The TUMOR Project: Integrating Cancer Model Repositories for Supporting Predictive Oncology

Vangelis Sakkalis[1], Stelios Sfakianakis[1], Kostas Marias[1], Georgios Stamatakos[2], Fay Misichroni[2], Dimitra Dionysiou[2], Steve McKeever[3], David Johnson[3], Thomas S. Deisboeck[4], and Norbert Graf[5]

[1]Institute of Computer Science at FORTH, [2]Institute of Communication and Computer Systems National Technical University of Athens, [3]Department of Computer Science, University of Oxford, [4]Harvard-MIT (HST) MGH, [5]Saarland University Hospital, Paediatric Haematology and Oncology

Correspondence: sakkalis@ics.forth.gr ICS-FORTH, Vassilika Vouton, GR-70013 Heraklion, Crete, Greece

In order to perform *in silico* modeling of cancer there is a need to study the various phases and scales describing different levels of biocomplexity using mathematical modeling and simulation. Such computational multiscale models often bypass the initial tumor genesis stage and focus mainly on the growth phase. In order to better relate various phenomena occurring in different scales, it is necessary in the one hand to account for microscopic processes when trying to predict macroscopic tumor growth, but on the other hand one needs to be able to correlate microscopic variables with a number of clinically meaningful parameters related to macroscopic phenomena often measured in clinical practice [1,2].

So far, significant but highly fragmented efforts have been made on both sides of the Atlantic to develop and use models of pathophysiology in order to better understand human function and promote individualized, patient-specific optimization of disease treatment. The TUMOR project, an EU FP7 funded project, is developing a European clinically oriented semantic-layered cancer digital model repository from existing EU Virtual Physiological Human (VPH) related projects designed to be interoperable with the US grid-enabled semantic-layered digital model repository platform at CViT.org (Center for the Development of a Virtual Tumor). Models and data will drive advances in cancer modeling with the ultimate goal to build a strong bridge that will pave the way for an integrated, interoperable transatlantic research environment offering the best available models and tools for clinically oriented cancer modeling and serving as an international validation/clinical translation platform for predictive, *in silico* oncology.

To achieve this ambitious goal, the TUMOR project sets the following specific objectives:

- To develop a European clinically oriented, semantic layered cancer multi-scale digital model and data repository from existing EU projects in order to amalgamate interoperability work with various EU project platforms thus saving resources and ensure that sensitive data issues are addressed altogether by the hospital provider.
- To build-up interoperable interfaces between this repository and the US semantic-layered digital model repository, CViT.
- To implement and demonstrate an integrated, interoperable transatlantic 'predictive oncology' workflow environment prototype with the following features: remote data access, application of tools/ models, and visualization of results in the 'transatlantic' context.
- To address specific clinical questions and scenarios that will drive the development of all the above and demonstrate the added value of TUMOR in applying shared models/data in common, clinically relevant 'transatlantic' studies.

To enable interoperability between EU and US executable models, a generic multilevel simulation execution environment is being developed. The models are mostly based on the so called "top-down" approach [3,4] that is based on how cells change their state based on factors outside of the cells themselves, where these models can be generalized to clusters of cells [5,6]. For the clinical application, drugs and radiotherapy are regarded as environmental factors. The US CViT models of reference also use a variety of modeling techniques, but predominantly take an agent-based approach where the individual cells have behaviors built into them. For example, molecular pathways are basically models of the chemical processes that cause a cell to shift state or result in an action such as cell division, cell death (necrosis/ apoptosis) and even proliferation or migration [7,8]. In this case, the "bottom-up" approach is being able to model within a single cell the molecular and chemical processes, and their resulting actions.

