IST-2004-027749
HeC
Health-e-Child

Integrated Project

Thematic Priority: IST

D.1.3 b Annual Report
Periodic Activity report

Reporting period: from January to December 2007   Date of preparation: 16 February 2008

Start date: 1st January 2006   Ending date : 31 December 2009

Project Coordinator   Joerg Freund   Revision: Joerg Freund
Siemens Medical Solutions

PUBLISHABLE EXECUTIVE SUMMARY

The Health-e-Child (HeC) project aims at developing an integrated healthcare platform for European paediatrics, providing seamless integration of traditional and emerging sources of biomedical information. The long-term goal of the project is to provide uninhibited access to universal biomedical knowledge repositories for personalised and preventive healthcare, large-scale information-based biomedical research and training, and informed policy making.

The objectives of the HeC project are: i) to gain a comprehensive view of a child's health by vertically integrating biomedical data, information, and knowledge, that spans the entire spectrum from genetic to clinical to epidemiological; ii) to develop a biomedical information platform, supported by sophisticated and robust search, optimisation, and matching techniques for heterogeneous information, empowered by the Grid; iii) to build enabling tools and services on top of the HeC platform, that will lead to innovative and better healthcare solutions in Europe and across the world, namely:

– Integrated disease models exploiting all available information levels;

– Database-guided biomedical decision support systems provisioning novel clinical practices and personalised healthcare for children;

– Large-scale, cross-modality, and longitudinal information fusion and data mining for biomedical knowledge discovery.

The partners involved in HeC are: Siemens AG (Siemens), Lynkeus Srl (Lynkeus), I.R.C.C.S. Giannina Gaslini (IGG), University College London – Great Ormond Street Children's Hospital (UCL), Assistance Publique Hopitaux de Paris – Necker (APHP), European Organisation for Nuclear Research (CERN), University of the West of England (UWE), University of Athens (UoA), Università degli Studi di Genova (DISI), French National Institute for Research on Computer Science and Control (INRIA), European Genetics Foundation (EGF), Aktsiaselts Asper Biotech (Asper), Gerolamo Gaslini Foundation (FGG), Maat G Knowledge (MAAT).

A new unfunded partner for the coming two years is going to be the Ospedale Pediatrico Bambino Gesù (OPBG) in Rome.

The Project Coordinator is Dr. Joerg Freund (Siemens), the Project Management Team Leader is Prof. Edwin Morley-Fletcher (Lynkeus), the Clinical Coordinator is Prof. Giacomo Pongiglione (IGG), and the Technical Coordinator is Dr. Martin Huber (Siemens). The Chairman of the Scientific Committee is Dr. Dorin Comaniciu (Siemens), the Chairman of the Ethical and Legal Review Committee is Prof. Alberto Martini (IGG), the Chairman of the IPR Committee is Prof. Yannis Ioannidis (UoA), the Chairman of the Executive Board is the Project Coordinator, the Chairman of the Governing Board is Dr. Alok Gupta (Siemens).

The HeC project is organised in 16 Work Packages, which have achieved the following results in months 12-24 of activity:

In the course of 2007 the management of HeC continued to be taken care of with the support of the core team composed of the Project Coordinator, the Clinical Coordinator, the Technical Coordinator and the Project Manager.

An important role was also played by the Project Management Team (PMT) in its enlarged form (as agreed on at the end of 2006, with all partners of HeC being represented in it), which has become the forum where preliminary analysis of issues and strategic decisions are regularly discussed in order to prepare for, and smooth, the decision making process of the Governing Board. The PMT met five times (once by telephone conference (TC)) within 2007. The Governing Board met in person three times, and once via TC. The core management team had weekly TCs, guaranteeing a constant monitoring of current developments and assuring day by day governance.

Specific priorities were:

Assuring the vertical integration of data and the technological integration of HeC's Grid Gateway

with the various applications being developed, by better alignment and coordination of the activities guided by the various relevant WP leaders.

Preparing for the possible entry of two new clinical partners into the consortium, correlated also to solving the issue of the partially incomplete nature of the data collected for the cardiology tests, with regards to the genetic data received from two of the original three clinical partners. The solution found was to have one new clinical partner, OPBG, allowing for a consistent increase of available data in cardiology, globally aligning HeC with the required level of correlated genetic tests.

Visiting John's Hopkins University Hospital in Baltimore in July 2007, starting the ongoing negotiations for having JHU to become an official member of HeC by the first or second quarter of 2008. Establishing such a relationship between HeC and JHU has been considered as particularly interesting also with a view to enhancing EU-USA cooperation in eHealth and healthcare research.

