interTwin

An Interdisciplinary Digital Twin Engine for Science

CERN-IT: Matteo Bunino, Alexander Zoechbauer, Kalliopi Tsolaki, Ilaria Luise, Maria Girone, Sofia Vallecorsa, Alberto Di Meglio

JSC: Rakesh Sarma, Mario Ruettgers, Andreas Lintermann

interTwin consortium

Funded by the European Union

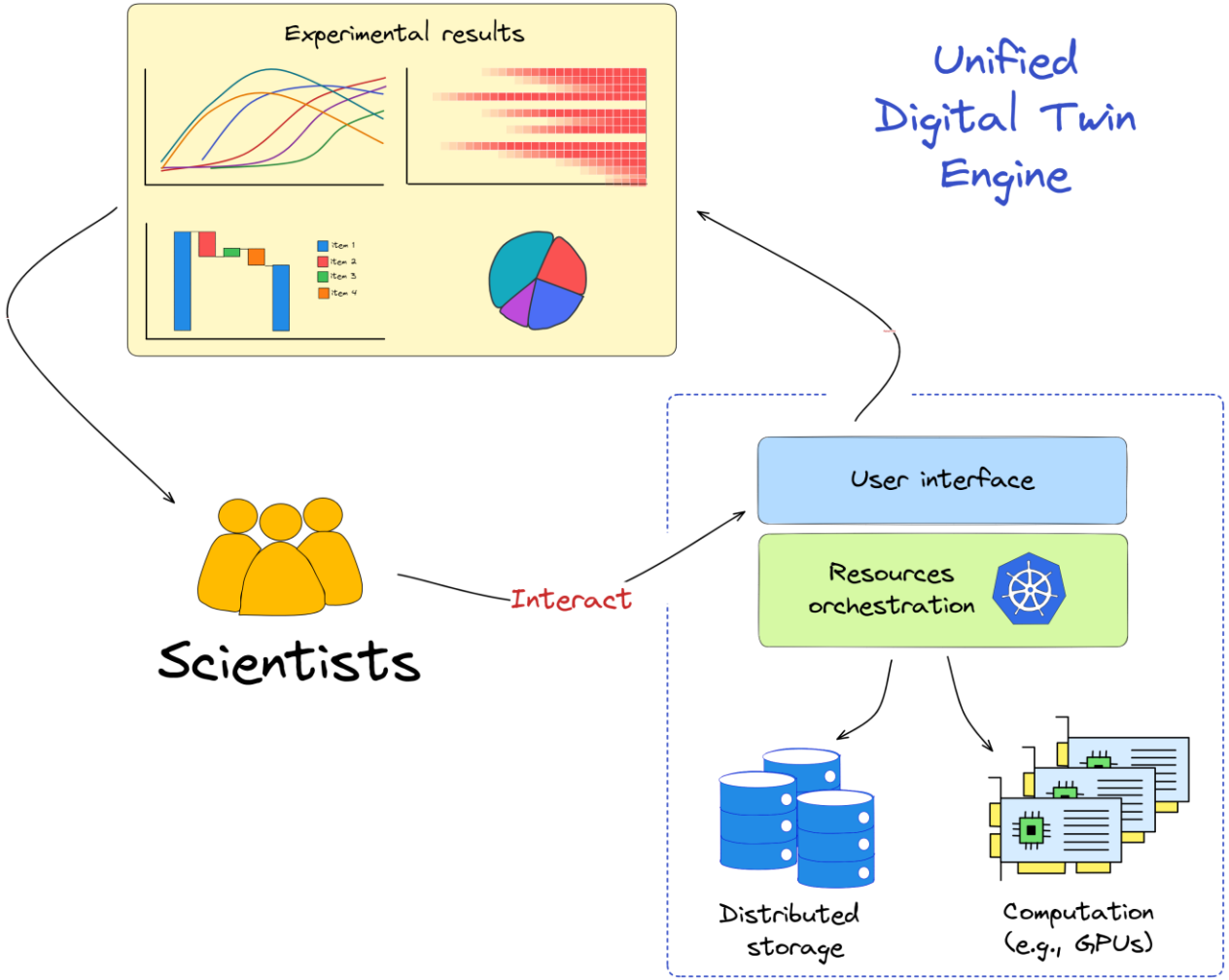
The interTwin project is funded by the European Union - Grant Agreement Number 101058386

interTwin: a Digital Twin Engine for science

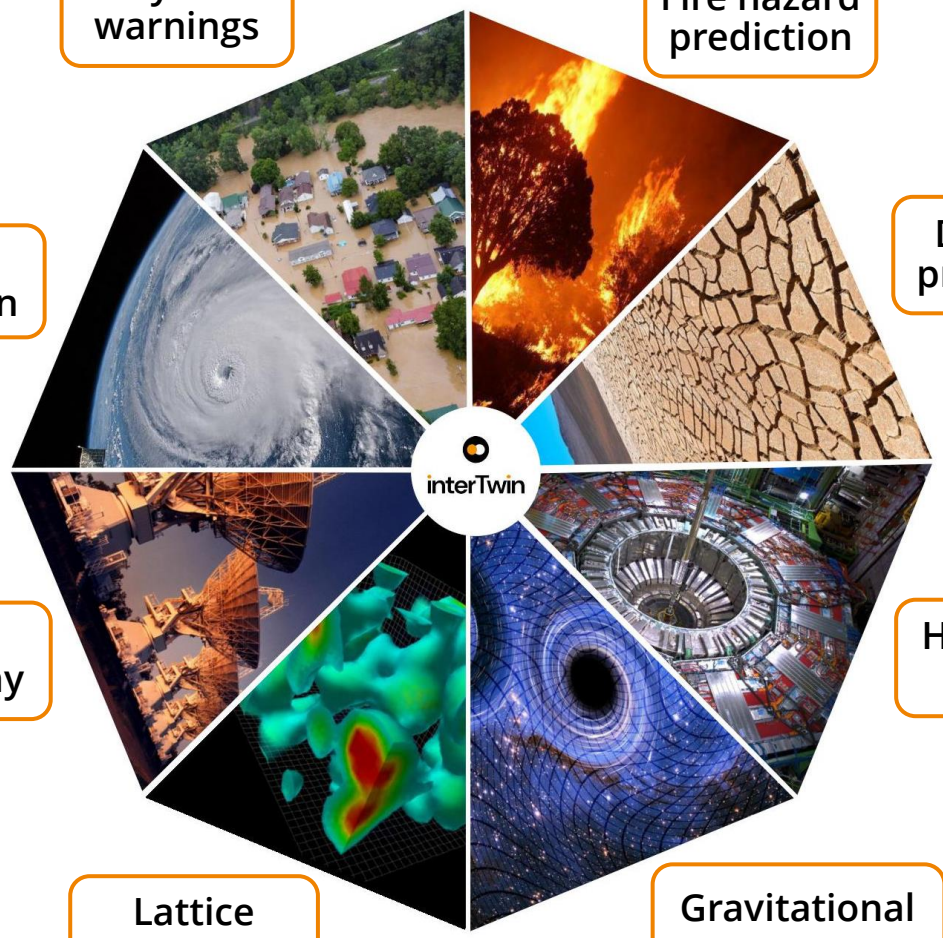
Co-design and implement the prototype of an **interdisciplinary Digital Twin Engine**

Open-source platform based on **open standards**

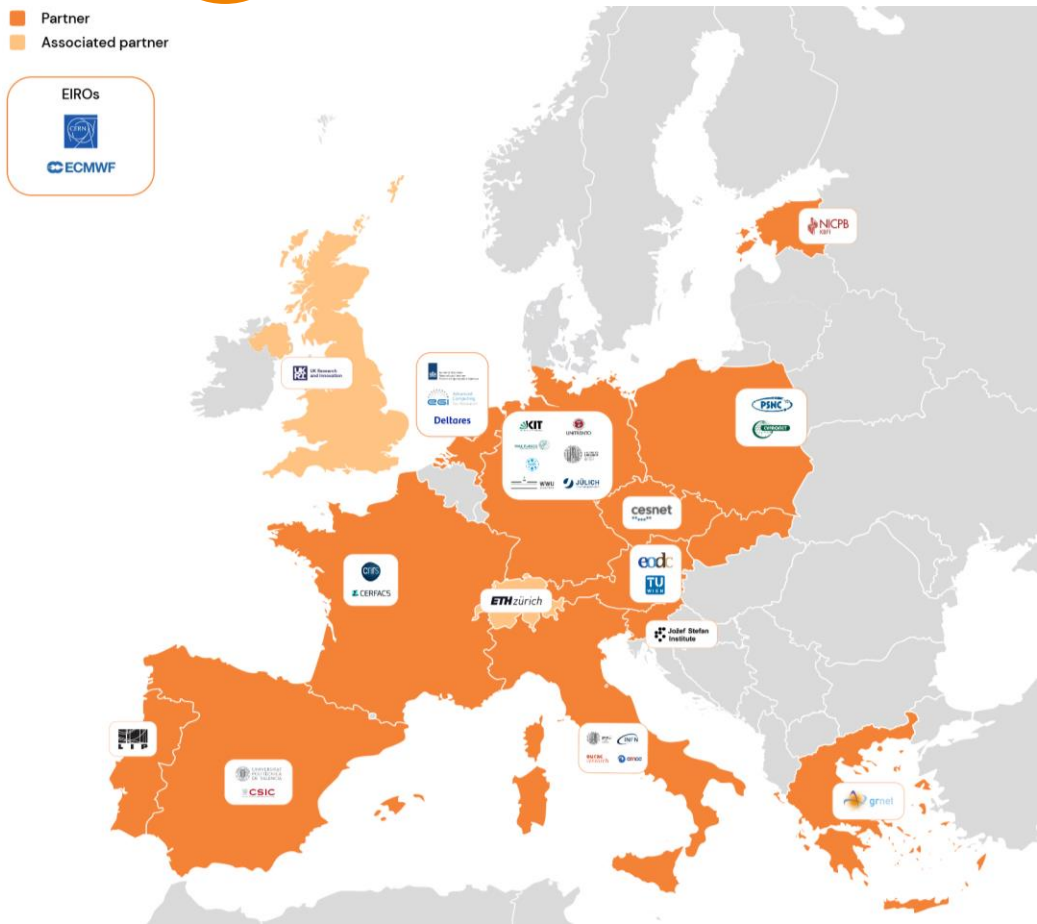
Large spectrum of **diverse use cases** from **physics** and **earth observation** sciences



