AGI 1: Hacky Hour I (joint session AGI/AKjDPG)

Time: Wednesday 14:00–15:45 Location: AGI-H20

Invited Talk AGI 1.1 Wed 14:00 AGI-H20

Practical semantic data management with CaosDB — •Alexander Schlemmer[1,2,3], Ulrich Parlitz[1,3,4], and Stefan Luther[1,3,5] — [1]Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany — [2]IndiScale GmbH, Göttingen — [3]German Center for Cardiovascular Research (DZHK), Partner Site Göttingen — [4]Institute for the Dynamics of Complex Systems, University of Göttingen — [5]Institute of Pharmacology and Toxicology, University Medical Center Göttingen

In practice, scientific data management comprises many different tasks and workflows that are typically accompanied by software in varying degrees. It is a common issue to find the right balance between standardization and flexibility, automation and interactivity, complexity and comprehensibility.

CaosDB is an Open Source (AGPLv3) research data management system (RDMS) that combines multiple data management concepts and practical tools for efficiently integrating daily research data management into scientific workflows. Especially noteworthy are the flexible semantic data model, the intuitive semantic query language CQL and the file crawler framework for automatic data integration.

In this talk the software and the central concepts will be discussed presenting use cases from daily scientific research. A practical introduction to the graphical user interface, the query language, the API and the crawler framework will be given to demonstrate how these concepts can facilitate data management and provide a deeper insight into complex and heterogeneous research data.

AGI 1.2 Wed 14:45 AGI-H20

Snakemake: Making data workflows easier and more reproducible — •Johannes Hampp — Center for international Development and Environmental Research, Justus Liebig University Giessen

Daily scientific work often involves handling research data from experiments or simulations. Necessary data wrangling and analysis steps are usually repeated following predefined steps. Snakemake aims to make this process easier, faster, less error-prone, improving transparency and reproducibility. Individual steps are split into standalone rules, which are flexibly combined into workflows. Workflows are defined in a simple and human-readable format. They are automatically executed to keep any data dependencies up-to-date. Snakemake thus ensures ordered, transparent and documented data workflows, significantly reducing human errors from manual workflow execution or from improvised, self-written workflow solutions. Snakemake is open source software and supports popular programming languages like R, Python and Julia. Furthermore, integration with other programming languages or programmes is possible as long as they offer a command line interface. Many more features are available.

For yourself, Snakemake makes your life easier, more productive and more fun. For other researchers, well-documented and automatic workflows increase the accessibility and reproducibility of your research and research data.

AGI 1.3 Wed 15:15 AGI-H20

Controling laboratory equipment using Python and `pylablib` — •Alexey Shkarin — Max Planck Institute for the Science of Light, 91058 Erlangen, Germany

As experiments become progressively more complicated and generate more data, there is a need for automation of the equipment control and data acquisition. This often requires orchestrated control of multiple devices, which demands custom experiment-specific software. For a long time LabView has been a de-facto standard in this domain, but over the last decade Python has been gaining more traction due to its universaility, simplicity, and its already wide support on the data processing side.

In this talk I will introduce basics of device control, specifically focusing on the Python libraries which are most useful in these tasks. Then I will present `pylablib`, a software package dedicated to control of specific devices. Finally, I will show how `pylablib` can be used in a couple of simple examples where several devices need to be controlled at the same time.