Chapter XXIV
Gridifying Biomedical Applications in the Health-e-Child Project

David Manset
maat Gknowledge, France

Frederic Pourraz
maat Gknowledge, France

Alexey Tsymbal
Siemens AG, Germany

Jerome Revillard
maat Gknowledge, France

Konstantin Skaburskas
CERN, Switzerland

Richard McClatchey
University of the West of England, UK

Ashiq Anjum
University of the West of England, UK

Alfonso Rios
maat Gknowledge, Spain

Martin Huber
Siemens AG, Germany

ABSTRACT

The Health-e-Child project started in January 2006 with the aim of developing a Grid-based healthcare platform for European paediatrics and providing seamless integration of traditional and emerging sources of biomedical information. The objective of this chapter is to share experiences, and present major issues faced, solutions found and a roadmap for future work in developing the Grid infrastructure for interactive biomedical applications in the project, as Health-e-Child approaches its final phases. This chapter starts with a brief introduction of the project itself, followed by a description of its architecture on the Grid. It then illustrates the approach with the description of a concrete example of one integrated key application, the Health-e-Child CaseReasoner, which is intended for biomedical decision support over the Grid, and is based on similarity search and advanced data visualization techniques.

1. INTRODUCTION

In recent time demand has risen for more holistic views of patients' health so that healthcare can be delivered at the appropriate time, by the appropriate clinician, with the appropriate means at the level of individual patients. The Health-e-Child (HeC, pronounced "healthy child") project ("Health-e-Child," 2008) aims to provide data integration across heterogeneous biomedical information in order to facilitate improved clinical practice, scientific research and ultimately such personalised healthcare. As one of the largest integrated projects of the 6th Framework Programme of the European Commission, HeC brings together three major paediatric medical centres with several European companies, university groups and research institutions specialised in Grid-based biomedical information integration and related technologies.

The main objectives of the HeC project are:

- To gain a comprehensive view of a child's health by vertically integrating biomedical data, information and knowledge that spans the entire spectrum from the genetic through clinical to the epidemiological;
- To develop a biomedical information platform, supported by sophisticated search, optimisation and matching techniques for heterogeneous information, empowered by Grid technology;
- To build enabling tools and services on top of the HeC platform, that will lead to innovative and better healthcare solutions in Europe, based on:
    - Integrated disease models exploiting all available information levels.
    - Database-guided biomedical decision support systems provisioning novel clinical practices and personalized healthcare for children.
    - Large-scale, cross-modality, and longitudinal information fusion and data mining for biomedical knowledge discovery.

The realization of these project goals requires an infrastructure that is highly dependable and reliable. Indeed, physicians demand guarantees that the system will be always available and that the processes which integrate and manipulate patient data will be reliable, even in the case of failures. The infrastructure will have to allow for transparent access to distributed data, to provide a high degree of scalability, and to efficiently schedule access to computationally intensive services by applying sophisticated load-balancing strategies. Consider a scenario where a similarity search across the entire HeC patient population is needed to make a better decision over a critical case. In order to support such a search possibility on demand intensive query processing, feature extraction and distributed similarity calculations have to take place. All these steps require significant computing power, storage capacity and an acceptable quality of service (QoS) over the infrastructure resources.

Consequently, the HeC project has as one of its primary objectives, the delivery of a complete suite of Grid-based and cost-efficient tools for individualised disease prevention, screening, early diagnosis and therapy and associated follow-up for paediatric diseases across three different domains; cardiology (e.g. Right Ventricle Overload caused by Atrial Septal Defect or the Tetralogy of Fallot), rheumatology (Juvenile Idiopathic Arthritis), and neuro-oncology (e.g. Pilocytic Astrocytoma). To facilitate this, it has started building a gLite-enabled European network linking leading clinical centres to enable them to share and annotate biomedical data, to validate systems clinically and to disseminate clinical excellence across Europe by establishing new technologies, clinical workflows and standards in the domain.

The project brings together three heterogeneous communities, in a well-balanced configuration, which can be described as three equally important cornerstones: