ACGT
Advancing Clinico Genomic Trials on Cancer

ACGT:

Open Grid Services for Improving Medical Knowledge Discovery

Stelios G. Sfakianakis, FORTH

http://www.eu-acgt.org

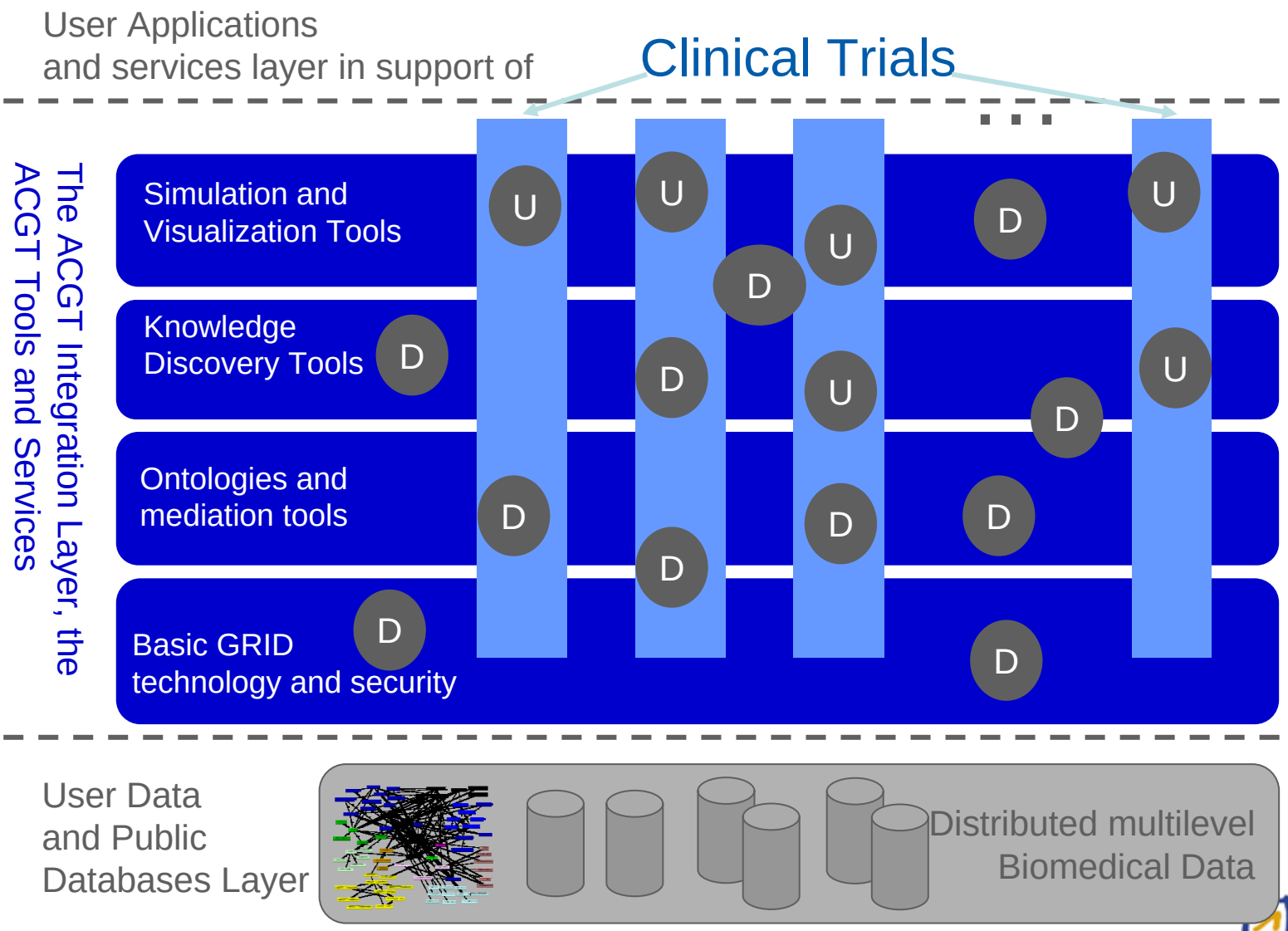
Information Society and Media

The ACGT project (FP6-2005-IST-026996) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.

