



The CYBELE Case: On Making Agriculture Data Smarter and More Efficient



CYBELE

Petrou Petros
Senior SW Architect



Learn more here:
<https://www.cybele-project.eu/>



Technical & Industrial Challenges

- Large volumes of data requiring diverse and online computing modalities for collection, processing and further analysis;
 - Need for global and local learning preserving privacy and contributing to advanced decision making at strategic level;
 - Need for distributed processing and speed up of time demanding simulations, model training, complex computations, etc.;
- On data analysis, complex, multimodal and dynamic workflows require intelligent mechanisms for transformation, aggregation and training pressing for convergence of the Big Data and HPC worlds (distributed vs. parallel computing);
 - Currently these two worlds seem decoupled;
 - There is no ease of use for end users; a huge and diverse knowledge gap needs to be bridged;
- When data are loaded & stored at the testbeds require efficient and distributed data services (curation, anonymization, enrichment);
- Voluminous analysis results require adaptable and non-blocking visualization services.





Value Proposition: To Innovate, Design Solutions, Integrate & Boost Interoperability for Advanced Decision Making in PA

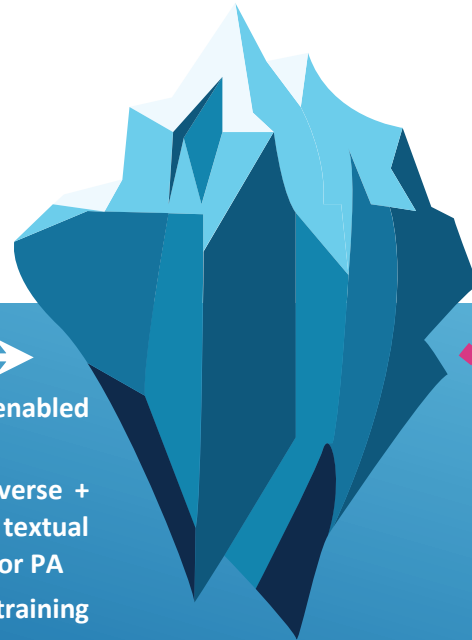
- Efficient and optimized Big Data and HPC e-infrastructure for seamless and voluminous data operations, parallel and distributed computing tasks
- **Our Tech Journey:** High-tech convergence for various optimizations on real-world cases
 - Enriched Data Services Layer & Secure Data Operations for several use cases in Precision Agriculture
 - Hybrid Big Data and HPC capabilities through a Parallel and Distributed Management Layer
 - Embedded Workflow Management and Experiments Composition Layer
 - Reusable AI Models Library
 - Enriched Visualization and Reporting Layer



Challenges & Outcomes

Challenges

- Development of secure and large-scale HPC-enabled industrial pilot (HPC/BD/Cloud) test-beds
- Efficient management and aggregation of diverse + large satellite, imagery, time-series and textual datasets for the creation of new applications for PA
- Time-consuming simulations and AI models training requiring days/weeks to be performed



Outcomes

- Advanced Big Data and HPC e-infrastructure Testbeds to efficiently process data and execute distributed/parallelized algorithms
- Optimized and secure Data Services and algorithms for distributed and parallel execution targeted at PA/PLF
- User-defined + Automatic Analytic Applications and Simulations configuration, resource allocation and execution in hybrid Big Data – HPC ecosystems





How HPC, Big Data & AI co-exist in CYBELE

- HPC, Big Data & AI convergence lies at several PA use cases
 - **Demo 1** (organic soya yield and protein-content prediction; overlaid maps): tasks parallelization / execution speed up;
 - **Demo 2** (Spark-fueled food safety & risk assessment);
 - **Demo 3** (climate services for organic fruit production): distribution / parallelization over HPC partition;
 - **Demo 4** (Spark-map / MPI+Python in Slurm autonomous robotic systems within arable frameworks);
 - **Demo 5** (Spark-based optimized computations for crop yield forecasting);
 - ...and more coming from the Livestock Farming domain.



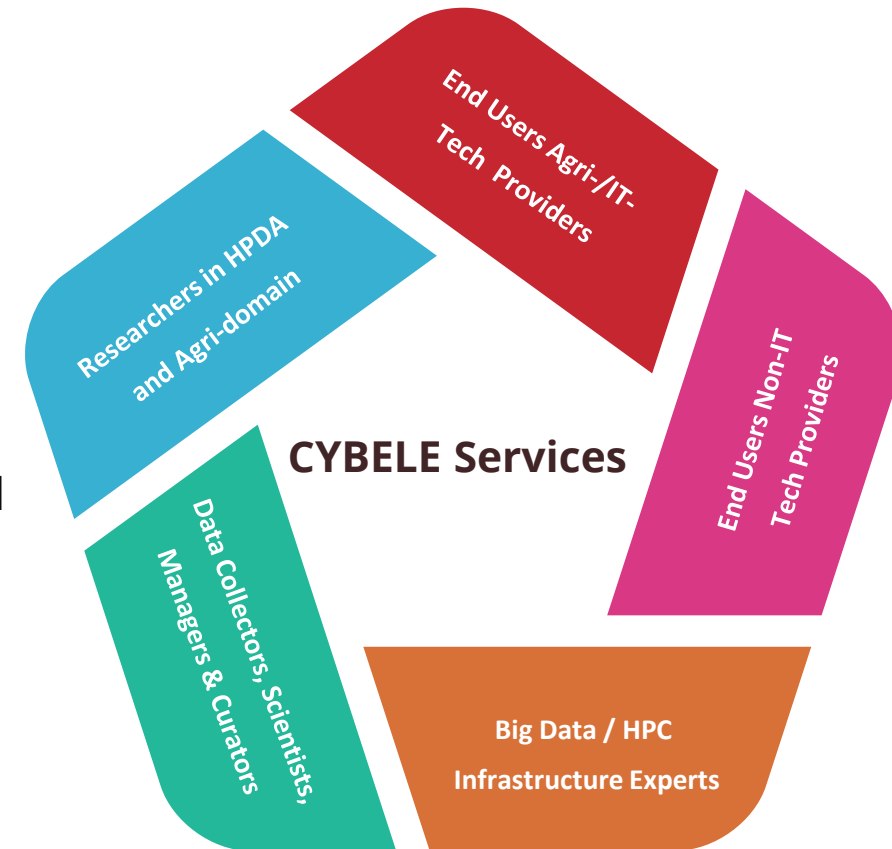
How CYBELE helps

- CYBELE Components are being used to help SMEs, Agri-tech companies and SW providers bring new products to market that make predictions, classify food and predict its quality for agriculture use cases;
- CYBELE Components increase PA business impact;
- Agri-tech companies now experience:
 - ✓ *Saving costs related to deploying ML training in HPC resources;*
 - ✓ Savings related to easily customizing ML templates;
 - ✓ Faster time to market for new products with analytic capabilities;
 - ✓ Higher accuracy in predictions, classifications, etc.
- ALL Demo partners are facilitated by fast ML/DL models training and / or simulations on HPC-BD tasks.



How CYBELE helps

- **Impact & Actual Exploitation**
 - Demonstrated effectiveness of new technologies
 - Affordable technological solutions
 - Involve end-users in the development procedure improving automation-assisted decision support
 - Stimulate business development & services in the agriculture supply chain
 - Business and financial value from the adoption of HPC and Big Data tools and services
 - Increase CYBELE social reputation & accountability → Participation in large venues, industrial events & BDVA workshops
 - One joint patent; One new product- FOODAKAI; New elastic DB - LXS capabilities; CYBELE Semantic Model enriched with OGC Definition Server





Thank you!

Petrou Petros
Senior SW Architect
ppetrou@ubitech.eu

