ACGT is a European Commission co-funded project supported by grant FP6-2005-IST-026996.

NEWSLETTER

ACGT on Cancer

Advancing Clinico Genomic Trials

EDITORIAL

Welcome to this first edition of the ACGT Newsletter!

ACGT is an EC co-funded project of the 6th Framework Program focusing on the development of open-source computer grid-based infrastructure and services that will support clinico-genomic trials in the area of cancer research.

Newsletter Edition

Issue 1 . December 2007

Editor-in-Chief

Dr. Andreas Persidis
andreasp@biovista.com

Editorial Board

Prof. Dr. Norbert Graf
Norbert.Graf@uniklinikum-saarland.de
Dr. Alberto d'Onofrio
alberto.donofrio@ieo.it
Mr. Nathanael Verhaeghe
nathanael.verhaeghe@healthgrid.org

Contact ACGT

Web: www.eu-acgt.org
Scientific Director:
Dr. Manolis Tsiknakis
tsiknaki@ics.forth.gr
Administrative Manager:
Mr. Remi Ronchaud
remi.ronchaud@ercim.org

Newsletter Design

HealthGrid
www.healthgrid.org

We are just about completing our 2nd year of existence and getting at a point where many of the ideas of the project are taking shape and being implemented in actual systems. At the same time a lot of interesting developments are taking place in the wider areas of clinical trials, IT infrastructures and semantic technologies to merit a newsletter that is dedicated to keeping practitioners in these fields up to date and in touch with each other.

We have designed this newsletter with the broader community in mind, not just the members of the ACGT consortium, and so we cover work that takes place within the project itself as well as developments that are broadly relevant to computer grid research, clinical trials, semantic technologies, IT services for life science research and eHealth in general. The newsletter will always host a feature article and for this first edition we have chosen to introduce ACGT itself in a bit more detail. We will also be hosting the views of prominent members of the community on interesting and important issues and for this edition Prof. Gordon McVie discusses the clinical possibilities arising from ACGT.

Our Clinical Trials and Grid News sections will be covering developments in both of these fields while the Products and Services section will be updating you on software or other tools that are available to the community for use.

We hope that you will find this newsletter interesting and look forward to you joining the growing ACGT community!

Information Society Technologies

Clinical Trials News

Latest developments in the world of clinical trials in cancer

Clinical Cancer trials are ongoing for almost every type of cancer. Most of these trials are solely based on clinical data which to a certain degree reduces our ability to better understand how this disease operates. In contrast ACGT focuses on the integration of multilevel biomedical data significantly increasing clinicians' ability to better understand cancer and thus help with the design of better therapies with higher cure rates.

ACGT clinico-genomic trials are devoted to archiving and analysing individual patient data on associations to molecular genetic findings in order to develop more individualized treatments for cancer patients. At the moment two clinical trials use the ACGT platform: the TOP trial for breast cancer and the SIOP 2001/GPOH trial for nephroblastoma, the most common childhood kidney cancer.

The multicentric prospective TOP trial aims to identify biological markers associated with pathological complete response to anthracycline therapy (epirubicin), one of the most active drugs used in breast cancer treatment. To this end, the neoadjuvant approach is very attractive, as it provides an in vivo assessment of treatment sensitivity without affecting adversely survival. To identify these predictive markers, the trial mainly uses gene-expression profiling based on microarrays as well as on genotyping technology.

The SIOP 2001/GPOH trial is a multicentre prospective randomized trial to evaluate the necessity of anthracyclines (doxorubicin, a cytostatic drug with the potential risk of heart disease (cardiomyopathy) as a late effect) in the treatment of unilateral non-metastasized nephroblastoma. Diagnosis is solely done by imaging studies. Patients enrolled in the trial will receive preoperative chemotherapy followed by surgery of the

tumour. Nephroblastoma is one of the success stories in cancer, where clinical trials did help to reverse prognosis during the last decades. Today "the integration of molecular biology into clinical trials is an absolute must which is why work being carried out by ACGT is of central importance" says Prof. Dr. N. Graf, Chairman of the SIOP trial .

Clinicians and other practitioners involved in the ACGT trials are supported by a variety of software resources. One of these resources is called ObTiMA (Ontology based Trial Management for ACGT). This tool which is currently under development enables the chairman of a clinical trial to set up and manage multi-centric trials from a single interface. From a clinical point of view ObTiMA will help to increase the number of clinico-genomic trials by facilitating the workload involved in creating new trials.

