

# Life Sciences Research Community Workshop

Enabling distributed Processing and  
Management of Biological Data through fusion  
of Grid

and Web Technologies

Metabolic Engineering & Bioinformatics  
Program, National Hellenic Research  
Foundation (partner of HellasGrid)

HealthGrid 2010, Paris  
28 June 2010

# Current situation in Greece

In the last five years the bio-medical informatics community is continuously growing, covering a broad range of activities.

- Bioinformatics and Systems Biology ( transcriptomic analyses, motif analysis, translational clinical research, in-silico pathway modeling, Next Generation sequencing, metagenomics, etc.)
- Biological and Medical Imaging –images & video (Feature-based characterization and categorization of images, 3D-reconstruction, registration, storage & management of medical data, recognition of salient objects, compression etc.)
- Molecular modeling (Protein structure prediction, drug design, crystallographic studies)

Stakeholders: University Groups (13) and Research Institutes (7), scattered around Greece (mainland and Crete)

# Problems

Though diverse, these activities are not coordinated, remain unconnected, geographically isolated, with minimal if any reuse.

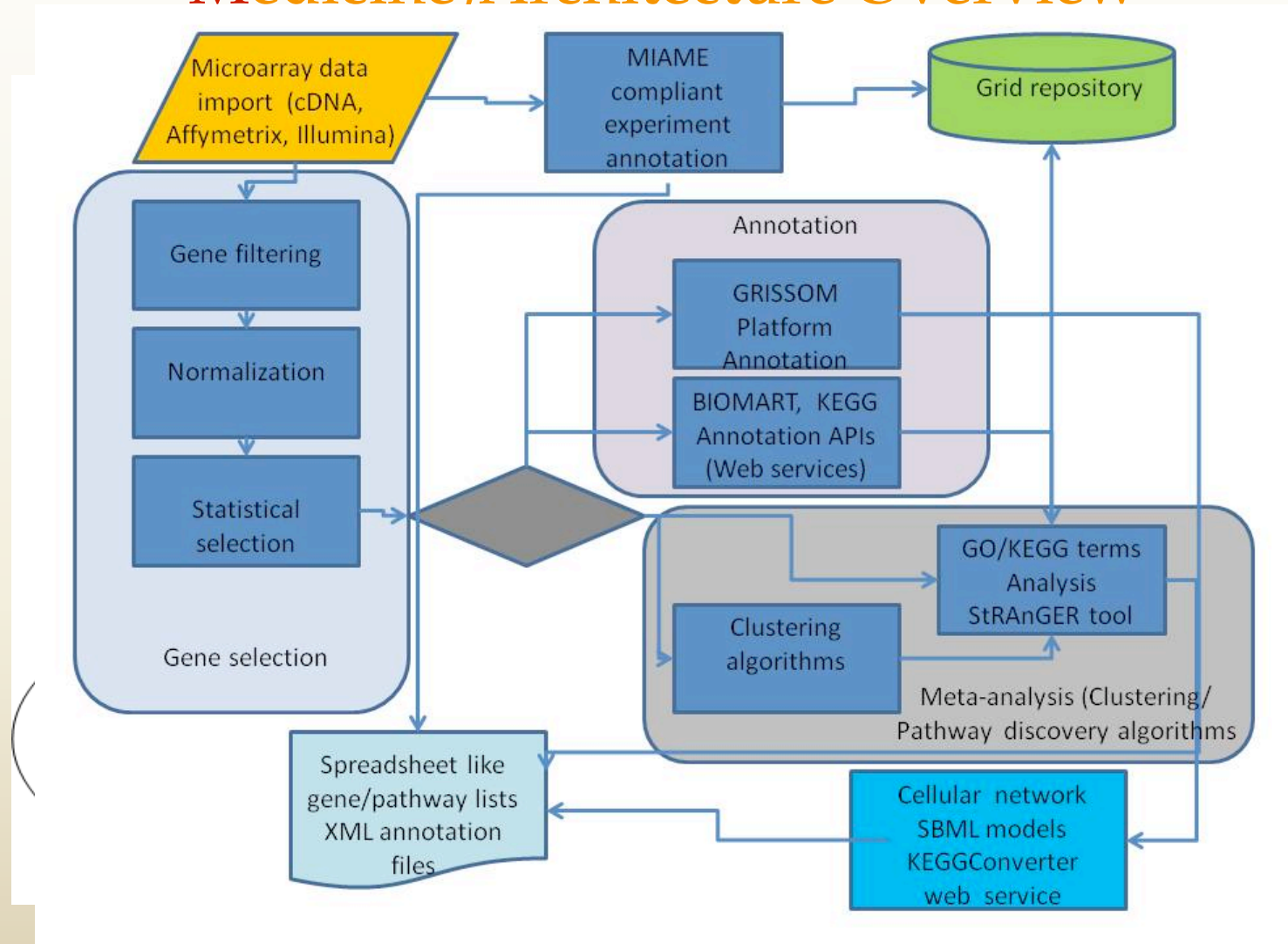
This picture is a miniaturization of the international situation.

