Many Successful Examples World-Wide (BIRN, MAMMOGRID, AGIR, TRENCADIS, GEMSS).