For more information of this article visit: http://eu-acgt.org/fileadmin/newsletter/20_11_2007/index.html

Products and Services

News on the latest products or services in our area of interest

When clinicians, bio-researchers or any other users think of IT supporting systems and tools, they think in terms of their daily tasks and routines and how such tools can help them perform these tasks more easily and more effectively.

ACGT is an infrastructure project but of course like any infrastructure its success will be judged on the basis of how much it is used by its target users and, in this case, by how many resources are available through it. ACGT is presently focusing on the significant amount of work needed to put this infrastructure in place but at the same time it is also looking to make available a number of services and other resources that will help end users perform their daily activities. These resources will span a broad area of interests including cancer tumor growth simulations, clinical trial design and management, literature research, statistical analysis tools and many more.

At the core of many of these resources will be a relatively recent technology going by the name of 'ontologies'. Ontologies are a promising development in computer science that capture reality in a way that allows software systems to better understand the semantics of the data being processed. Understanding the semantics of data offers significant advantages if done right, not least of which is the ability to integrate distributed resources in a way that is transparent to the user who does not need to worry about where the data

comes from but can rather focus on what this data is and how best to use it for the task at hand.

Ontologies represent a central resource for ACGT and one of the major contributions of the project to the community is the ACGT master ontology (MO). Currently including over 1000 concepts together with their relations and properties, the ACGT ontology focuses on diseases, patients, treatment opportunities and clinical trials. The MO is being developed with a modular design, integrating existing ontologies such as the Gene Ontology and the Foundational Model of Anatomy.

To make use of this resource, the consortium is developing a number of associated tools the first one to come online being the ACGT Ontology Viewer. This tool, which is freely accessible by the public, currently serves 3 purposes:

- To provide a view and entry point to the important concepts in the area of clinico-genomic trials on cancer
- To help create electronic health records using well structured and semantically exploitable information
- To allow the use of a multitude of databases through the mapping of their data in a semantically correct way

The Ontology Viewer can be freely accessed:
http://62.103.163.162:8080/OntologyViewer/

Additional information on the ontology can be found here:
http://www.eu-acgt.org/ontology.htm

Grid News

Latest developments in the world of computer grid research

After almost 2 years of background research and the specification of user requirements, ACGT has completed the architecture design of its Grid infrastructure. Put simply, a grid infrastructure is a collection of servers and communication protocols that allow highly complex and compute-intensive tasks to be shared by the computers in the grid in a safe and efficient manner.

Based on standard Grid technologies like Globus and Gridge, the ACGT Grid infrastructure is managed by the Poznan Supercomputer Center in Poland and at the time of publication of this newsletter includes computing nodes in Malaga, Brussels, SanKt Augustin, Crete and Poznan with plans to add more in the near future. In total 5 servers running mostly Linux provide a stable, distributed startup test bed with sufficient computing horsepower to handle a variety of compute-intensive applications such as the National Technical University of Athens' Oncosimulator tool.

The sensitivity of the personal data being accessed and the life-science applications that are at the core of ACGT together with the computational complexity of some of these, have put very strong security and load distribution efficiency requirements on the infrastructure.

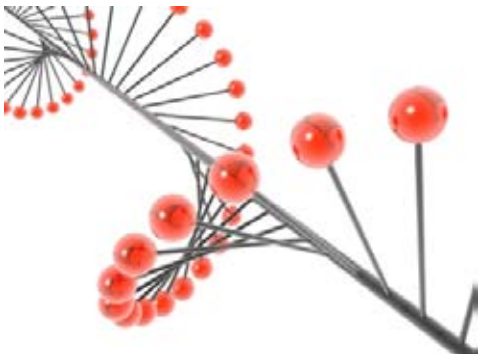
To meet these, ACGT has adopted standards commonly used in the Grid technology area; for example the security element is based on GSI technologies. Security tools are concerned with establishing the identity of users or services (authentication), protecting communications, and determining who is allowed to perform what actions (authorization). All ACGT users and services are obliged to identify themselves using certificates.

GSI technology supports authentication based on X.509 certificates while authorization decision is provided by the Grid Authorization Service – a higher level tool fully compliant with GSI.

For load distribution, ACGT exploits the grid meta-scheduling system called GRMS (Grid Resource Management System). GRMS which is part of the PSNC Gridge Toolkit, is able to react dynamically to changing states of the Grid environment and balance the workload of the available Grid nodes according to the current load of the processors.