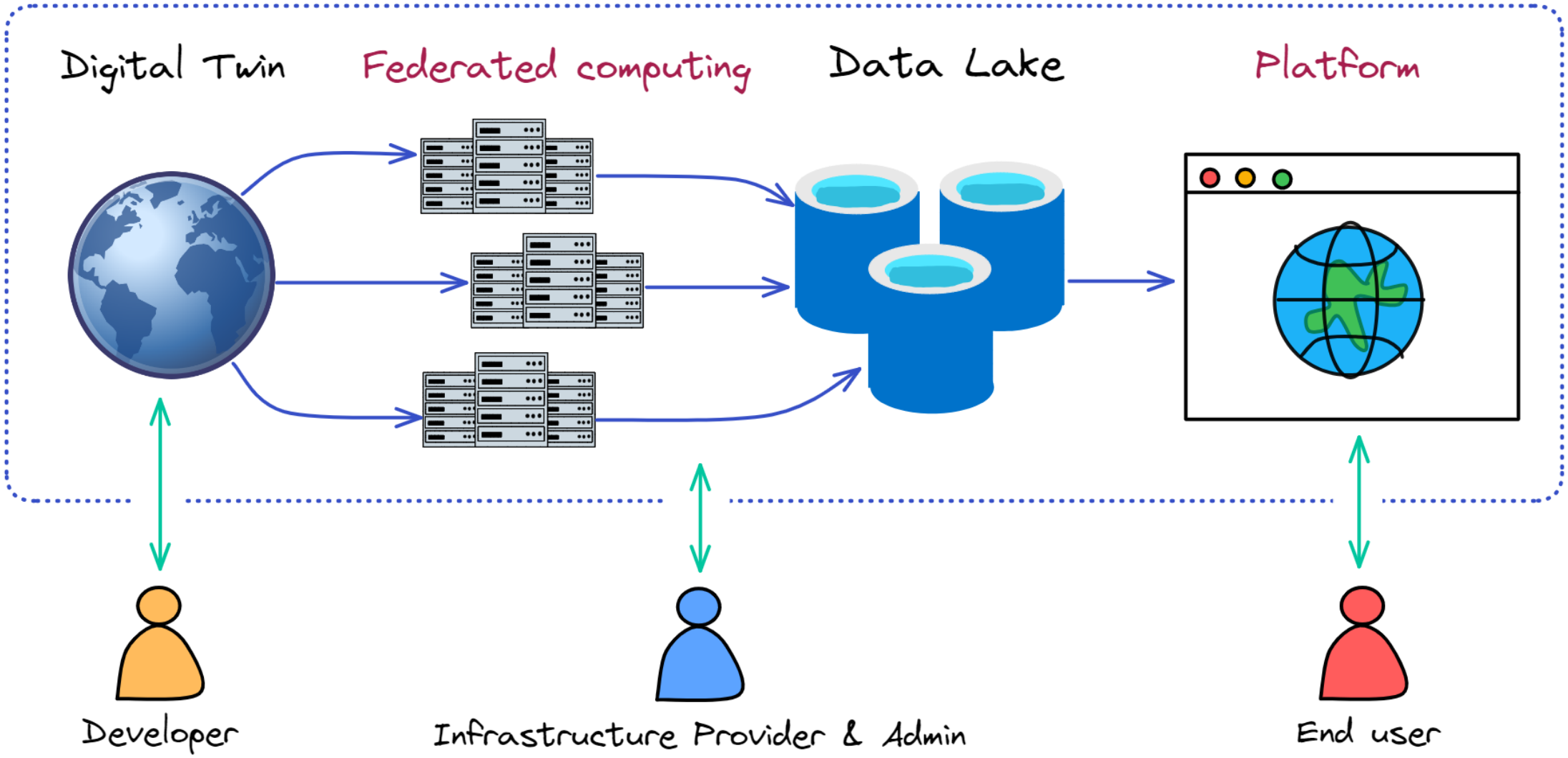
interTwin: a Digital Twin Engine for science

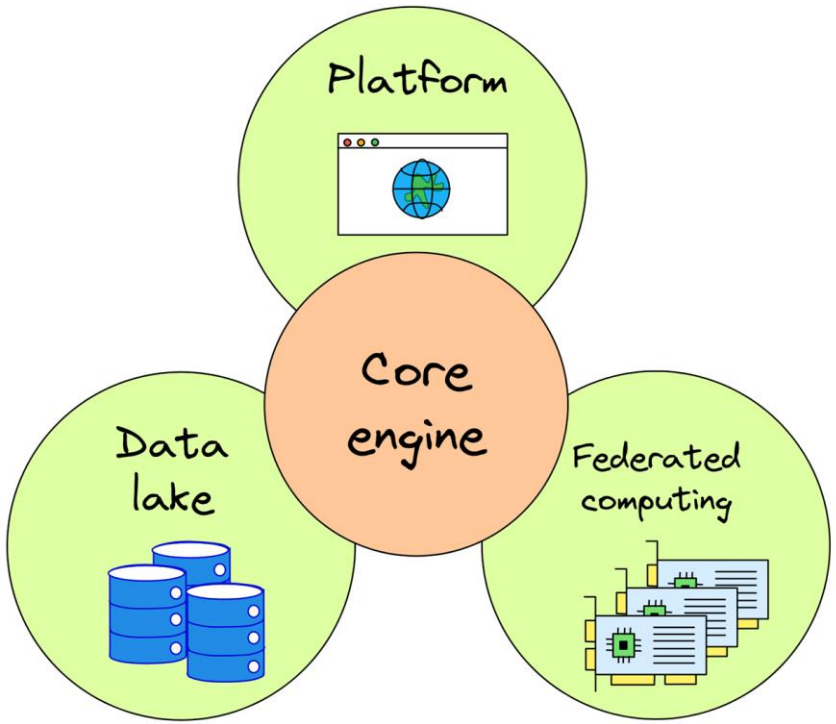


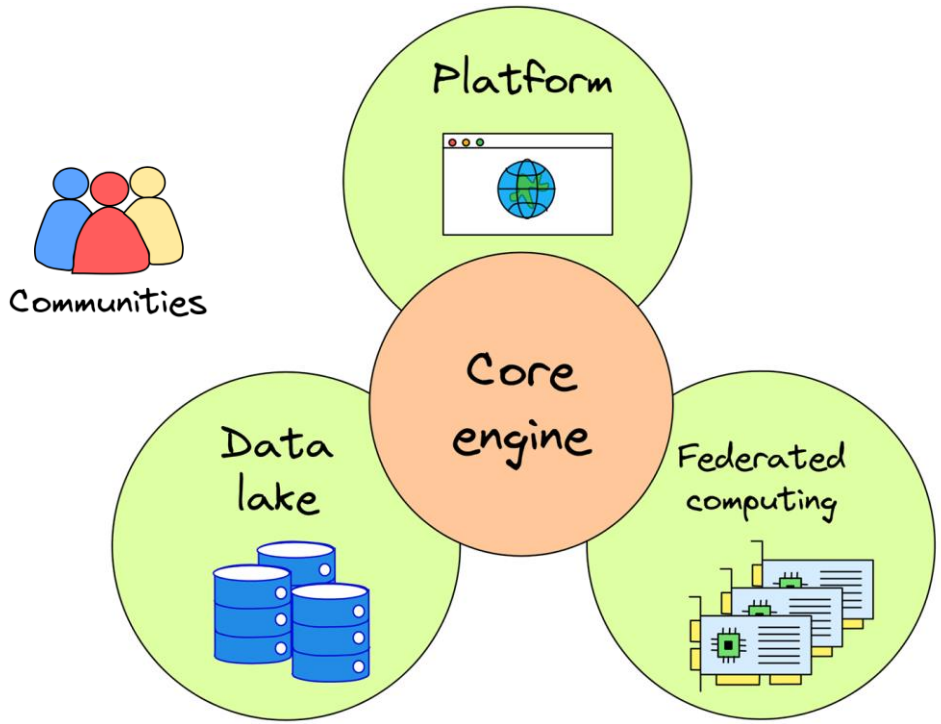
29 **Participants,** including sciences, technology and resource providers.

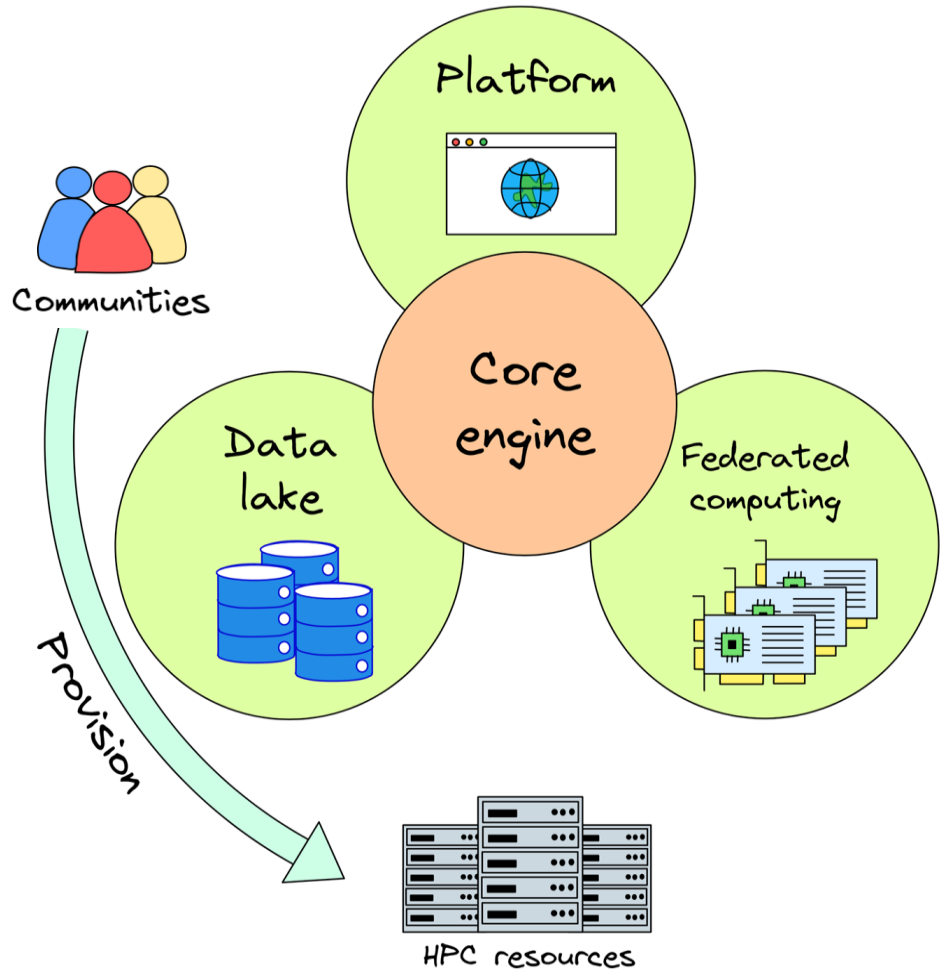


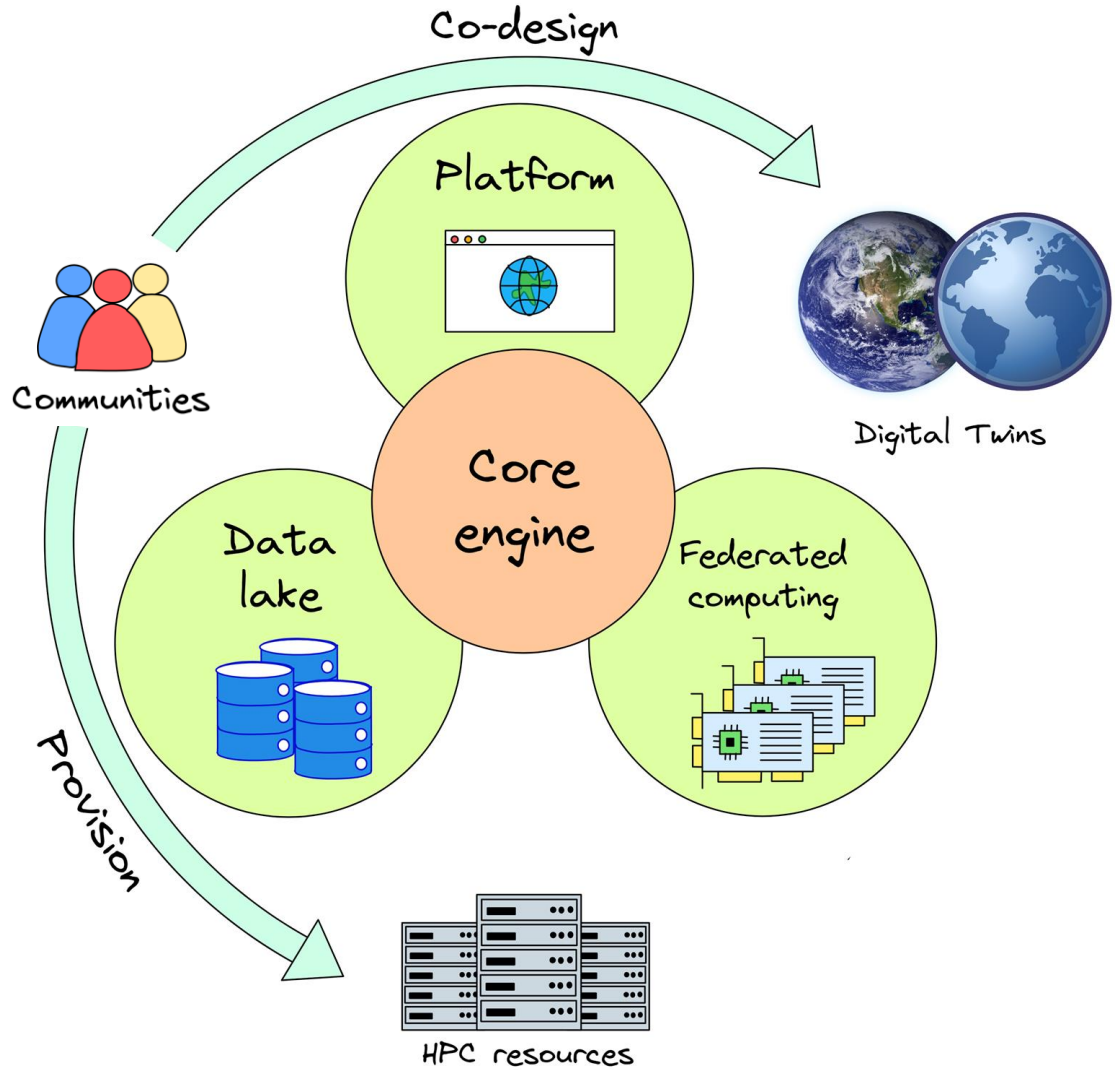
interTwin: a Digital Twin Engine for science

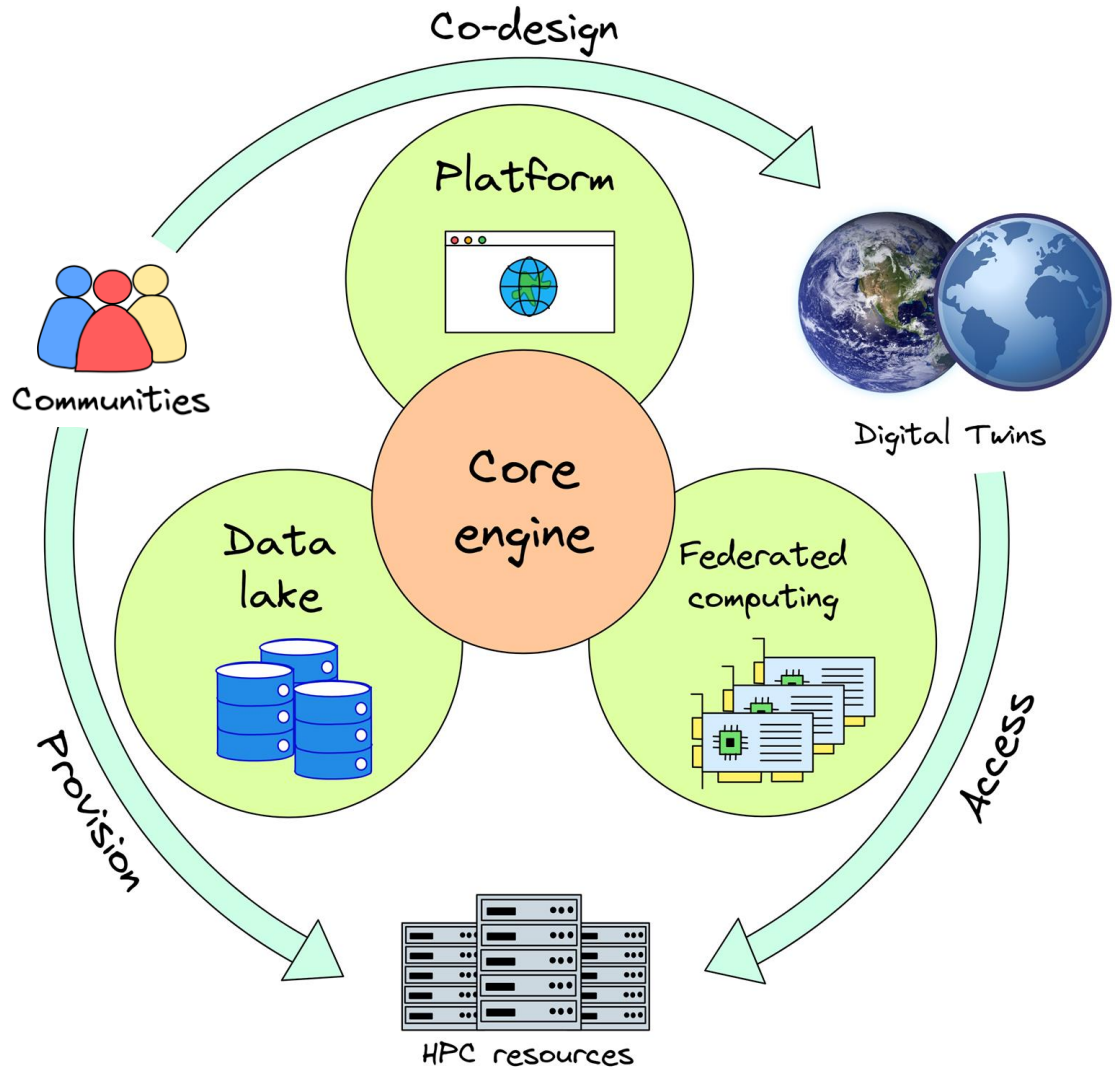



Platform
Core engine
Data lake
Federated computing


Platform
Communities
Core engine
Data lake
Federated computing


Platform
Communities
Core engine
Data lake
Federated computing
Provision
HPC resources


Co-design
Platform
Communities
Digital Twins
Core engine
Data lake
Federated computing
Provision
HPC resources


Co-design
Platform
Communities
Digital Twins
Core engine
Data lake
Federated computing
Provision
Access
HPC resources

Climate research and monitoring

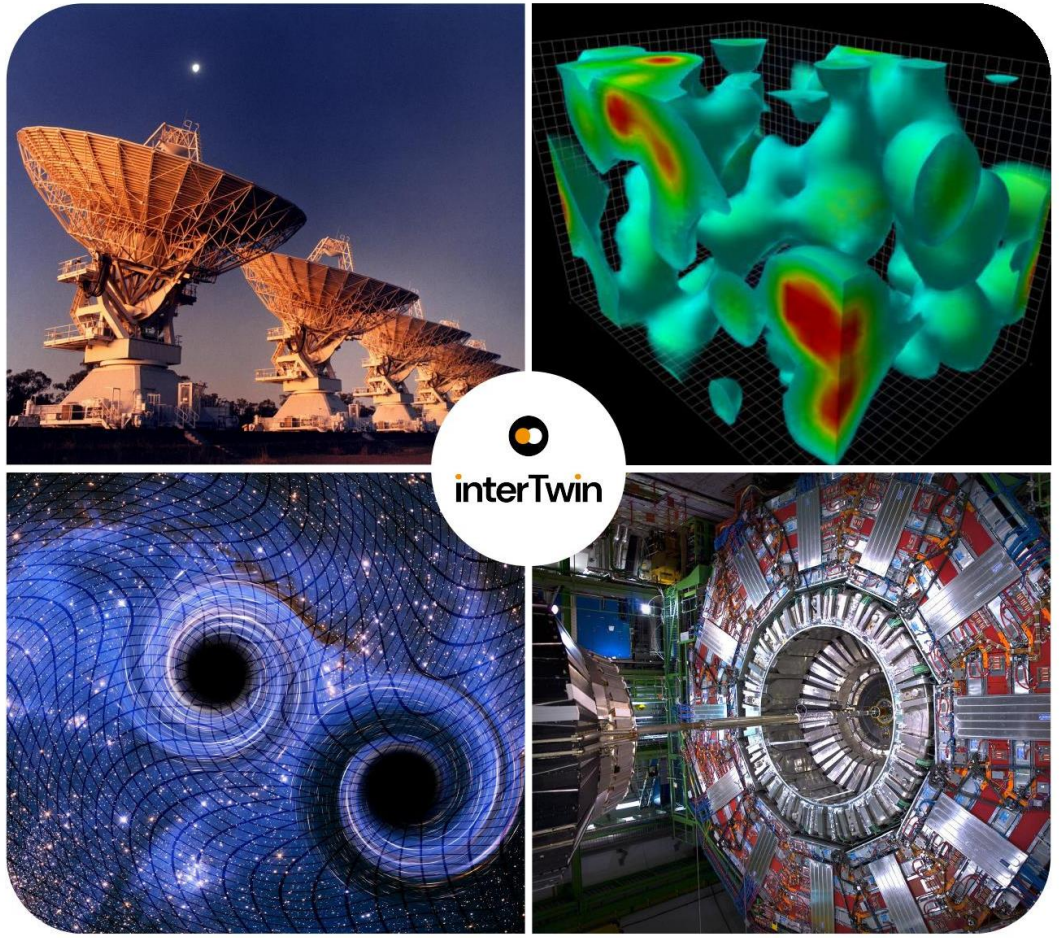
Climate Change Future Projections of Extreme Events (storms & fire)
CMCC, CNRS (IPSL), UNITN

Climate Change Impacts of Extreme Events (storms, fire, floods, drought)
CERFACS, Deltares, EURAC