The ACGT vision & principles

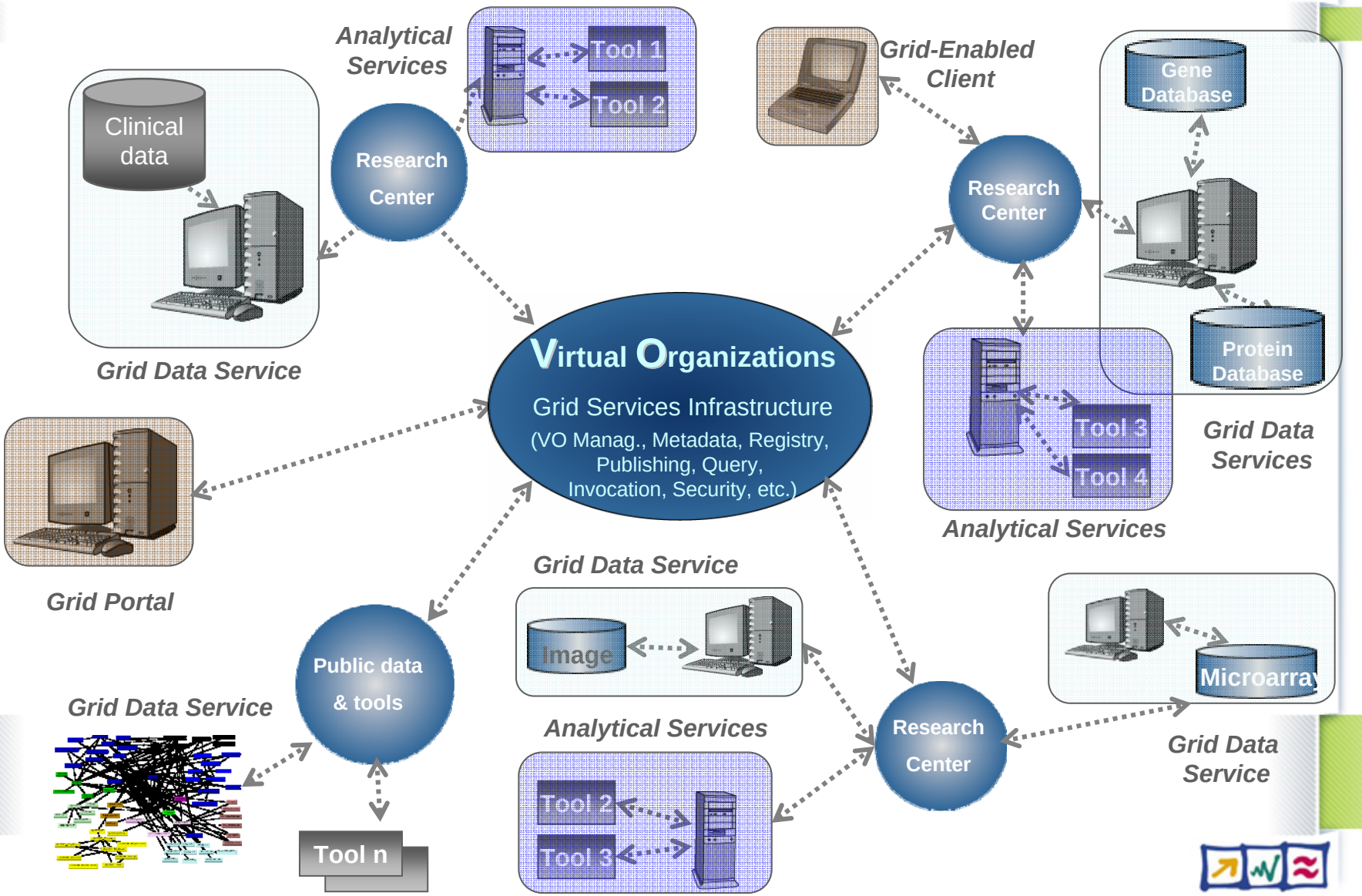
- The ultimate objective of the ACGT project is the provision of a unified technological infrastructure which will facilitate
  - integrated access to multi-level biomedical data
  - development or re-use of open source analytical tools, accompanied with the appropriate meta-data allowing their discovery and orchestration into complex workflows.
- ACGT will deliver a European Biomedical GRID infrastructure offering seamless mediation services for sharing data and data-processing methods and tools, and advanced security;
- ACGT
  - focuses on clinical trials on Cancer (Wilms tumor, Breast) and
  - is based on the principles of
    - Open access (among trusted partners)
    - Open source
  - is not a standards generating exercise but a standards adopting one.

Enabling dynamic Virtual Organizations

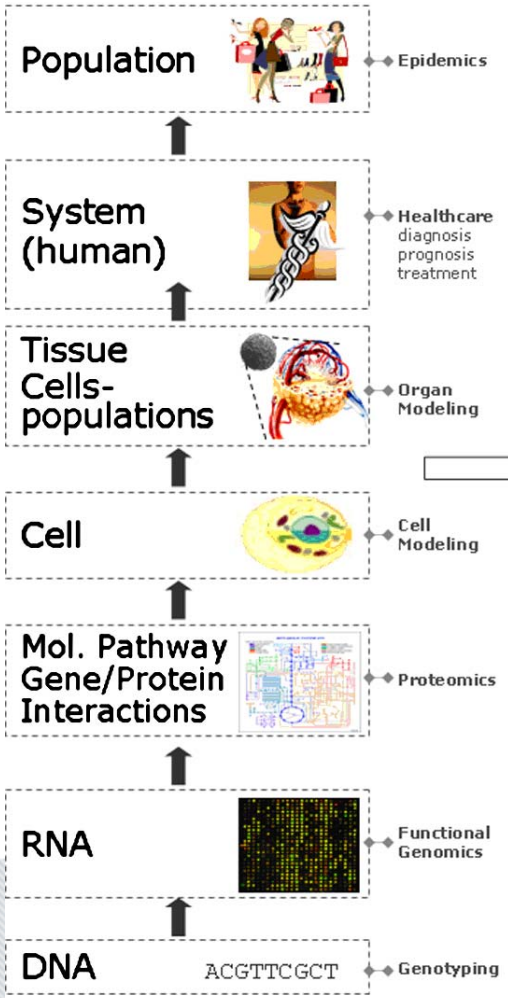


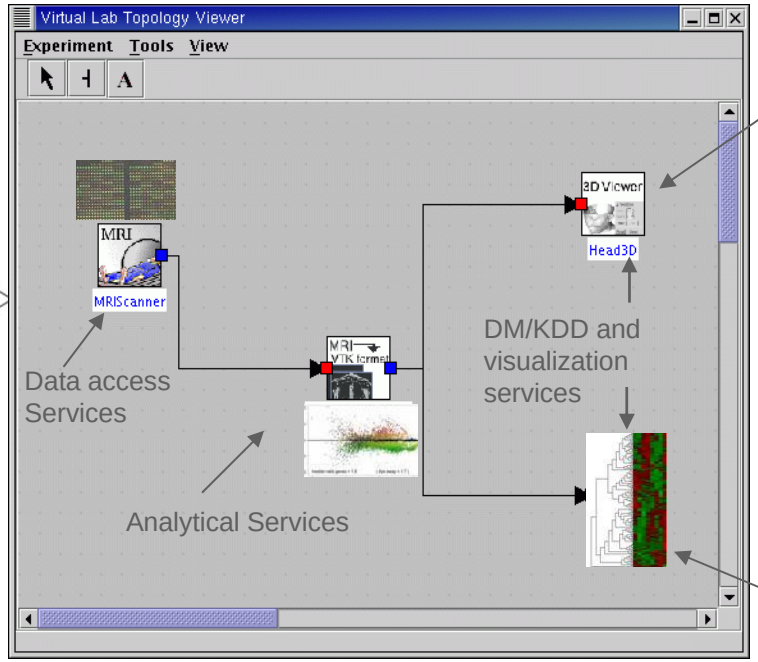
Advancing Clinico Genomic Trials on Cancer

