WP3: ML-based analytics

- Lead: Bruno Raffin (Inria) & Thomas Moreau (Inria)

The team involved	DataMove, MIND, Thoth, SODA, STATIFY, CEA MdlS

WP-leaders

B. Raffin
Inria
T. Moreau
Inria

Main objective: support automated data analytics to reduce the need for human inputs

- **[T3.1]** Analysis of software packages and (ML/DL based) methodologies for simulation-based analytics, design with other WPs of PC3 of a componentized architecture
- **[T3.2]** Iterative/incremental ML parallel algorithms for online/out-of-core large scale data analysis
- **[T3.3]** Ensemble-based online scalable data analytics framework
  - For classical statistical approaches (sensibility analysis, data assimilation) and neural approaches (simulation based inference, deep reinforcement learning, adaptive experimental design, deep surrogate training)
- **[T3.4]** Statistical analytical tools for large distributed data
  - Supporting scikit-learn on HPC platform behind task based Python frameworks (Ray, Dask,Joblib) in collaboration with WP1,augmented with accelerated kernels (in coordination with PC2).

WP3: Budget 840 k€ HR

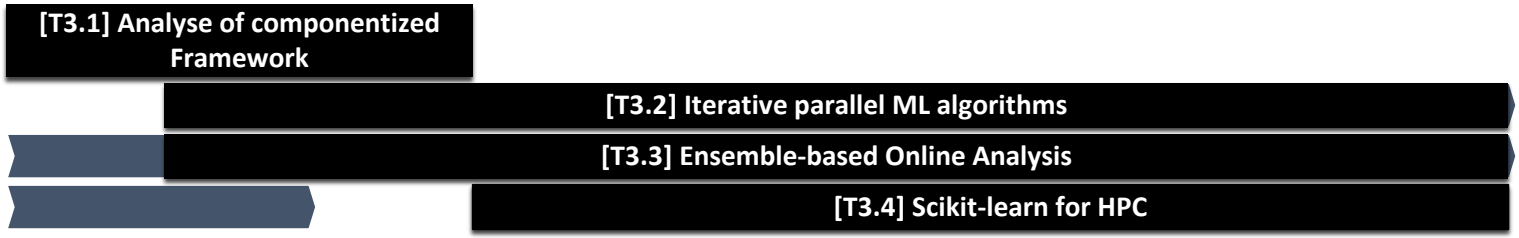
Core teams

DataMove (Inria)

MIND (Inria, CEA)

\+ collaborators

- 1 PhD Inria MIND (+ MdlS + Thot)
- 1 PhD Inria Datamove (+Satisfy)
- 2 y Postdoc Inria (MIND)
- 5y Eng Inria Datamove
- 5y Eng Inria Mind (+MdlS)



2.5y Eng. + 3y PhD

2.5y Eng. + 3y PhD

5y Eng.

WP3: Hiring

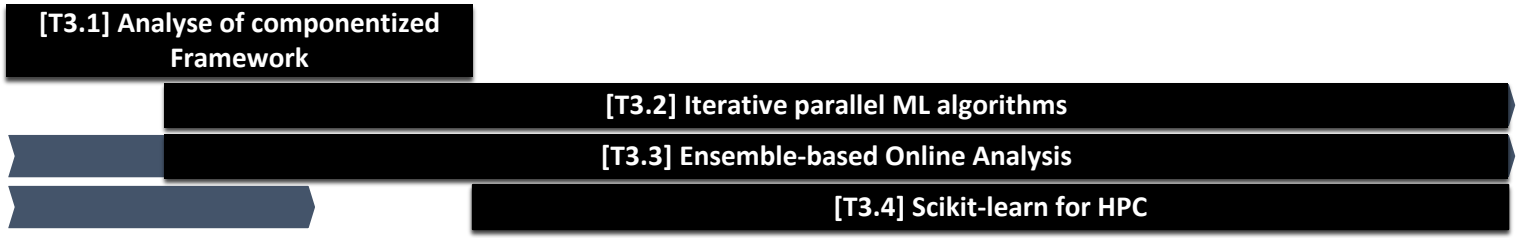
Core teams

DataMove (Inria)

MIND (Inria, CEA)

\+ collaborators

- 1 PhD Inria MIND (+ MdlS + Thot)**-> Try again in 2025 (M2 first)**
- 1 PhD Inria Datamove (+Satisfy) **-> Try again in 2025 (M2 first)**
- 2 y Postdoc Inria (MIND) **-> Mansour Benbakoura** (w/ Gysela, MdlS)
- 5y Eng Inria Datamove **-> Abhishek Punrandare**
- 5y Eng Inria Mind (+MdlS) **-> Some first contact with candidates**



WP3: Actions

- Melissa (A. Purandare, S. Dymchenko, B. Raffin):
  - Adios2 as transport layer (hard to get a performance boost vs ZMQ)
  - Adaptive input parameter sampling (1 paper @ AI4Science, SC24 workshop)
  - Talk at WANT workshop, ICML 24
- SBI (P. Rodrigues, B. Raffin, T. Moreau)
  - First Steps into active learning for SBI (Camille Touron, M2 internship)
  - Discussions with EDF on generative flows and SBI
  - Participate in sbi package (TU Tubingen)
- AI4Science (P. Rodrigues, T. Moreau, J. Le Sommer, B. Raffin)
  - GAP workshop @ Grenoble: https://gap2024.sciencesconf.org/
  - Chair Proposal @ Grenoble AI Cluster (MIAI)
  - Sacl-AI4Science workshop @ Saclay

WP3: Actions

- Data Challenge (T. Moreau, M. Mancip, M. Lobet):
  - First data challenge on Hot Jupiter pattern classification
  - Organizing a second on Laser-Plasma interactions
- Event detection in simulation (M. Benbakoura, T. Moreau, V. Grandgirard)
  - Literature review on event detection in computer vision
  - Discussions with Gysela on a reduced application (Tokam2D)
- Distributed dimension-reduction (H. Hendrickx, T. Moreau)
  - Many interviews to find a PhD candidates but hard to find the right profile!

A word about the GT IA

Missions:

- List IA use in NumpEx and interactions w. HPC in general
- Drive transverse initiave around IA (workshops, formations, …)
- Help identify ML/HPC interactions for science in the project

Members: E. Franck, T. Moreau

J. Bobbin, J-P. Villotte, T. Deutsch, P. Helluy, M. Clausel, A. Buttari, J Lesommer

Don’t hesitate to send us interesting joint use of IA/HPC.