Feature Article

The Vision and Objectives of ACGT



Advances in post genomic research have created significant opportunities for offering personalized treatment and better health care services to the population at large. At the same time clinical trials have become a bottleneck in terms of complexity, effectiveness and, in their present form, fitness for purpose. In the realm of information technologies on the other hand advances in semantic technologies and grid computing have reached a stage where multi-dimensional applications requiring the combination of heterogeneous data and software resources can be realistically tackled.

ACGT (Advancing Clinico-Genomic Trials on cancer: Open Grid Services for improving Medical Knowledge Discovery) is an Integrated Project (IP) funded in the 6th Framework Program of the European Commission under the Action Line "Integrated biomedical information for better health".

ACGT aims to develop methods and systems for improved medical knowledge discovery and understanding through the integration of biomedical information and the use of modeling, visualization, data mining and grid technologies. Biomedical data and information include not only clinical information relating to tissues, organs or personal health-related information but also information at the level of molecules and cells that is produced by genomics and proteomics research.

The project will design, develop and validate an integrated Grid-enabled technological platform in support of post-genomic, multi-centric Clinical Trials with an initial focus in the domain of Cancer. The driving motivation behind the project is our committed belief that the breadth and depth of information already available in the research community at large, present an enormous opportunity for improving our ability to reduce mortality from cancer, improve therapies and meet the increasing demand for personalized health care.

The problem with data integration

In addition to its shear volume, data collected for use in clinical trials and other analyses related with the provision of healthcare services is often published without the background information (method of capture, sample preparation, statistical techniques applied) that is needed to reproduce results. Even within a single laboratory, researchers have difficulty integrating data from different sources not only because of a lack of common standards but also because of other technological, medical, legal and even ethical issues. As a result, very few cross-site studies and clinical trials are performed and in most cases it isn't possible to seamlessly integrate multi-level data. The benefit of such multi-centric trials is hence never realized.

Apart from problems in sharing and re-using data, what is even more critical is the fact that clinical researchers or molecular biologists often find it hard to exploit each others' expertise due to the absence of a cooperative environment which enables the sharing of resources and tools, as well as a uniform platform that supports the seamless integration and analysis of disease-related data at all levels.

With this in mind ACGT aims to develop a European Knowledge Grid infrastructure offering high-level tools and techniques for the distributed mining and extraction of knowledge from data repositories available on the Grid, leveraging semantic descriptions of components and data and offering knowledge discovery services in the domain of Cancer research. Special emphasis will be given to the trust mechanisms that need to be embedded in the platform as well as the relevant ethical issues, thus creating conditions that promote service uptake.

ACGT technologies and approach – an overview

ACGT is a 4 year project run by a 25 member international consortium and represents a significant undertaking by any measure. To meet its objectives, the project is developing solutions in a multitude of technological areas, the main ones of which are:

• Focusing on the semantic integration of data but also on the discovery, integration, and management of sharable data assets (i.e. data and tools operating on such data). As a result the issue of metadata becomes of paramount importance for the successful achievement of the project objectives. The project is developing standards and metadata specifications describing and exposing web services (semantics), scientific services, and the properties of data sources, datasets, scientific objects, and data elements;

• Developing new, domain-specific ontologies, built on established theoretical foundations and taking into account current initiatives, existing standard data representation models and reference ontologies; The ACGT Master Ontology on Cancer plays a pivotal role in enabling (a) semantic information integration and (b) ontology based implementation of end-user applications;

• Developing innovative and powerful data exploitation tools, for example multi-scale modeling and simulation, considering and integrating from the molecular to the systems biology level, and from the organ to the living organism level, as well as technologies for interactive visualization on the grid;

• Developing an open biomedical GRID infrastructure offering seamless mediation services for sharing data and data-processing methods and tools;

• Set up cross-disciplinary task forces to propose guidelines concerning issues related to sharing legal, regulatory, and ethical data. ACGT is also developing advanced security tools (including anonymisation and pseudonymisation of personal data) for data protection in a web (grid) services environment in accordance with existing European legal and ethical regulations;.

• Developing an ontology based Trial Builder for helping to easily set up new clinico-genomic trials, to collect clinical, research and administrative data, and to put researchers in the position to perform cross trial (meta) analysis;

• Developing data-mining services in order to support and improve complex knowledge discovery processes;

• Creating an easy to use workflow environment, so that biomedical researchers can easily design their "discovery workflows" and execute them securely on the grid.

The ACGT consortium is currently focusing on the development of this core set of tools and components up to a stage where they can effectively support in silico investigations. Later stage developments will subsequently allow the consortium to improve and refine the capabilities of the services offered via the ACGT infrastructure.