Analysing under-spending issues concerning some of the partners (especially 2 of the 3 originally participating hospitals), which led to investigating structural under-spending (due to a long ramp up phase, specifically for the clinical partners) and functional under-spending (resulting from the slower speed at which some parts of the project had been developed).

Two major Milestones and altogether 23 deliverables were all complied with on time, while the updating of two deliverables, D16.1 and D4.1, was slightly delayed.

On the whole, HeC's management system coped with the consortium's governance while providing also some strategic foresight. It succeeded in increasing cooperation and productive interaction among all partners, and attracting potential new unfunded contractors.

HeC's Grid platform and related applications, a new MRI-based scoring system for rheumatology, and Ontology Driven Uncertainty Modelling are the elements showing the greatest innovation vis à vis the current state of the art. The HeC Grid platform ranked in the top 5 out of 50 proposed demonstrations at the EGEE07 Conference in Budapest.

Other important achievements in 2007 were the following:

Setting up the HeC grid gateway at several partner sites, deploying grid nodes at 6 IT partners and 2 hospitals. Grid nodes will be deployed at the remaining two hospitals (including the new clinical partner, OPBG) in the first quarter of 2008.

Improving Data Collection for clinical and imaging data, making it likely that the goals which were self-assigned will be met by June 2008.

The status of data collection improved in the second year. While the data collection for the baseline visits is phasing out, the data collection for follow-up visits in cardiology and rheumatology patients is ramping up. This follow-up data is crucial for building and validating the predictive models.

Focusing on Machine Learning based feature extraction methods and presenting the first steps towards personalized disease models for cardiology.

Work Packages main activities:

WP2 – User Requirements Specification: New and more complex use cases were engineered, and these use cases were analyzed and documented primarily from the clinical point of view, putting more emphasis on clinicians' perception of the system. The first URS document was revisited and updated.

WP3 – **Ethical and Legal issues:** Only one family at IGG refused (in the Heart Disease Study) to take part in HeC out of all clinical centres and all diseases.

WP4 – **Privacy & Security:** Integrated peer to peer patient privacy (P2P3), One Time Password (OTP) and the GRID infrastructure were evaluated with respect to security and privacy.

WP5 – **Grid Platform:** WP5 released on time D5.1. "Report on Prototype Grid Platform" which provided detailed technical information about HeC requirements, the Grid infrastructure and gLite's integration in the Gateway software. A second version of the HeC Infrastructure Questionnaire was circulated. New requirements were communicated to WP5 in December 2007, most requests related to the Knowledge Discovery and Decision Support Systems.

WP6 – Medical Vertical Knowledge Representation: The modelling of clinical data based on the protocols and existing databases was completed and the UML model for brain tumours was produced in addition to the models for cardiology and rheumatology. The REMIND tool used for the data acquisition was put online and RVO forms have been developed.

WP7 – **Data Management Layer & Data Integration Mechanisms:** The first prototype of the HeC Gateway was released for the project's annual review at the European Commission (March 2007). DICOM analysis, Database abstraction Layer Client library, Peer-2-Peer Patient Privacy (P2P3), and other functions were added to the platform and integrated into a usable Live demonstration. Significant modification of the authentication client GUI made it more attractive and user-friendly. Modification of the Patient Browser GUI client to make it more attractive and user-friendly too, as well as to enable it to deal with the new Portal interface.

WP8 – **Medical Query Processing:** A preliminary definition of the Medical Query Language (MQL) was produced for the Medical Processing Engine (MPE), as well as an initial query processing and optimization strategy (planner, optimizer), a simple implementation of a work-flow processor to execute the optimized query, and the framework of the custom operator development. Both, MQL and MPE-DEMO were presented at the consortium meeting in London in July 2007. The second half of the year was focused on MPE integration.

WP9 – **Data Collection: The Paediatric Heart Diseases** study is proceeding at APHP, IGG, and UCL, and will get additional support from the new clinical partner, OPBG, in the course of months 25-42. 75 patients (70 RVO - 39 ASD, 30 TOF, 1 PAPVR - and 5 CMPs being 2 hypertrophic CMP and 3 dilated CMP) had been enrolled at **IGG** (Month 24) - 19 follow-ups. 129 patients (80 with TOF or ASD and 49 with CMPs) were recruited at **APHP/Necker**, and 58 patients were enrolled at **UCL/GOSH**. For the **Inflammatory Diseases,** 53 JIA patients were enrolled at **IGG**, 17 of which have had 1 year follow-up, 52 patients were enrolled at **APHP**, and seven of them have had a one-year follow-up visit. At **UCL** recruitment started in July 2007. By the end of December, 36 patients were recruited. For **Brain Tumours a** total of 49 patients were enrolled.