Early Warning for Extreme Events (floods & droughts)
Deltares, EURAC, TU Wien

Physics research

Noise Simulation for Radio Astronomy
Univ. of Heidelberg, Max Planck Society



Lattice QCD Simulations for High Energy Physics
ETHZ, CSIC

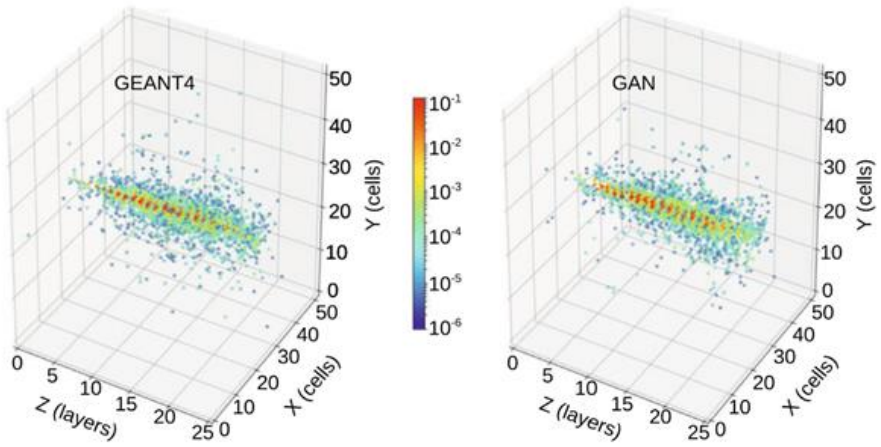
VIRGO Noise Detector for Astrophysics
INFN

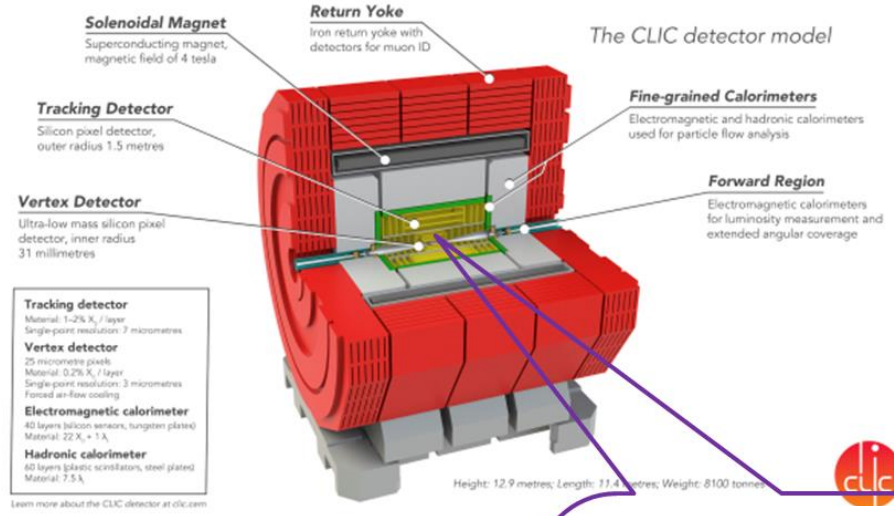
Detector Simulation for High Energy Physics
CERN, CNRS (IN2P3-IRES)

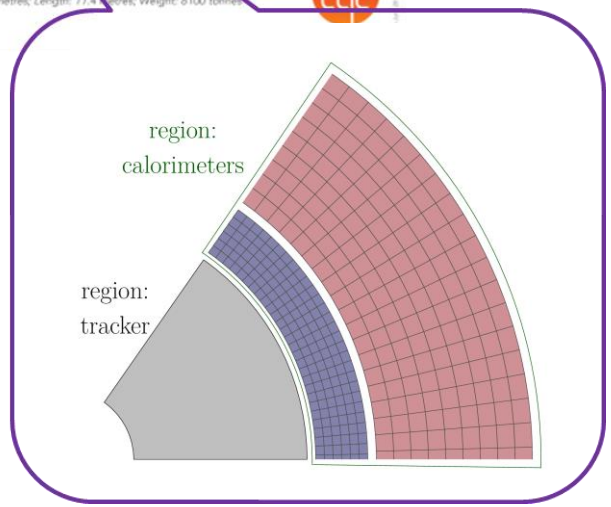
HEP Digital Twins of particle detectors

Detector Prototyping & Optimization
Build data-driven tool that **simulates detector response** and integrates operation conditions from experimental setups (test-beams).

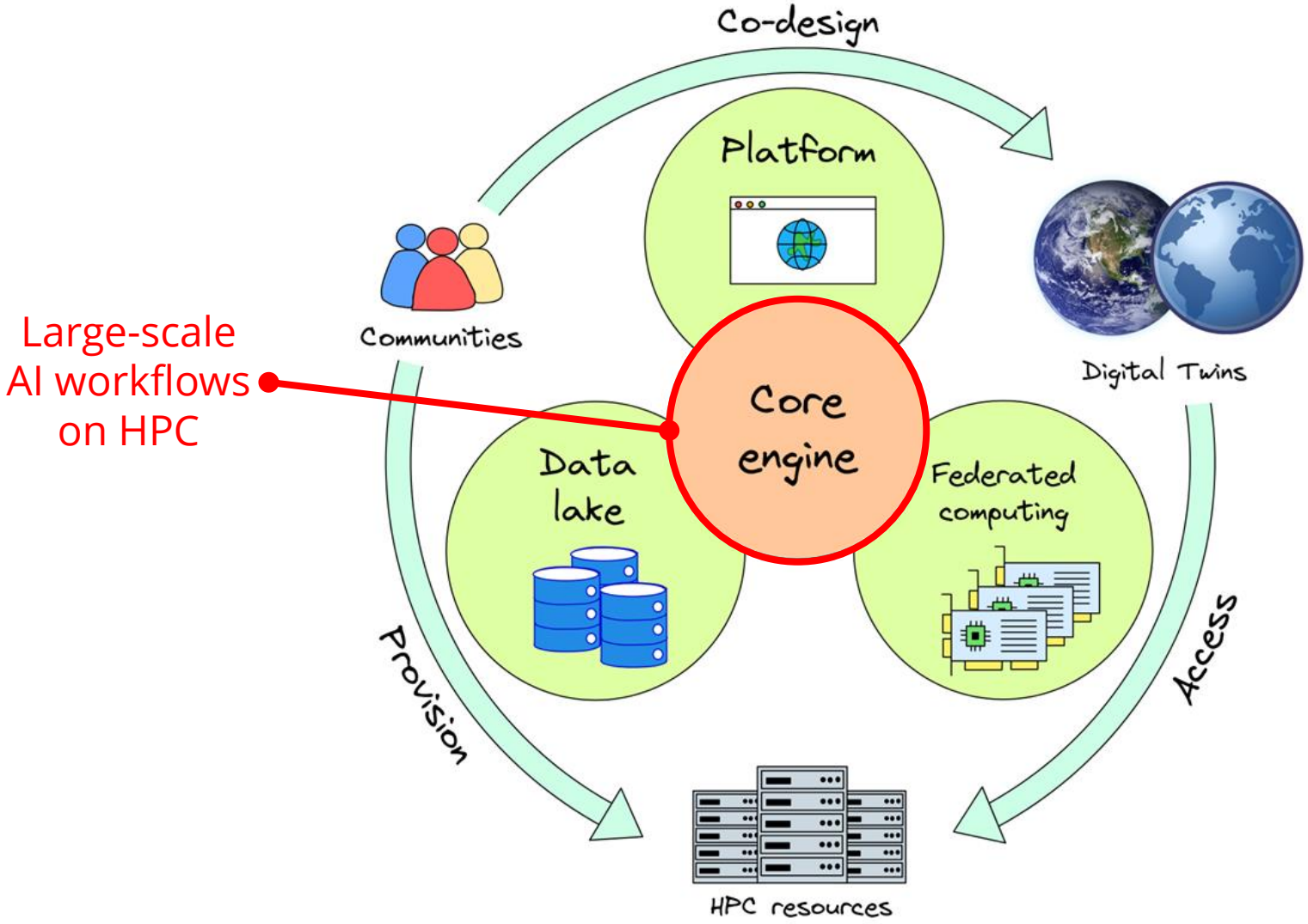
Online ML for Detectors
Adapt **real-time** detector and/or data acquisition configuration with respect to run conditions







Fast Simulation of a High Granularity Calorimeter by Generative Adversarial Networks.
Gul Rukh Khattak et al. https://arxiv.org/abs/2109.07388 DOI: https://doi.org/10.48550/arXiv.2109.07388