To introduce the project to its intended target groups, a range of demonstrators, are currently under development. First prototypes are expected early in 2008 enabling the consortium members to begin evaluation and at the same time gather additional and more concrete requirements to guide future refinements and work.

ACGT opts for Open Source

The project promotes the principles of open source and open access, and has conceived an overall architecture for an integrated biomedical sciences platform. The infrastructure being developed enables efficient management of Virtual Organizations crated in the context of multi-centric, post-genomic clinical trials and uses a common set of services and service registrations for the entire clinical-trials-in-cancer community.

ACGT and the Community

ACGT is an infrastructure and related services project and as such its success ultimately hinges on its adoption by the intended user communities. Demonstration and awareness activities have already begun and will continue for the duration of the project. One of the planned highlights will be the ACGT competition, expected to take place in the middle of 2009. Open to everybody, it will bring together representatives from the healthcare, life sciences and IT communities rewarding entries that most convincingly demonstrate the use of the ACGT infrastructure and offer true value to the end users.

For more information on ACGT visit: www.eu-acgt.org

Community View

Invited contributions from non-ACGT members of the wider research community

ACGT clinical possibilities

by Professor J Gordon McVie MD, FRCP, FRCS Ed, FMedSci, DSc(Hon)

It is widely predicted in oncological circles that research is fast driving clinical care towards so-called personalised medicine. The first hint of this came from the unraveling of the human genome (and that. of other species too) . Say there are around 32,000 genes which regulate every cell in our body, each coding for how many proteins? And around each gene the "epigenome" whose loss of integrity may prove every bit as relevant to the carcinogenesis process. It's not clear what the possible numbers of combinations of dysfunctional regulatory molecules an emerging tumour cell may boast, but a conservative guess could be hundreds.

No longer can we defend the newspapers expectations of "Finding THE cancer gene", and "Fixing it with a replacement kit called gene therapy". It turns out that it's not as simple as changing the brake blocks, or mending the puncture in the tyre! The famous wiring diagram (originally from Science) of intracellular pathways , any combination of which could be disturbed in a cancer cell, is part of every Powerpoint presentation on "Targetted therapy". So of the 32,000 genes and over a million relevant proteins, and clever epigenetic camouflage, how do we get to "personalised medicine"? Sadly, only by collecting all these disparate bits of subcellular intelligence from cohorts of cancer patients who although superficially suffering from the same cancer, and are treated following a strictly observed protocol, will eventually display heterogeneity in outcomes. And only by collating all the bits of information (hundreds or thousands of bits) and subjecting them to rigorous mathematical modeling, will we ever identify the "personalised signature" relevant to each individual patient and his or her particular cancer. Then and only then, will the Holy Grail of "personalised medicine" become a reality. That's it. That's what I see ACGT offering from a perplexed but optimistic clinician's point of view.

EVENTS

Information on upcoming events of interest

ACGT is following an active dissemination strategy in reaching its various end-users and stakeholders. For the immediate future ACGT has planned 2 events:

ACGT will participate in the **e-Science 2007, the 3rd IEEE International Conference on e-Science and Grid Computing**. ACGT will be presenting a paper entitled "GridR: An R-based grid-enabled tool for data analysis in ACGT clinicogenomics trials".

In this paper, an approach devoped by ACGT for making the R system available in a large-scale Grid environment is described.

ACGT will organize a special Session on "Knowledge Discovery and Decision Support Systems in Health Information Systems", in collaboration with the EU FP6 funded projects HEARTFAID and MyHEART. The session will take place during the HEALTHINF 2008 conference (http://www.healthinf.org/Special_Sessions.htm), 28-31 January 2008, Madeira, Portugal.

Data Protection

The latest thinking on legal, ethical and data security issues surrounding clinical trials

Genetics present a tremendous variety of possibilities for medical care offering a lot of hope for better patient-care. As vast amounts of all kinds of data need to be processed when genetic research is done and as the trials are so complex to build up it is in principle a good idea to share data and computer power in order to achieve common goals faster. But this is unfortunately only half the story.

The other half of the story is that genetic data are of an extremely sensitive nature. They can have impact on the patient's life and give important information not only about the person himself but also about relatives, who might not even be born yet! Needless to say that every employer, every insurance company and every state prosecutor would be happy to have access to such data...

Several issues arise here. First, from the legal perspective, the right to privacy is a fundamental human right that has to be protected. Never may this sensitive data become the object of misuse.