The vast majority of these efforts does not make use of National Grid Infrastructures

Reasons:

- Technical complexity,
- Lack of familiarization with the distributed environment
- nudity of the infrastructure from tools, code, to facilitate rapid deployment,
- underperformance of the infrastructure

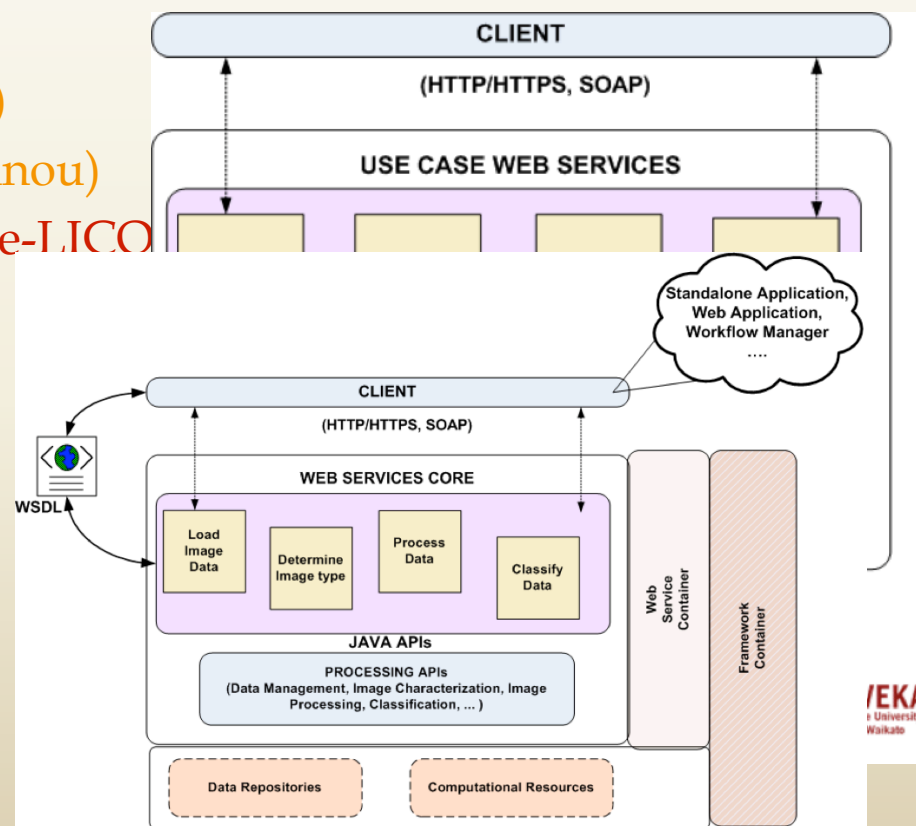
# GRISSOM (Grids for In Silico Systems biOlogy and Medicine) Architecture Overview



# The e-LICO image mining service oriented architecture

- Characterization of Kidney Obstructive nephropathy Images through a Web Service implementation
- Each step has been implemented as a Web Service function

Common work of University of Central Greece (Assist.Prof. I Maglogiannis) and NHRF (Res. Assist.Prof. A. Chatziioannou) FP7/ICT/Intelligent Content & Semantics/e-LICO



# Do we need a European structure for grid-related health and LS communities?

Yes, there is need and space for synergy in terms of:

- Data
- Tools/methodologies / architectures (reuse, expansion, assembly)
- Standards (data exchange, methodologies applied, infrastructures exploited)
- Performance improvement, Efficient community support, tangible benefits for the 'average researcher' user of the infrastructure (speed, quality, integration)
- Strategic expansion of LS research (public awareness, research policies, pressure for funding)

# What would be its role?

- Promotion of Biomedical distributed computing, through design of new architectures, environments, tools, services, applications, integration/assembly of popular tools.
- Promotion of standardization in semantic description, data formats, algorithmic design.
- Promotion of transparent, burden-free, integrative implementations for the user -development of virtual laboratories (analysis pipelines) for generalized tasks
- Gleaning of top performing code, refactoring for optimized grid performance, maintenance and upgrade.
- Effort for efficient integration of biological and medical information (translational research challenge)
- Push for operatively sound performance of the infrastructure (data integrity and security)
- Innovative Solutions for huge data volumes transfer, storage and management

# Who would be the targeted user communities ?

- As end-users, LS-related ESRIS, biomedical researchers, hospitals (strong need to ensure compliance with bioethics standards for data privacy), academia (for educational purposes-online help, tutorials, webinars required)
- For development purposes, the broader LS computing community (bioinformaticians, software engineers, parallel/distributed computing experts, application programmers, web technologies experts, AI experts etc.)

For this community a training roadmap should be shaped including online help&training, workshops, dissemination activities, but also calls for application development, plus policies, which promote access, in return for voluntary application development and support.



# What would be the supporting infrastructures ?

- EGI, ESFRIs when available
- National Grid infrastructures
- Local infrastructures that are keen in the promotion of Grid Computing.

Here the need for application development and support should be stressed, by recruiting active players of the LS computing community

# How would it be organized and how would it interact with the existing initiatives ?

- A flexible organizational scheme, which provides good coverage, primarily of all scientific areas pertaining to the field, but also inspires synergy among national partners, through visionary scientific targets (need for a strong scientific advisory board).
- Intense collaboration with EGI, NGIs and LS-related ESFRIS.
- Representation of all national players.
- Decisions should be made on scientific merit, technical feasibility, value for money, and determination to overcome obstacles, bureaucracy included.