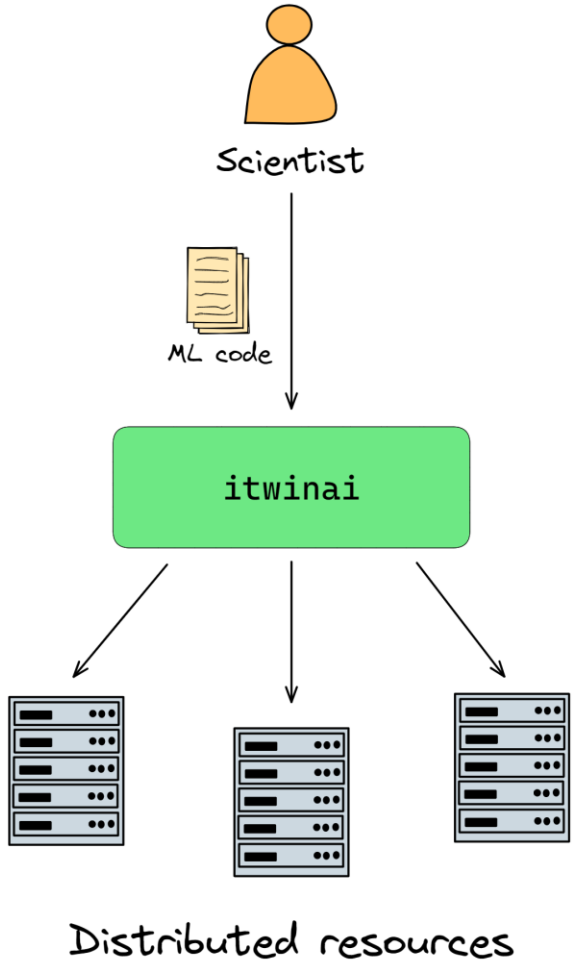


DT core engine – AI features

Our contribution to the Core Engine

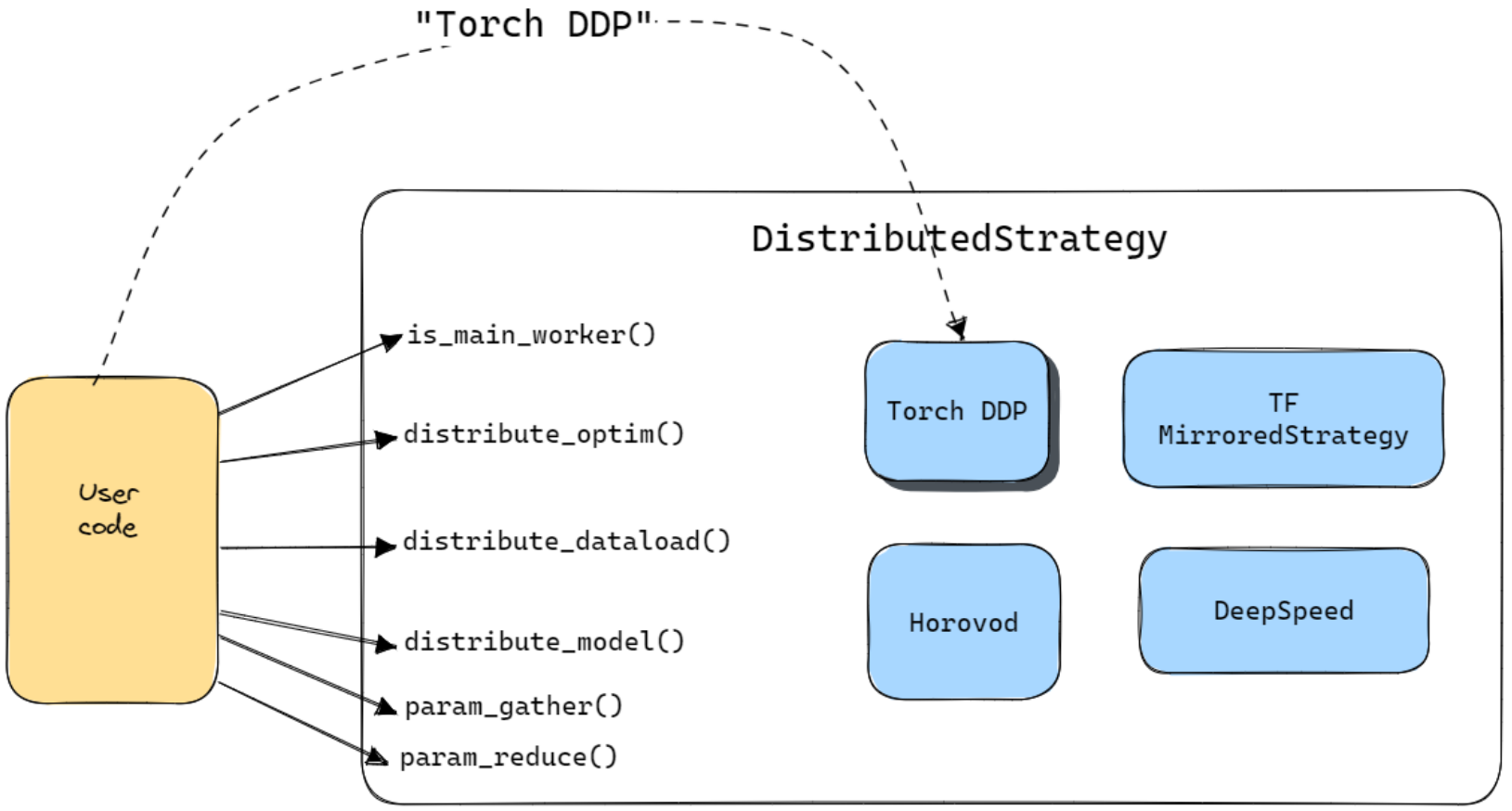
Support AI-based digital twin applications in science:

- **Reproducibility, Reusability, and Modularity**
- **Framework-independent** (e.g., PyTorch, TensorFlow, MLFlow, WandB)
- **UX/UI**: user-friendly GUI (e.g., JupyterLab)
- Off-the-shelf AI tooling, reducing engineering overheads:
  - **Hyper-parameters optimization**
  - Scalability (e.g., **distributed ML**)
  - State of the Art **models repository**
- **Seamless access to infrastructure** (cloud and HPC)



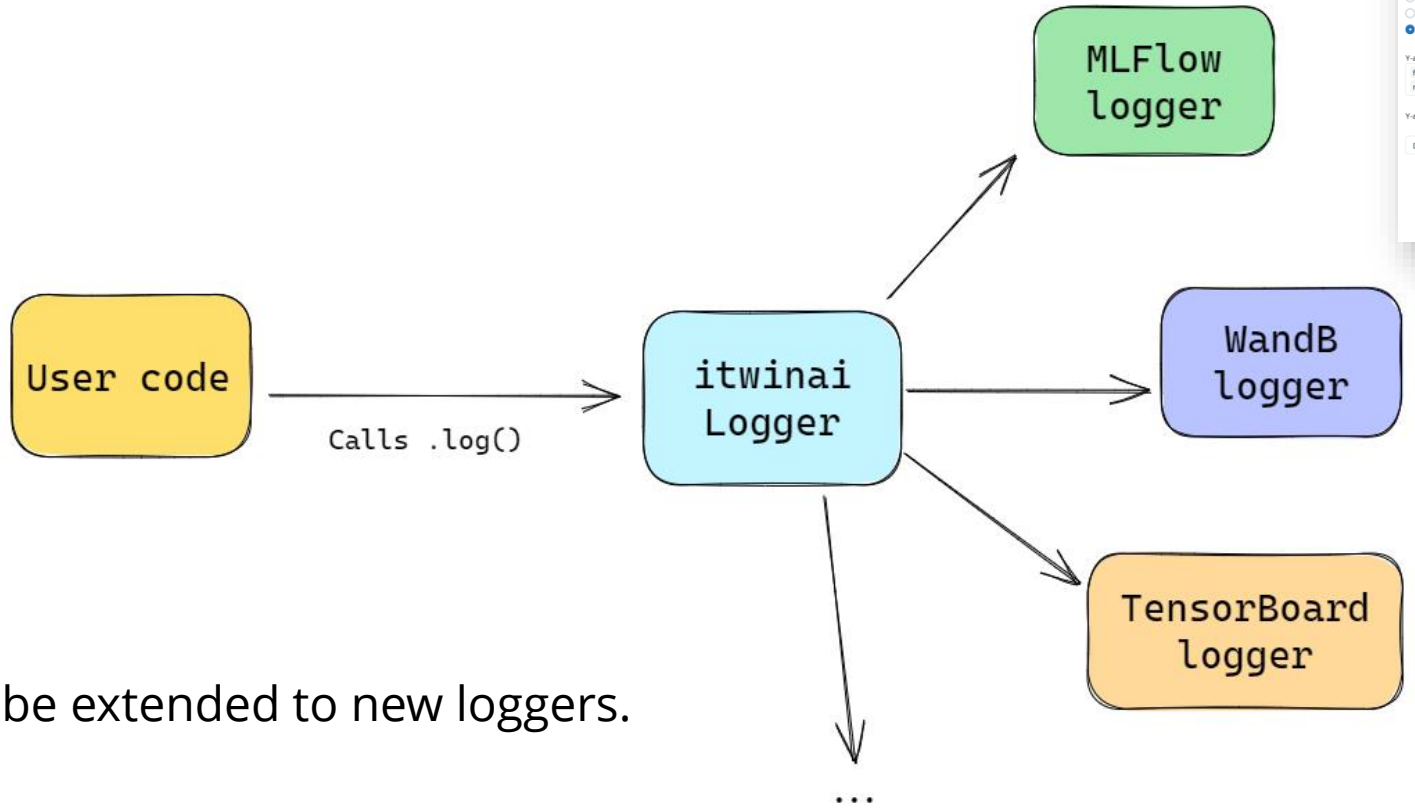
Our contribution to the Core Engine

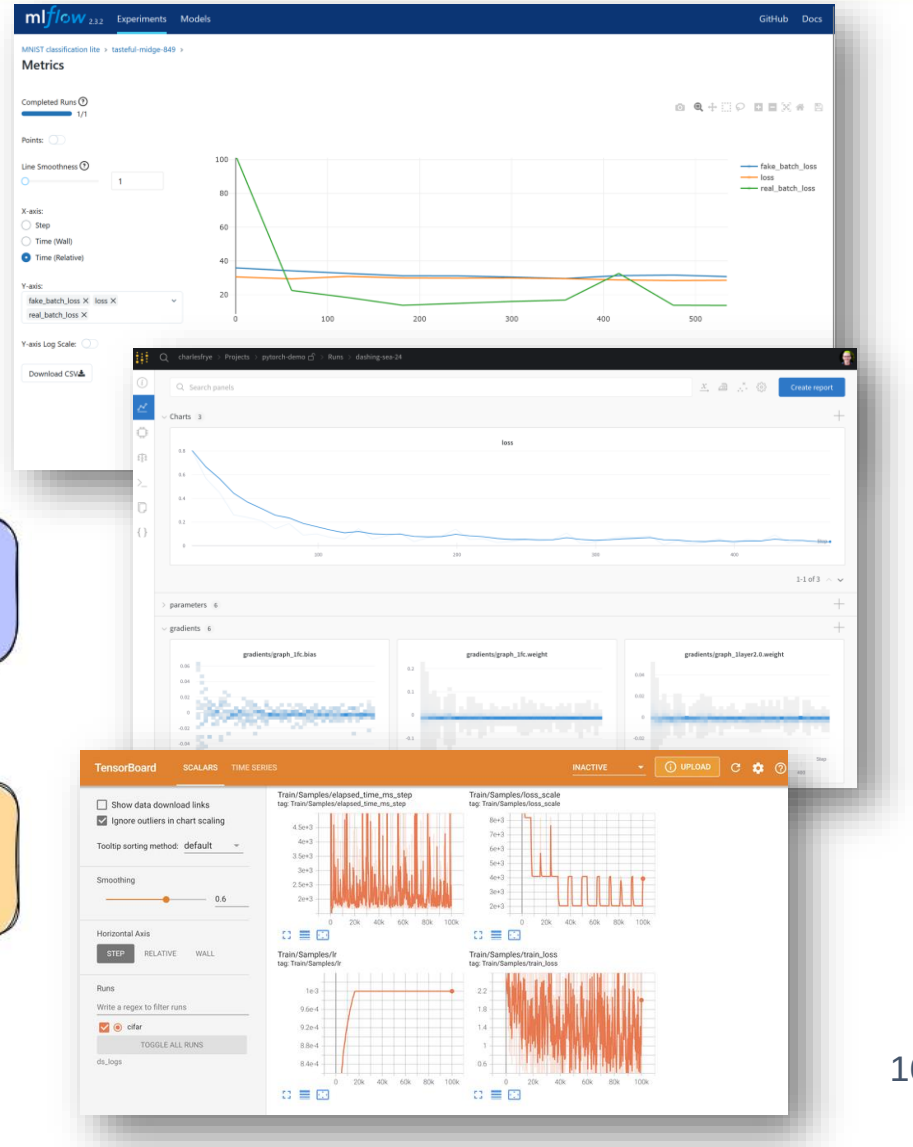
Unified distributed training (informal representation)



Our contribution to the Core Engine

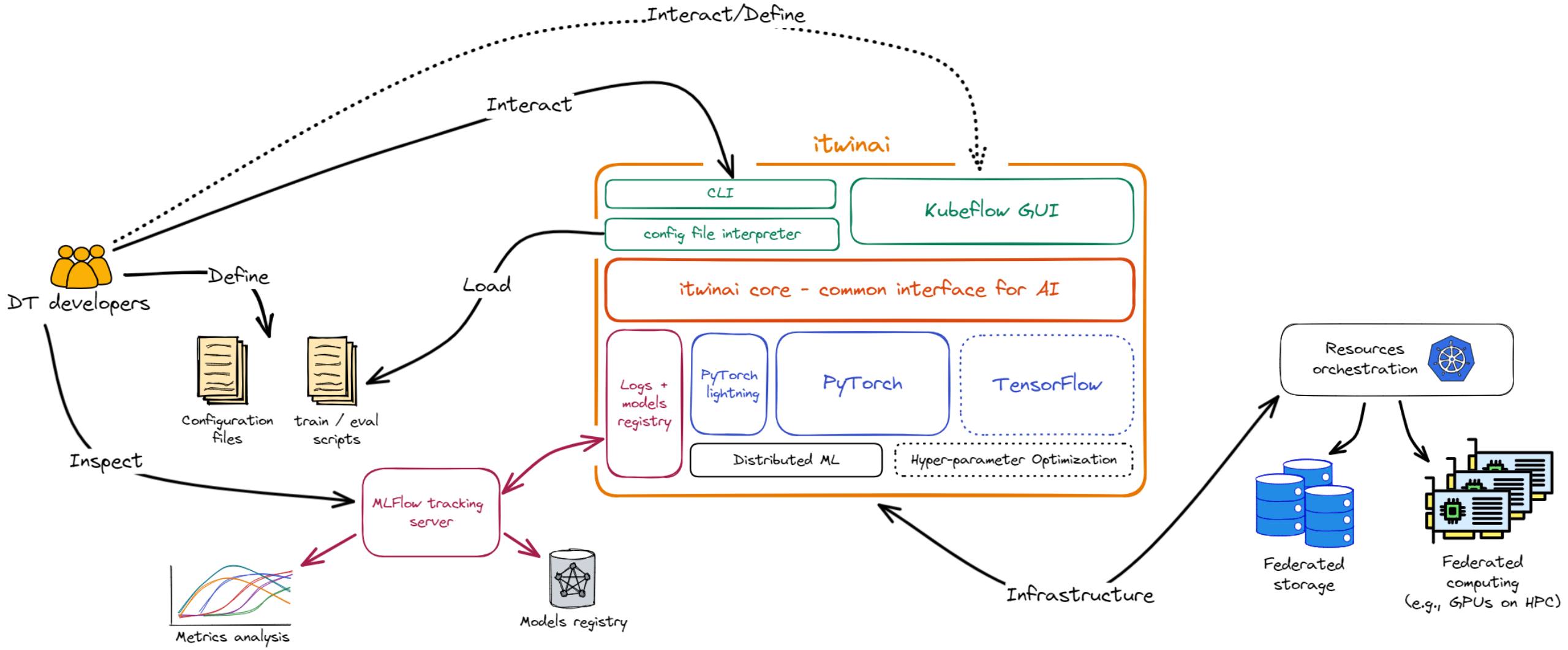
Unified logger (informal representation) to track ML metadata.





Can be extended to new loggers.

itwinai - ML tooling for DT applications



Latest developments

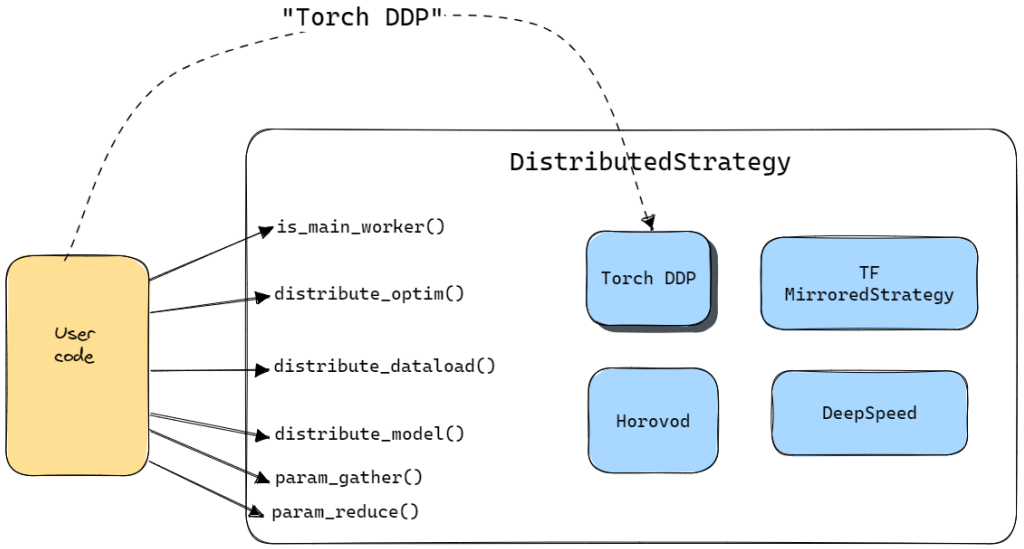
Development status of `itwinai` library:

- Support for **PyTorch**, **TensorFlow** under development.
- **AI workflows**: exploration of **KubeFlow** Pipelines.
- **Distributed ML**: integration of existing strategies (e.g., Toch DDP, Horovod, DeepSpeed).
- “Read The Docs”-style **documentation page**.
- **Link with the infrastructure**: Docker/Singularity container, offloaded through WP5’s **interLink** on cloud/HPC systems.
- **ML logs and models**: MLFlow tracking and WandB.
- EC review demo video.
- D6.2 First release of the DTE core modules: https://doi.org/10.5281/zenodo.10224213.

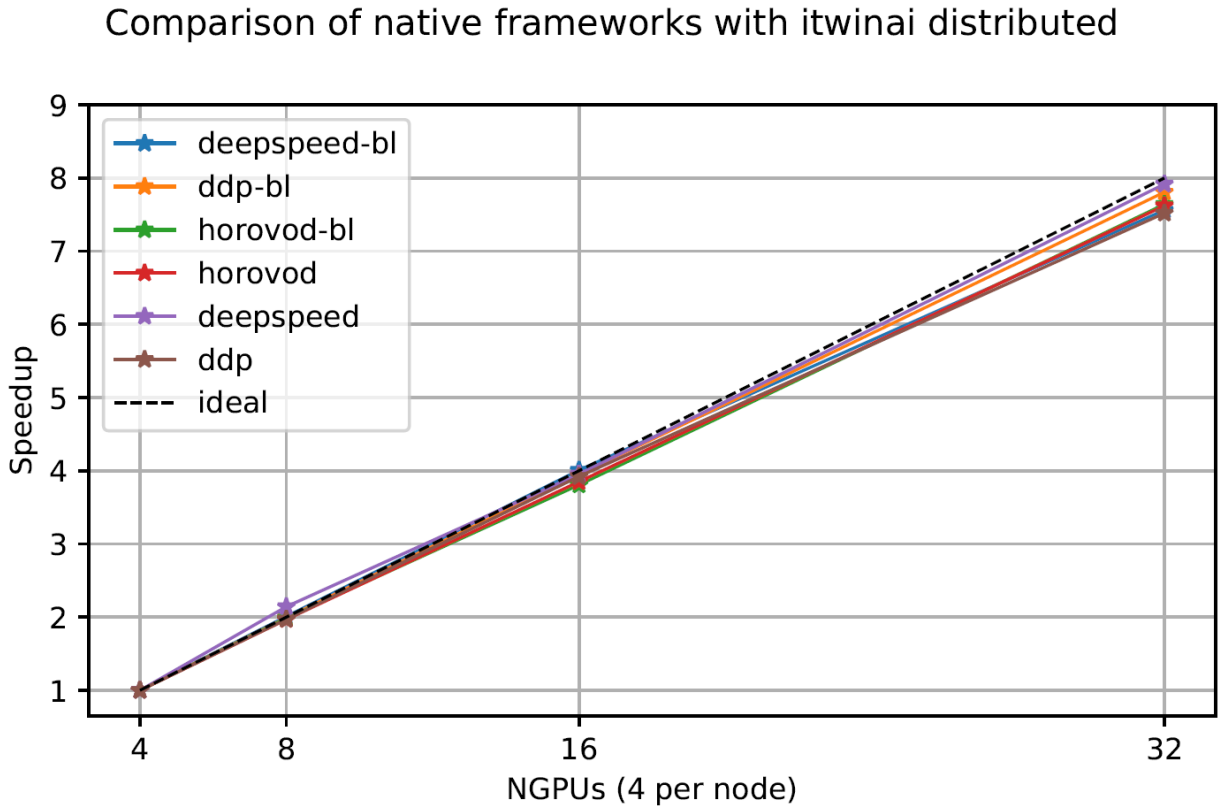
Preliminary results – Distributed training

Scalability comparison between `itwinai` distributed ML API and native frameworks (Torch `DistributedDataParallel`, Microsoft DeepSpeed, and Horovod) serving as baselines (marked with "-bl" suffix).

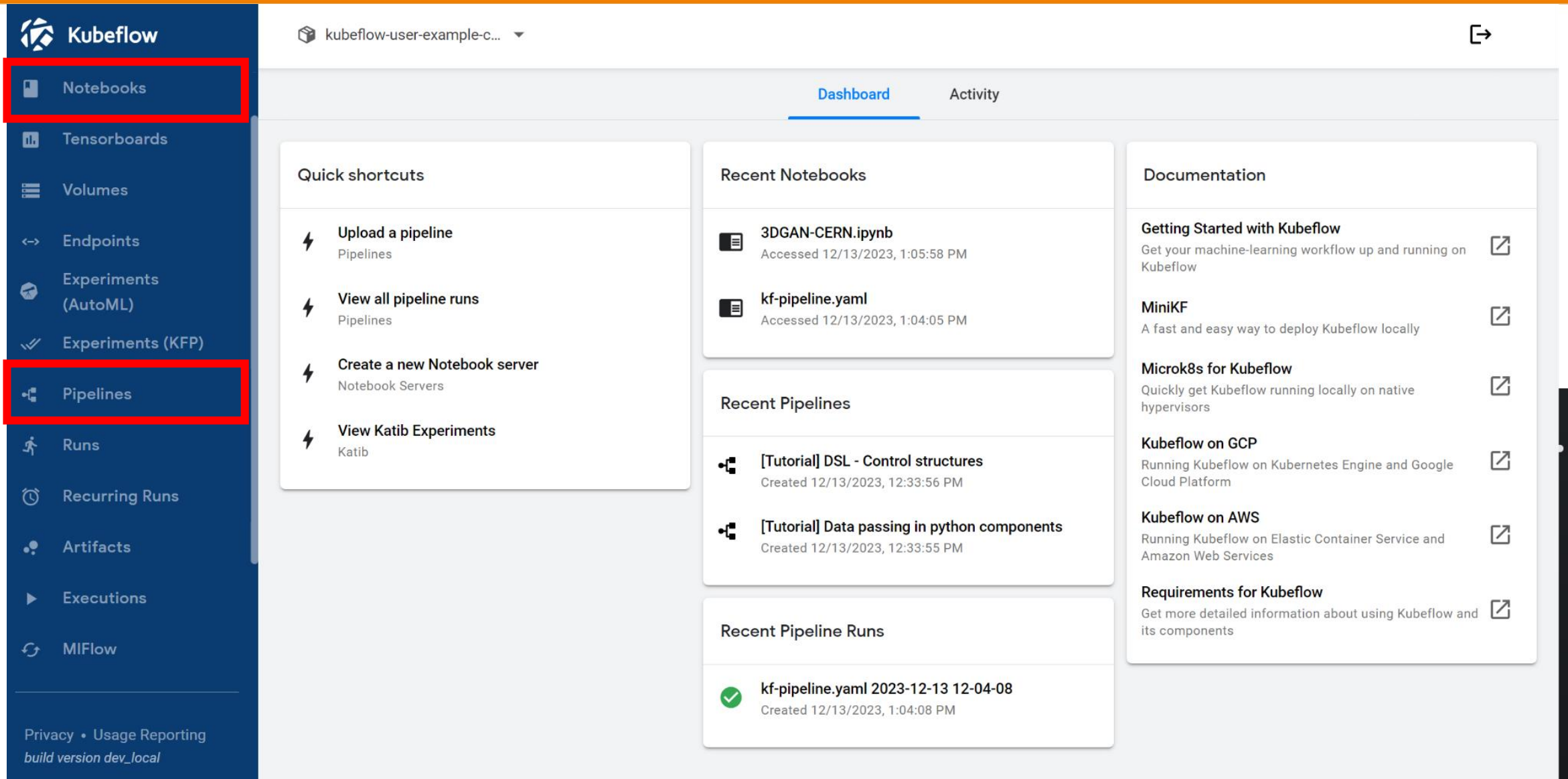
Scalability measured on ResNet152 on a subset of Imagenet.



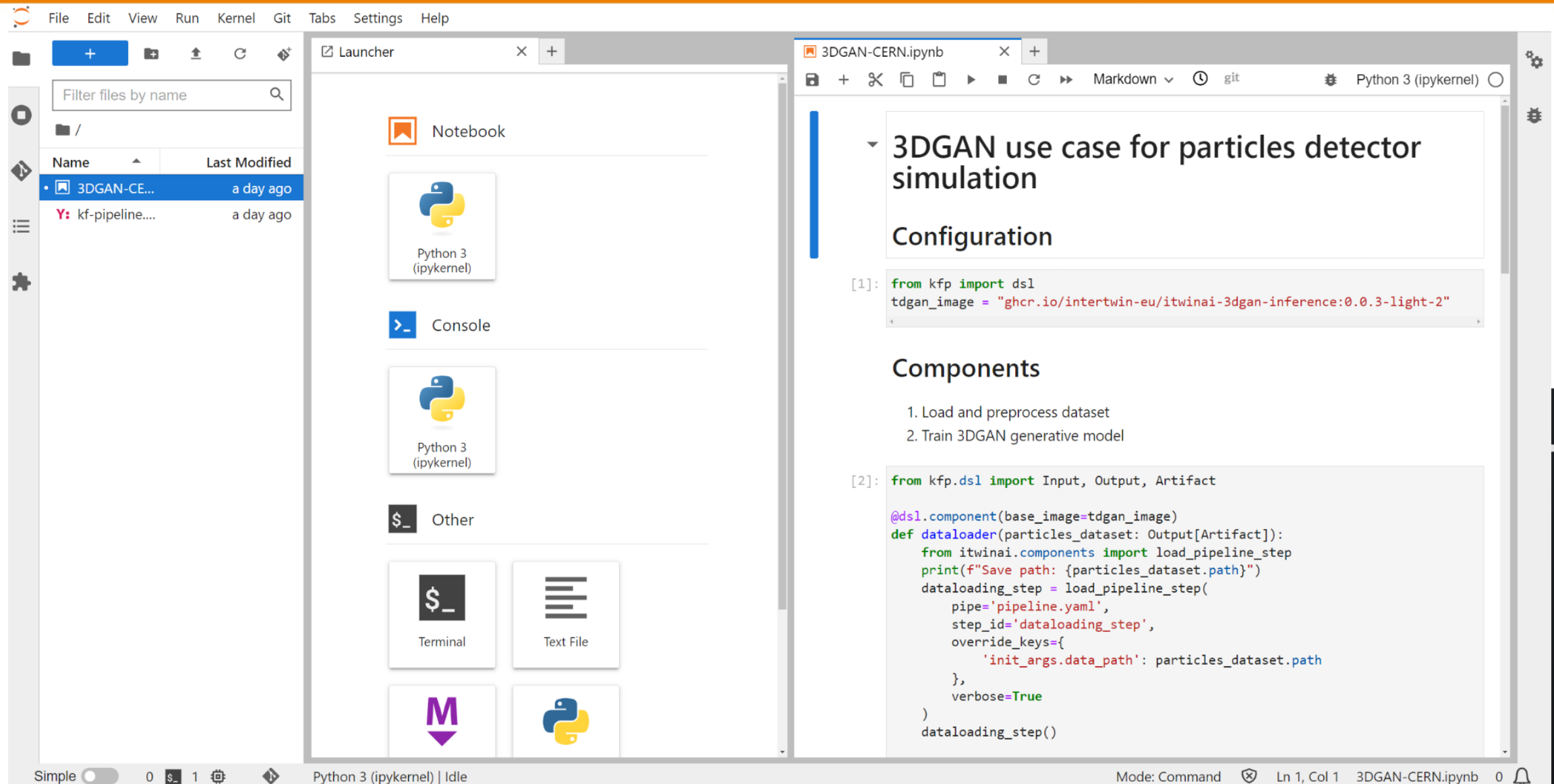
Intuitive representation of itwinai distributed ML API.



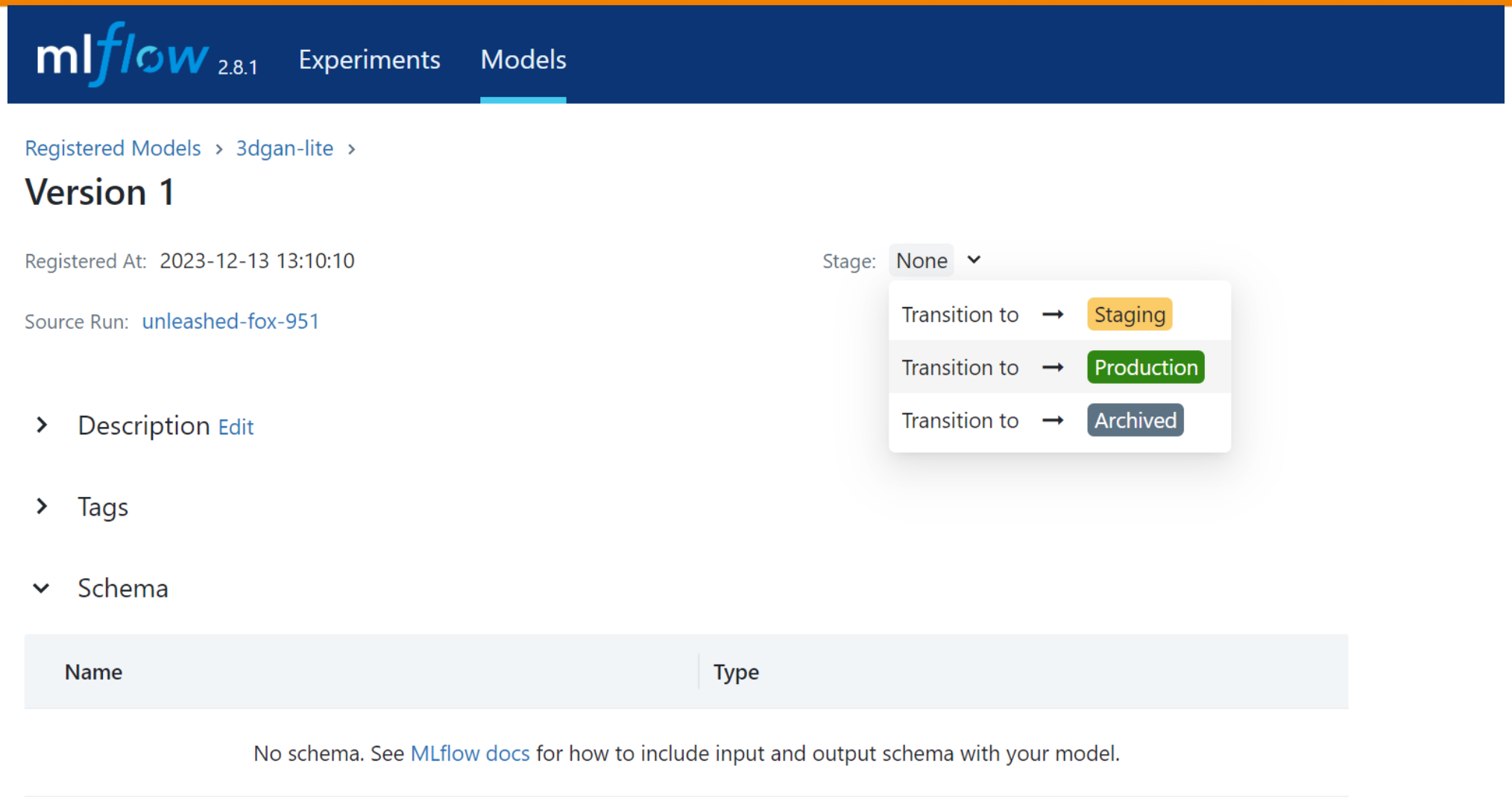
UX/UI: KubeFlow



UX/UI: KubeFlow and JupyterLab



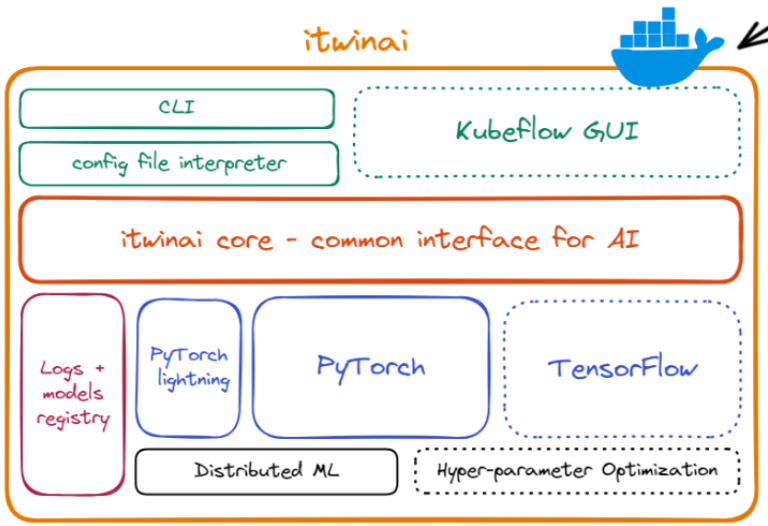
UX/UI: MLFlow models registry

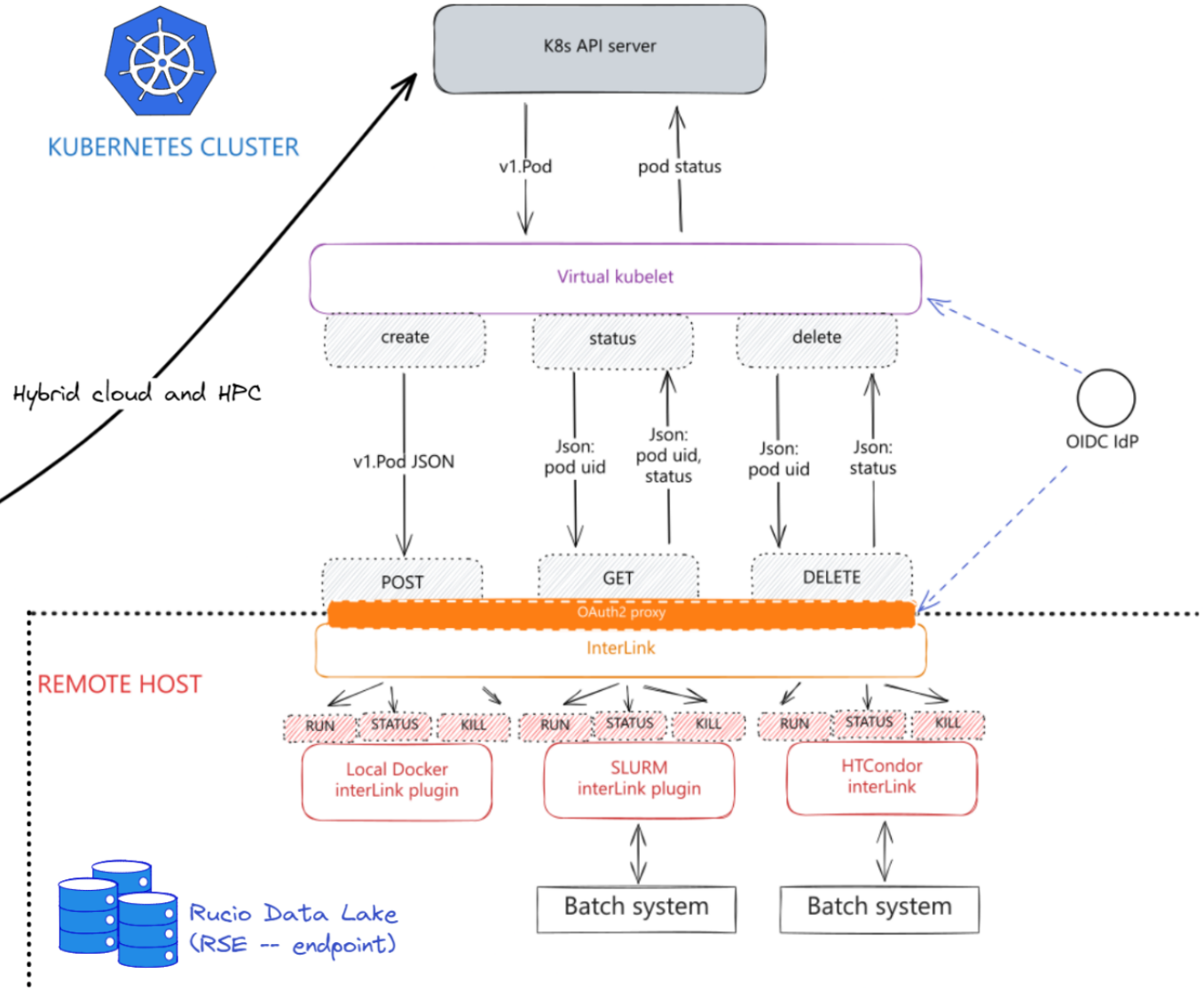


Federated computing

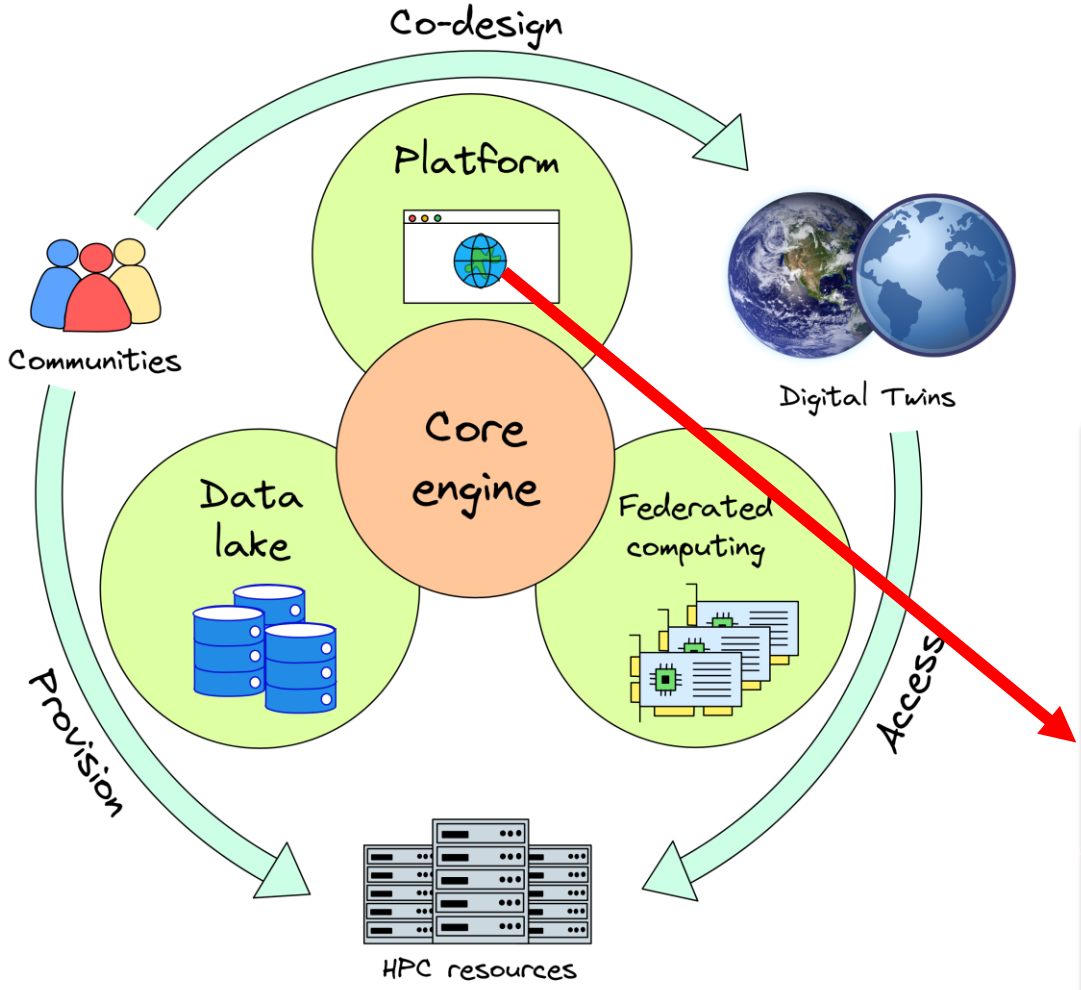
Access to infrastructure:

- SLURM on HPC
- InterLink: leverages k8s' Virtual Kubelet and microservice architecture of the DTE.
- Data management: Rucio data lake, OSCAR





Link with industry (CERN openlab)



Nvidia Omniverse:

- Visualization of scientific data (HEP)
- Support the design of new detectors and accelerators by integration and simulation tests
- Training of RL-based agents (e.g., robots) in simulated environments (IsaacGym)

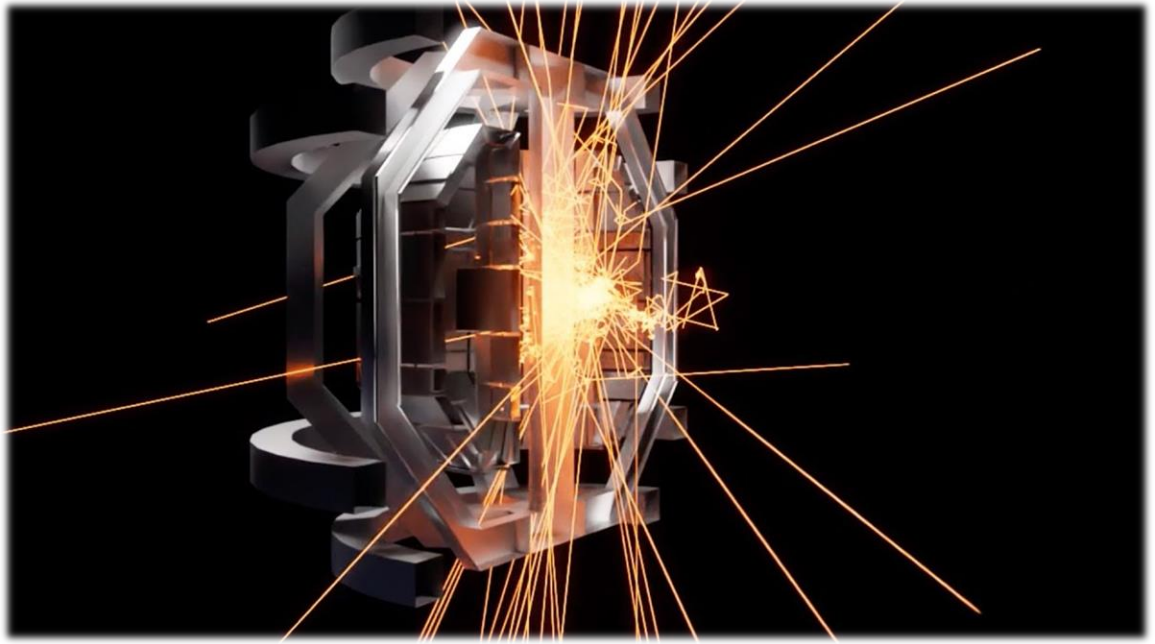


Image credits: https://blogs.nvidia.com/blog/ukaea-digital-twins-omniverse/

Thank you! Questions?

matteo.bunino@cern.ch

www.intertwin.eu

info@intertwin.eu

intertwin_eu

intertwin

itwinai

GitHub repository: https://github.com/interTwin-eu/itwinai

- For the moment, please refer to the “dev” branch: https://github.com/interTwin-eu/itwinai/tree/dev
- Some **tutorials** available, **more to come**. Check “tutorials” folder: https://github.com/interTwin-eu/itwinai/tree/dev/tutorials and https://github.com/interTwin-eu/itwinai/tree/distributed-strategy-launcher/tutorials/distributed-ml
- Additional examples are under “use-cases” folder: https://github.com/interTwin-eu/itwinai/tree/dev/use-cases
- Docs: https://intertwin-eu.github.io/itwinai/