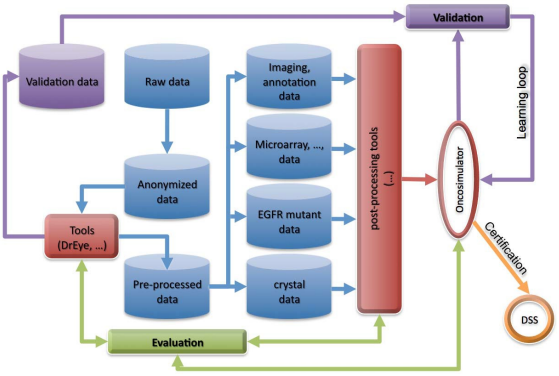

Figure 1 The workflow in more detail

To support the vision for model-assisted predictive oncology, a biologically meaningful, complex fusion of both bottom-up and top-down models is of utmost importance. To support this fusion, as well as the seamless integration with input data, a specialized workflow editor needs to include a number of tools dealing with the pre-processing of data, anonymization and pseudonymisation processes, the linking and execution of bottom-up and top-down models, and the visualization and validation of results. A clinician does not want to deal with the building of such a workflow; instead it is desirable to start a workflow and in an interactive and intuitive dialogue to proceed through the whole workflow until the model is executed (Figure 1). If the result of the model is to be used in a clinical setting, it needs to be automatically validated (with a validation tool within the workflow) and delivered in due time allowing its use in the decision process for treating the patient.

A critical requirement for TUMOR is that according to the current legal and ethical regulations and restrictions, in both Europe and the US, even the exchange of retrospective data is not a trivial task. Data needs to be stored locally in Europe or the US and not exchanged between legal domains. To overcome this problem, tools and models need to be exchanged and shared to run simulations with local data. This solution has significant implications for the type of infrastructure that TUMOR is developing.

TUMOR presents a number of integration, interoperability, and security related challenges. In designing the architecture we have followed the approach of views and viewpoints, which is standardized by IEEE/ISO [11]. Starting with the requirements and the functionality that the TUMOR platform aims to deliver, we have identified the following software components and their responsibilities:

- The European Model and Data Repository: This is the "main" model repository, located in Europe, where the cancer models of the European users and their anonymized data are stored and curated.
- The US Model Repository: This is the American model repository, located in the US (in MGH premises), and operated by CViT. This is where US-CViT users store their models and data. It can be accessed from the European side but only the models can be transferred due to the legal and ethical requirements.
- The Workflow Editing and Enactment environment, which is the Web-based application that allows the construction of simulation experiments through the linking of the available cancer models [12,9,10]. In order to do this, the Workflow Environment accesses the EU and US model repositories and selectively retrieves models from each. It is hosted within the EU and therefore has access to the data stored in the EU repository. Nevertheless since it is a Web application, it has to make authorization decisions based on each user's profile in order to restrict the data access mechanisms only to the European users. The execution of the workflows is taken care of a cluster of processing machines physically located with the workflow environment's server machine.
- The Common Access Point (CAP, for short): This is the main "entrance" to the platform. It is a Web portal for interacting with the majority of the TUMOR services. Behind this portal there will be the EU Model and Data repositories and also the users profile database.

The deployment architecture of TUMOR is shown in Figure 2.The models themselves are described using TumorML, a new markup language (ML) for describing cancer models [15]. The development of TumorML contributes to enabling some of the key interoperability aims of the TUMOR project. Firstly, by annotating cancer models with document metadata, digital curation is facilitated in order to make publishing, search, and retrieval of cancer models easier for users of the TUMOR digital repository. Second, markup is used to describe abstract interfaces to model implementations allowing the Workflow Environment to run simulations using models published to the model repository. Finally, TumorML markup facilitates the composition of compound models, regardless of scale and source, enabling multiscale models to be developed in a modular fashion, and models from the US CViT to be integrated with EU models in the

