

# Enabling organized computer simulations with patient cohort data

Kasztelnik, M.<sup>1</sup>, Sosnowski, D.<sup>1</sup>, Nowakowski, P.<sup>1</sup>, Bubak, M.<sup>1,2</sup>

<sup>1</sup>ACC Cyfronet AGH University of Science and Technology, Kraków, Poland

<sup>2</sup>Sano Centre for Computational Medicine, Krakow, Poland

In this paper we present an advanced version of the Model Execution Environment, the outcome of our work in the EurValve project [1, 2] where it was used to build virtual human heart models, enhanced with features appropriate for large-scale multi-model simulations required for investigation of two paediatric cancers [3]. This new environment enables virtual human computational models to be developed in a simple and organized manner, and easily deployed to the available HPC infrastructure.

We have discovered a set of patterns (as well as antipatterns) with regard to how calculations are executed and data is managed on supercomputers by researchers. First of all, researchers try to run models as quickly as possible and focus on results produced by the models, often forgetting about model versioning and traceability of results. Second, data is usually transferred to HPC from a local computer to the user's personal directory, which is not accessible by other team members. Third, once a calculation for a given use case is finished, other calculations are often executed in succession, which leads to result data being overwritten. It is therefore difficult, or even impossible, for other researchers to recreate the experimental setup.

The Model Execution Environment address these issues and thus simplify virtual human creation by promoting the following principles:

- Repeatable runs achieved through integration with git repositories. Each run record model stores a version (git SHA) which can be used to run the same calculation again (either by the same or by another user).
- Integration with HPC. To meet this goal, MEE is integrated with the PLGrid [4] infrastructure which allows us to delegate user rights from MEE to PLGrid supercomputers. As a result, the modeler can easily use a web-based interface to start the simulation without

deep knowledge of HPC technologies.

- An organized way to store patient data and calculation results. The main unit of data in the MEE environment is a single patient. Each patient has a dedicated storage space on HPC, where their data is stored. Inside, we also allocate space for calculation-specific inputs and results. Storage is secured in compliance with the POSIX standard and is only accessible to a restricted group of users.

This environment enables the development of new models, algorithms, and computational methods that may make medical treatment more efficient and shorter.

During the talk we will present the main building blocks of the Model Execution Environment, and explain how they promote virtual human creation with regard to repeatability, replicability and reproducibility.

**Acknowledgments.** This work was supported by the EU project PRIMAGE H2020-EU.3.1.5.3 826494 and by the PL-Grid Infrastructure ([www.plgrid.pl](http://www.plgrid.pl))

## References

- [1] The EurValve Project website, <http://www.eurvalve.eu/>
- [2] M Bubak, K Czechowicz, T Gubała, DR Hose, M Kasztelnik, M Malawski, J Meizner, P Nowakowski, S Wood The EurValve model execution environment, Interface Focus 11 (1), 20200006, 2021
- [3] The PRIMAGE Project website, <https://www.primageproject.eu/>
- [4] PLGrid Infrastructure website, <http://www.plgrid.pl/en>