WP10 – **Ground Truth and Clinical Knowledge Gathering:** The Annotation tool for cardiac MR is now able to visualize multiple 4D sequences fused together improving the accuracy of finding anatomical landmarks. A RV mesh can be placed in a volume by selecting 4 landmarks in ED and ES time frames. Testing of the tools started with echocardiography studies from APHP-Necker and IGG, cardiac MR studies from UCL-Gosh, and JIA studies from UCL-Gosh and IGG. The annotation tools specifically developed for HeC (RVO and JIA) were refined and used to annotate 35 cardiac MRI and 13 JIA MRI cases.

WP11 – **Integrated Disease Modelling:** Electro-mechanical model of the heart was calibrated to simulate a healthy child heart. Simulations with modified parameters in accordance with clinical observations were also performed for CM and RVO, giving the first disease-specific (but still patient generic) models. For rheumatology, a significant amount of work was devoted to testing and improving segmentation tools for synovitis. Hard threshold methods were evaluated, especially in terms of robustness and variational methods for foreground/background separation in image processing were investigated. The development of image segmentation algorithms to measure the synovial membrane volume and bone erosion in MRI was continued. New fast numerical schemes to estimate the tumour cells density that take into account correct boundary conditions were designed and implemented. An important step was reached for planning patient-adapted irradiation margins for radiotherapy on a single 3D MRI thanks to the tumour growth model: experiments with simulated tumours showed that this better targets tumour cells and avoids healthy tissue than the usual constant margin.

WP12 – **Decision Support System**: A first stand alone version of a case based reasoning tool for HeC (called CaseReasoner) was developed. The CaseReasoner was further refined, integrated into the overall HeC client and connected to the grid. Four tools that extract clinical features from clinical raw data were released and serve as a necessary pre processing step for knowledge discovery and decision support algorithms that rely on cleansed, normalized and quantified features.

WP13 – **Biomedical Knowledge Discovery:** Novel learning algorithms for (multi)classification and feature selection were developed, including statistical learning tools based on regularized

kernel methods, Bayesian learning methods and causal feature selection approaches, and a consistent statistical method for describing the geometric variability of anatomical shapes using currents was developed.

<u>WP14 –</u> **Deployment of the Data Management System & Grid Gateway:** The first complete client library was developed, making all the platform functionality available to clients abstracting them from the complexity of the grid. Focussing on the Pre-Production Environment, gLite services were deployed at the Siemens node, 12 gLite services were deployed at IGG and 13 Central services for the PROD environment were deployed at CERN. During the course of deploying the SIEMENS node, a first step towards integration of WP12 client applications was made. In the third and fourth quarters also UWE, and UOA were deployed. The necessary integration work was carried out together with INRIA and SIEMENS, to connect their 3D Registration Tool and CaseReasoner to the HeC Gateway. Final Hardware specification reached a consensus at GOSH. DISI provided VMWare virtual Machines with enough IP addresses; it was negotiated to put the servers outside of the DMZ. The demonstration of the integrated prototype system was submitted to the EGEE 07. It was ranked in the top 5, out of 50, in the competition.

<u>WP15 –</u> **Training:** The 1st HeC Course in *"Genetics of rheumatic and auto-inflammatory diseases"* was held in Bertinoro di Romagna (Italy) from April 1st to 4th 2007. The Course was a live web cast, post-produced and made available for on-demand streaming. The 2nd HeC Course in *"Cardiogenesis and Congenital Cardiopathies: from Developmental Models to Clinical Applications"* has been organized for June 7-10 2008.

<u>WP 16 – Dissemination</u>: The consortium broadened the geographic range of its dissemination activities. Having previously focused its dissemination activities in the EU, 2007 saw the developments and achievements of HeC disseminated across the globe. Another fundamental part of the consortiums dissemination activities involved the preparation of scientific papers, presenting some of HeC's results. The consortium also teamed up BioMed and Healthgrid for a combined booth and dissemination activities at EGEE 07. 2007 saw the beginning of the consortiums Active Clinical Dissemination strategy.