Second, from the scientific perspective, if patients suspect that their personal data they agreed to share for medical care and research work may be used for other purposes (and it is quite irrelevant whether this other purpose is a legal or illegal one) they will simply refuse to allow access, creating an obstacle for further research.

Third, from a medical perspective, simply anonymizing the data and processing it without bothering about the patients does not provide a real solution. Patients have a legally and ethically guaranteed right to be informed, especially if the research done with their data leads to a result with a possible impact on their disease.

From its inception, ACGT has identified these issues and has set up a strong legal and ethical team that closely cooperates with the persons in charge for IT-security in the project. Full compliance of ACGT with all legal and ethical regulations is the first goal of this group and is reflected in the whole ACGT-architecture itself.

This is an important differentiator and asset of the project since partners, doctors and patients can base their work on a clearly defined legal framework. This is also an important contribution to the wider e-health community since the solutions found in ACGT will be adaptable and usable by many other projects in the *ICT for health* space.

Life in ACGT

Earlier this year, ACGT presented a poster at the ECCO 14 Meeting (14th European Cancer Conference) in Barcelona held from 23rd to 27th of September.

The title of the Poster was *"Clinical requirements of "In Silico Oncology" as part of the integrated project ACGT"* authored by: Norbert Graf, Christine Desmedt, Alexander Hoppe, Manolis Tsiknakis, Dimitra Dionysiou and Georgios Stamatakos.

Drs A. Hoppe, N. Graf and G. Stamatakos (left to right)

The ACGT team at Nice, France

ACGT PEOPLE

Dr. Manolis Tsiknakis
FORTH
tsiknaki@ics.forth.gr

Dr. Tsiknakis is the Scientific Director of ACGT and Principal Investigator and Leader of the Center for eHealth Technologies at the Biomedical Informatics Laboratory of the Institute of Computer Science at FORTH. His research interests include Intelligent Health Information Integration, Medical Informatics and eHealth, Distributed Architectures for Health Information Networks, HealthGRIDs, Component-based Software Engineering, and Socio-economic aspects of Health Telematics and eHealth services.

Several institutes and research groups of FORTH (www.forth.gr) are involved in the ACGT research activities. Namely, the Biomecical Informatics and the Information Systems Laboratories of Institute of Computer Science (www.ics.forth.gr), the Post-genomics laboratory of the Institute of Molecular Biology and Biotechnology (www.imbb.forth.gr) and the Biological engineering and systems biology Group at the Institute of Chemical Engineering and High Temperature Chemical Processes (http://www.iceht.forth.gr/).

Prof. Dr. Norbert Graf
University of the
Saarland
graf@uks.eu

Prof. Graf is the medical director of the Department for Pediatric Hematology and Oncology at the University Hospital of the Saarland and Chairman of the SIOP 2001/ GPOH Trial on Nephroblastoma. He is Dean for study affairs at the Faculty of Medicine of the University of the Saarland and member of several national and international societies. His research interests include Pediatric Oncology, especially nephroblastoma and brain tumors, Clinical Trials, Ethical issues in medicine, eHealth, eLearning and Medical Education. He is the leader of Work package 2 "Users needs and requirements" and Quality Manager in ACGT.

The Saarland University Hospital was founded in 1947 in co-operation with France. The hospital is a centre of medical excellence in Saarland offering a comprehensive range of medical, surgical, diagnostic and healthcare services covering practically all areas of modern medical practice. The Clinic for Paediatric Oncology and Haematology specializes in the diagnosis, treatment and follow-up of children, teenagers and young adults with all types of cancer. Psychosocial support is an integrated part of the care for these patients. Focusing its research on nephroblastoma and brain tumours, the team in Paediatric Oncology works hand in hand with other medical specialities and keeps close ties with a local parents group.

JOIN ACGT

Membership in ACGT is open to all. Here are some benefits you enjoy as an ACGT member:

- Access to all member resources
- Support in solving problems in the areas of interest of ACGT
- Direct contact with ACGT experts in a variety of fields including clinical trials, cancer research, advanced software development, Grid implementations, legal, ethical and data security issues and much more
- Ability to contribute to the ACGT infrastructure and receive support for it.

Join us at: eu-acgt.org

SUBSCRIBE!

Subscribe today and be among the first to learn of all the latest developments in ACGT and post-genomic clinical trials research on Cancer

3 EASY STEPS

1. Visit us at www.eu-acgt.org
2. Type in your name and email
3. Press the Subscribe button