The ACGT Virtual Organizations

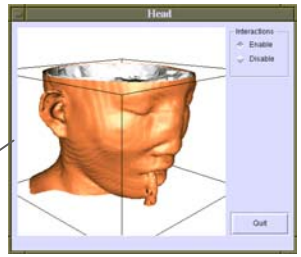


Discovery and Orchestration of Services

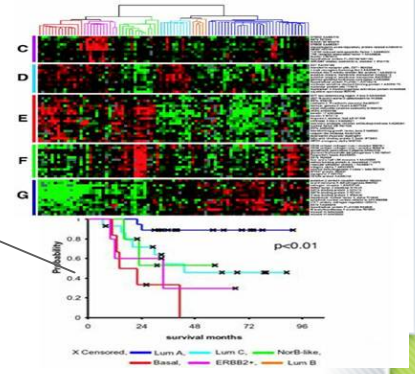




ACGT experiment topology



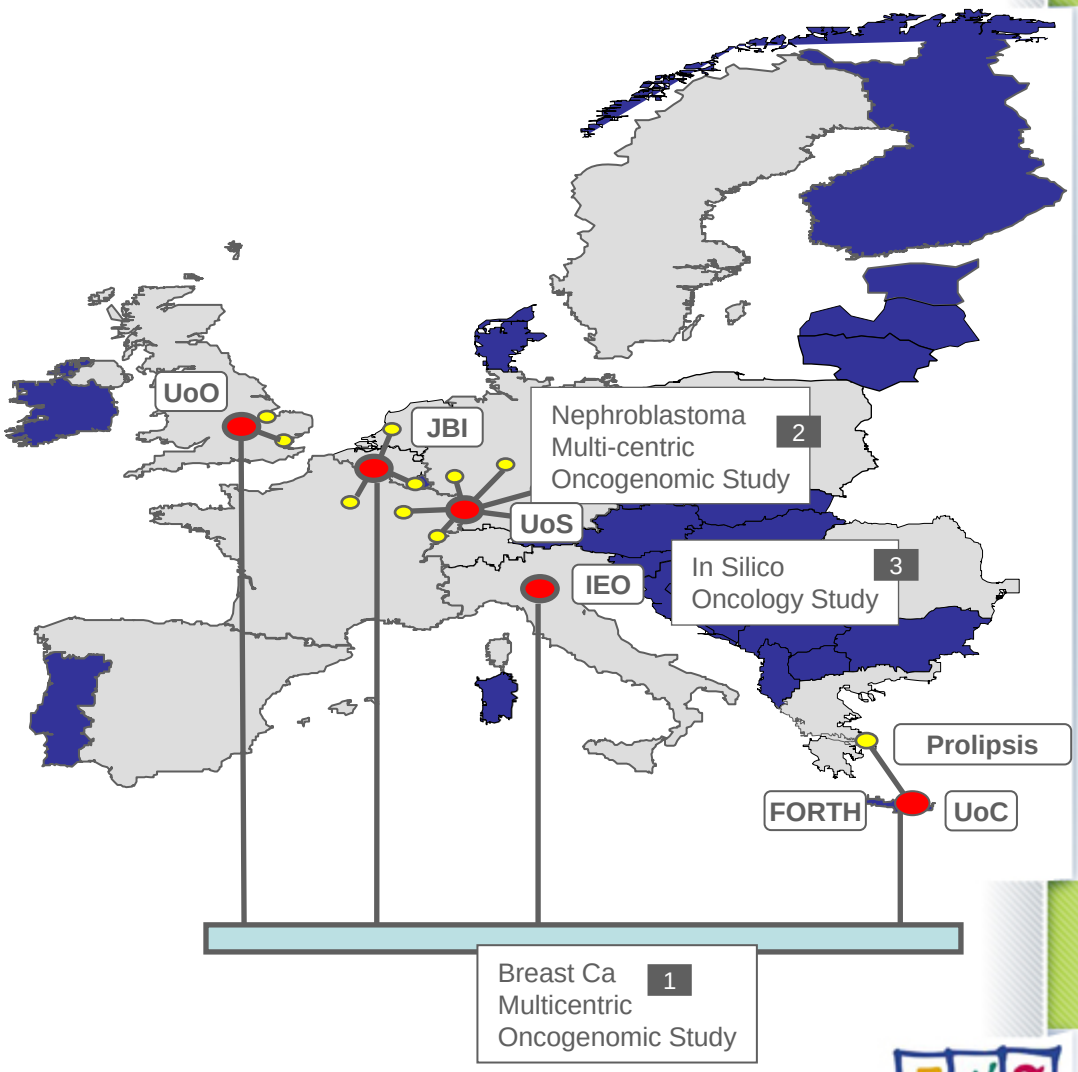
2D/3D visualization for in silico models



Microarray data processing for molecular classification of disease

The ACGT clinical trials

- Multicentric TOP trial – Breast Cancer
- SIOP 2002 – paediatric nephroblastoma
- In Silico modeling and simulation of tumor growth & response to treatment



Main challenges in ACGT

- **Grid middleware** services, enabling large-scale (semantic, structural, and syntactic) interoperation among biomedical resources and services;
- Master **ontology** (on Cancer) through semantic modelling of biomedical concepts using existing ontologies and ontologies developed for the needs of the project;
- Open source bioinformatic tools and other **analytical services**;
- Semantic **annotation** and **advertisement** of biomedical resources, to allow **metadata-based discovery** and query of tools, and services;
- Orchestration of data access and analytical services into complex eScience workflows for post genomic clinical research and trials on cancer;
- **Meta-data descriptions of clinical trials** to provide adequate provenance information for future re-use, comparison, and integration of results;

Major Challenge: Semantic Interoperability

- The bottleneck is not so much about:
  - computational needs,
  - the volume of data, or
  - performance issues in accessing/transferring data;
- It's integration and semantic interoperability;

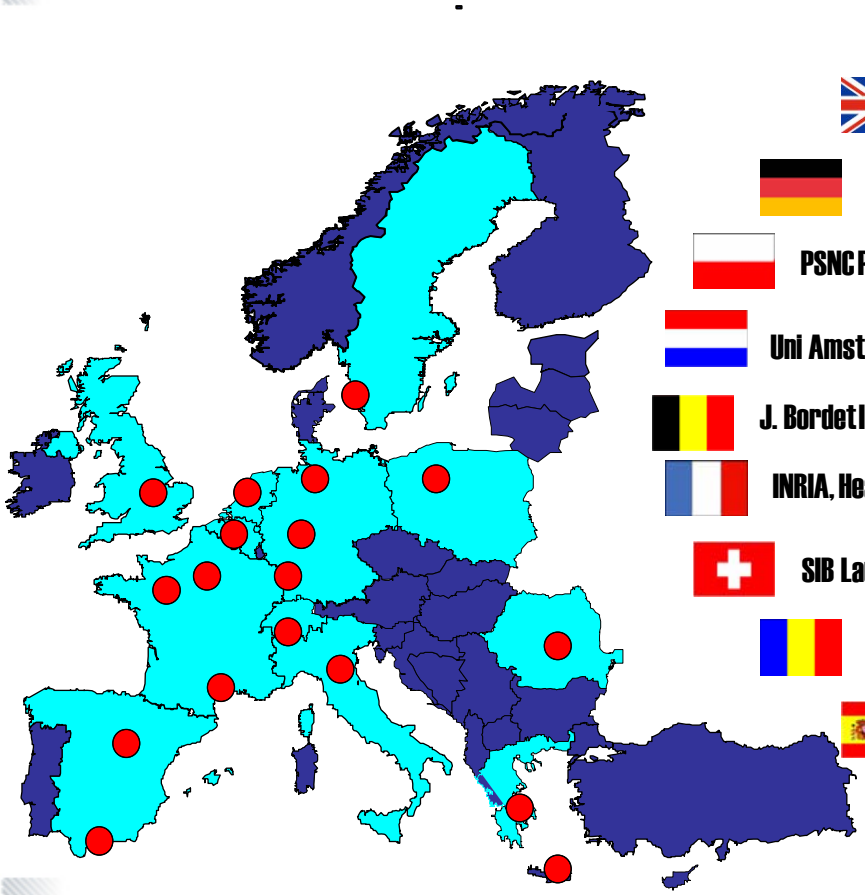
Data Integration Impediments

- Heterogeneity
  - Syntactic: Relational (SQL) Databases, web accessible databases, …
  - Structural: Different schemas and formats
  - Semantic: Different vocabularies and semantics
- Security related:
  - Different access policies: some data sources require authentication, whereas others are public
  - Sensitive and confidential data: patient names or other identifying traits should be hidden (anonymization, pseudonymization)

Required Services

- The primary services required for supporting the identified scenarios fall into four categories:
  - services for access and retrieval of data , that is: internal phenotypical (clinical and imaging) DBs and other "-omic" DBs, as well as external biomedical databases;
  - services that are the **analytical and visualization tools**, that is: computational analysis, simulations, knowledge extraction, exposed as Grid (web) services;
  - services for **forming and executing eScience Workflows**, that is:
    - workflow management services,
    - information management services, and
    - distributed database query processing;
  - **semantic services** for discovering services and workflows, and managing metadata, such as:
    - ontologies
    - metadata
    - provenance

The ACGT Consortium



Uni Lund

Uni Oxford

Uni Hamburg, Uni Hosp Saarland, IFOMIS
Fraunhofer (IBMT,AiS), Uni Hannover

PSNC Poznan

Uni Amsterdam, Philips

J. Bordet Institute, Custodix, Uni Namur

INRIA, HealthGrid, ERCIM

SIB Lausanne

SIVECO

Uni Madrid, Uni Malaga

FORTH, Uni Hosp Crete,
ICCS-NTU Athens, Biovista

Uni Hokkaido

Funding: ~18 MEuro, Time plan: 1/2/2006 – 31/1/2010

Thank you!