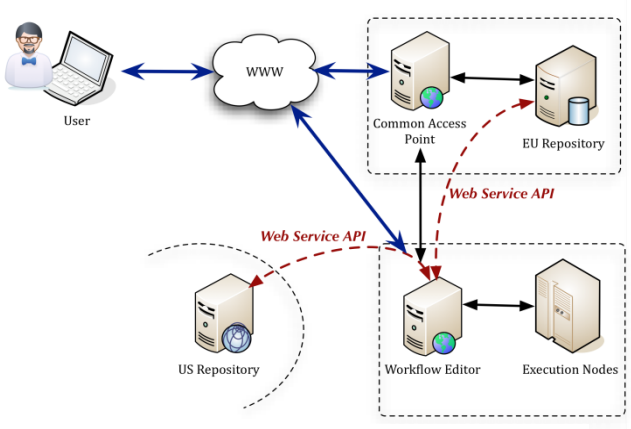

Figure 2 The deployment of the TUMOR platform

TUMOR transatlantic scenarios by using the Workflow Editor. The architecture therefore is designed with "service orientation" in mind, i.e. the software components expose a Web Service programmatic interface [13,14].

Acknowledgment

This work was supported in part by the European Commission under the TUMOR (FP7-ICT-2009.5.4-247754) project.

References

1. P.P. Delsantoa, et al., "A multilevel approach to cancer growth modeling", Journal of Theoretical Biology, Vol. 250, No 1, pp. 16-24, 2007.
2. T. S. Deisboeck and G. S. Stamatakos Eds, Multiscale Cancer Modelling. Chapman & Hall/CRC, Boca Raton, Florida, USA, 2010, ISBN: 9781439814406.
3. A. Roniotis, V. Sakkalis, K. Marias, I. Karatzanis and M. Zervakis, "In-depth analysis and evaluation of diffusive glioma models", IEEE Trans. Inf. Tech., Vol. 16, No. 3, pp. 299-307, 2012.
4. V. Sakkalis, et al., "Evaluation framework for the multilevel macroscopic models of solid tumor growth in the glioma case", 32nd IEEE- EMBC 2010, Buenos Aires, Argentina, August 31-September 4, 2010.
5. D.D. Dionysiou, et al., Critical Parameters Determining Standard Radiotherapy Treatment Outcome for Glioblastoma Multiforme: A Computer Simulation The Open Biomedical Engineering Journal 2, 43-51, 2008
6. G.S. Stamatakos, et al., "Exploiting Clinical Trial Data Drastically Narrows the Window of Possible Solutions to the Problem of Clinical Adaptation of a Multiscale Cancer Model", PLOS ONE 6(3), e17594, 2011.
7. RL Ho and LT Bartsell, "Biosimulation software is changing research," Biotechnol Annu Rev, Vol. 10, pp. 297-302, 2004.
8. TS Deisboeck et al. "In silico cancer modeling: is it ready for prime time?," Nat Clin Pract Oncol., Vol. 6, pp.34-42, 2009.
9. A. Belloum, E. Deelman, Z. Zhao, "Scientific workflows," Scientific Programming, vol. 14, no. 3-4, p. 171, 2006.
10. G. Fox and D. Gannon, "Special Issue: Workflow in Grid Systems: Editorials," Concurrency and Computation: Practice & Experience, vol. 18, no. 10, pp. 1009–1019, 2006.
11. ISO/IEC 42010:2007. "Systems and software engineering -- Recommended practice for architectural description of software-intensive systems". October 9, 2000.
12. S. Sfakianakis, et al., "Web-Based Authoring and Secure Enactment of Bioinformatics Workflows," Grid and Pervasive Computing Conference (GPC), 2009 pp.88-95, 4-8 May 2009, doi: 10.1109/GPC.2009.14
13. F. Curbera, et al., Unraveling the Web services web: an introduction to SOAP, WSDL, and UDDI, Internet Computing, IEEE 6 (2) (2002) 86–93.
14. N. Shadbolt, W. Hall, T. Berners-Lee, The semantic web revisited, Intelligent Systems, IEEE 21(3) (2006) 96–101.
15. D. Johnson, J. Cooper, and S. McKeever, "TumorML: Concept and Requirements of an In Silico Cancer Modelling Markup Language," Conf Proc IEEE Eng Med Biol Soc. 2011, pp.441-4, 2011.