# eFlows4HPC

Enabling dynamic and Intelligent workflows
in the future EuroHPC ecosystem

# An end-to-end workflow for climate data management and analysis integrating HPC, Big Data and Machine Learning

A. D'Anca[1], S. Scardigno[1], J. Ejarque[3], G. Accarino[1,2], D. Peano[1], F. Immorlano[1,2],

D. Donno[1], E. Scoccimarro[1], R. M. Badia[3], G. Aloisio[1]

[1] Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Lecce, Italy, [2] Università del Salento, Dept. of Engineering for Innovation, Lecce, Italy, [3] Barcelona Supercomputing Center (BSC), Barcelona, Spain

## eFlows4HPC PROJECT

**eFlows4HPC** provides a **software stack** that simplifies the development, deployment, orchestration and execution of complex end-to-end workflows combining **HPC** simulation and modeling with **artificial intelligence** and **data analytics**. It develops the concept of **HPC Workflows as a Service (HPCWaaS)** to facilitate the use and reuse of the workflows across HPC infrastructures.
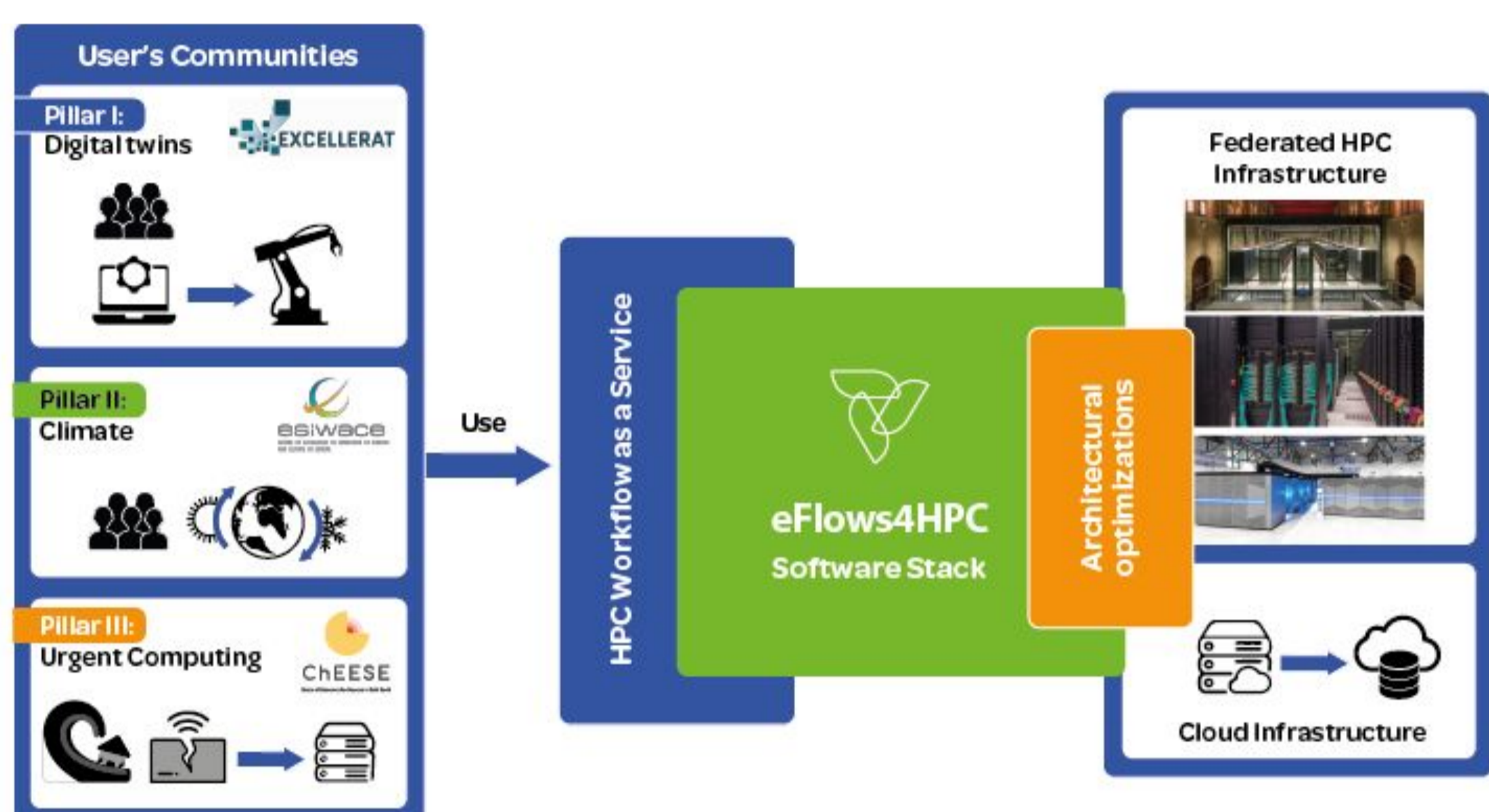


## PILLAR II USE CASE

**Earth System Models (ESM) simulations** represent one of the most challenging HPC use cases due to the very high computational cost, intensive Input/Output patterns, very large data volumes produced, and the necessity of post-processing them to distill knowledge and extract relevant information.
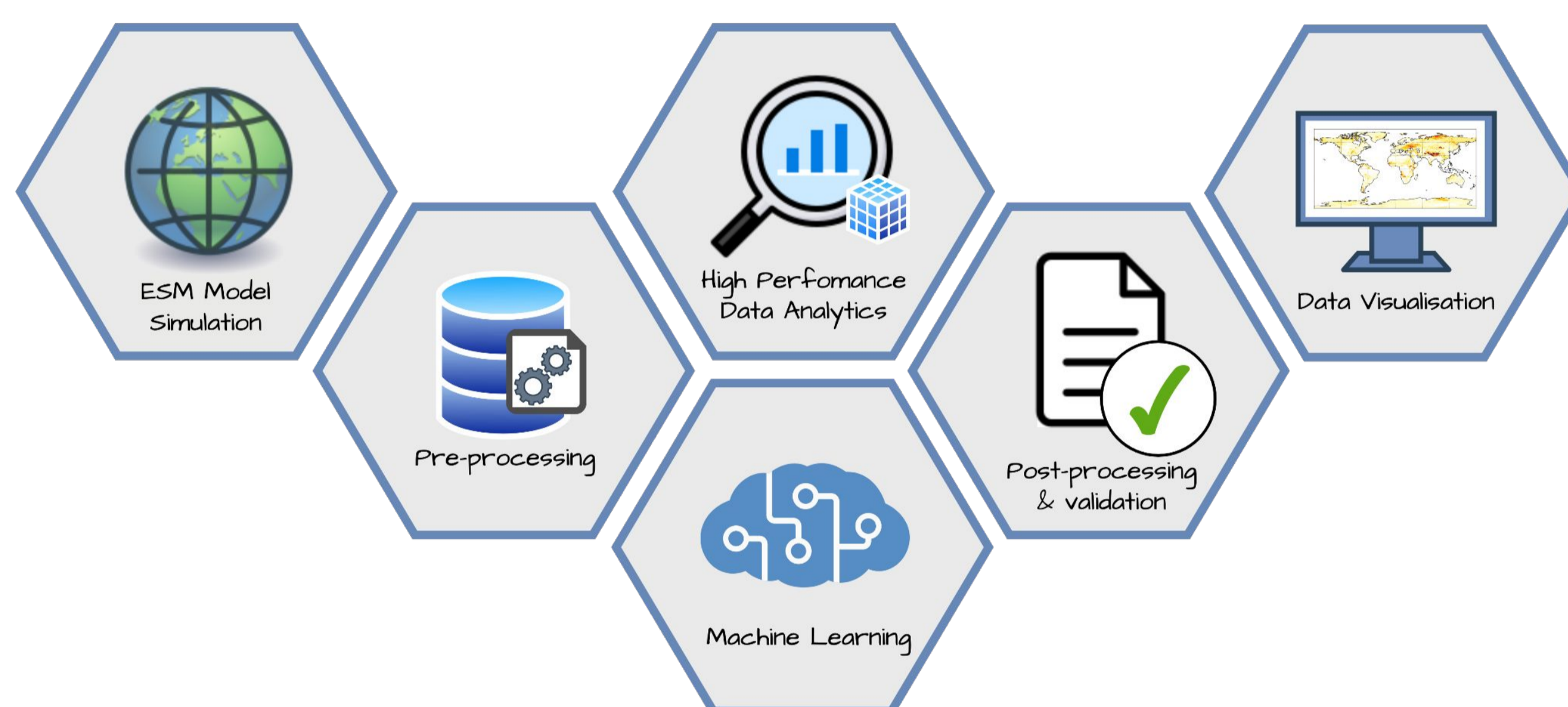


## FEATURE EXTRACTION WORKFLOW

Our work targets ESM workflows, integrating multiple components in a transparent and efficient way: the CMCC-CM3 model execution, post-processing tasks, **HPDA** and **ML** techniques. **PyCOMPSs** is exploited for the management of parallel pipelines, **PyOphidia** for climate data analytics and ML frameworks for data-driven models. Through this, users can run complex end-to-end workflows in a seamless way on both Cloud and HPC.
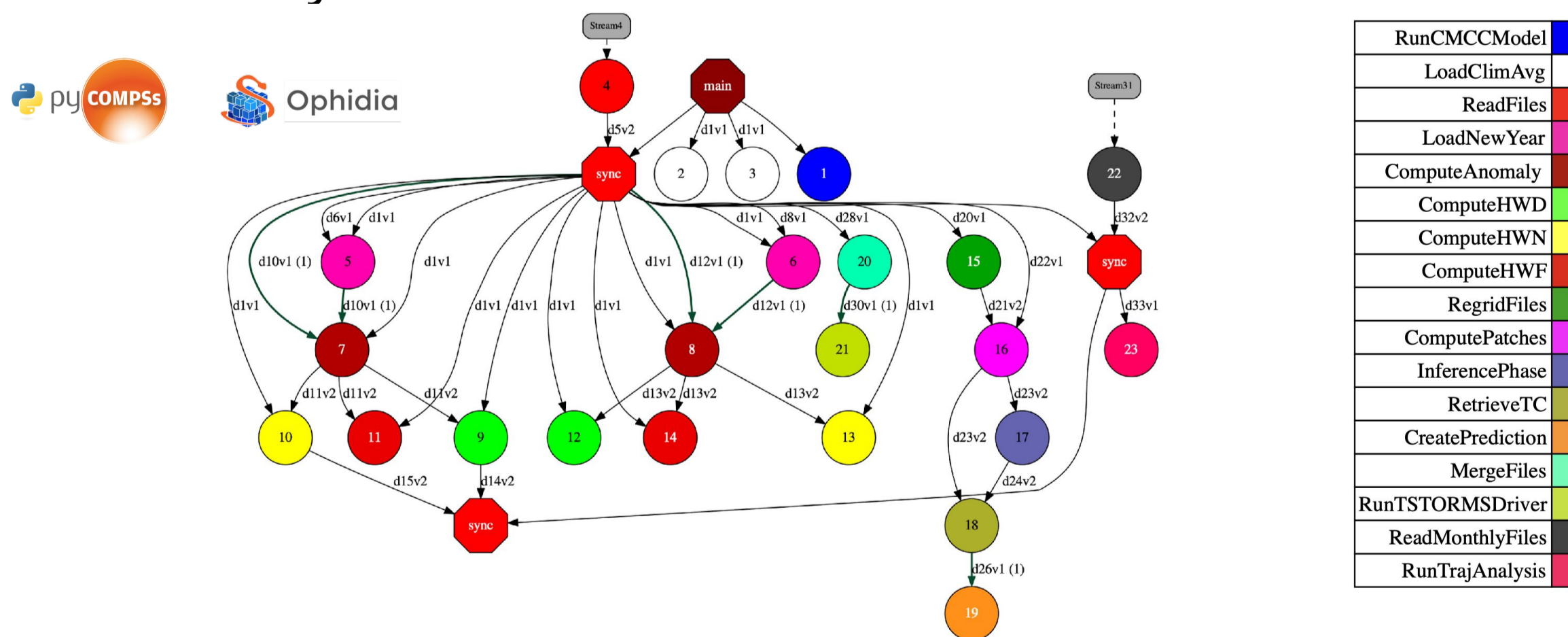


## EXTREME EVENTS COMPUTATION

**Key results** of this integration:
- automating the creation of added-value products from the output of the ESMs simulation.
- allowing scientists to increase their knowledge of the climate system through novel techniques.
- Evaluation of data-intensive vs. data-driven approaches for TC analysis .

| | |
|---|---|
| HWD/CWD | Heat/Cold Wave Duration |
| HWN/CWN | Heat/Cold Wave Number |
| HWF/CWF | Heat/Cold Wave Frequency |
| TC Detection | Detection of Tropical cyclone eye |

**www.eFlows4HPC.eu**

@eFlows4HPC    eFlows4HPC Project