SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

Dagstuhl Seminar 23352

Integrating HPC, AI, and Workflows for Scientific Data Analysis

(Aug 27 – Sep 01, 2023)

Topics for Discussion

CURRENT CHALLENGES

Workflow Dynamics and Management (Fred, Rosa B., Silvina, Wolfgang, Domenico, Kyle)

- Managing AI+HPC workflow dynamics
- ML-heavy vs. normal workflows

Human Interaction and Accessibility (Mathias, Paolo, JJ, Dan, Rafael)

- Addressing "human in the loop" in AI+HPC
- Expertise balance in big data technologies
- Bridging HPC and 'data people' gap

Integration and Standardization (Rosa F., Bertram, Shantenu, Ana, Jeyan, Ulf, Ilkay)

- Integrating HPC, AI, and workflow modeling
- Universal APIs for workflow components
- Continuous benchmarking of AI+HPC workflows

Sustainability Concerns (Timo, Dejan, Laure, Christine, Sean)

- Reducing carbon emissions in workflows
- Crafting sustainable AI+HPC workflows

FUTURE PERSPECTIVES

Advanced Techniques and Innovations ()

- Using synthetic data in AI+HPC
- LLMs in AI+HPC workflow design
- Systematic uncertainty in workflows
- Transfer learning in AI+HPC

Scalability and Deployment ()

- Exploring federated workflow engines
- Scalability and easy deployment of dynamic workflows

Optimization and Efficiency ()

- ML for HPC energy optimization
- Enhancing AI+HPC workflow efficiency
- Performance efficiency in running workflows and improving productivity in their development and deployment

Collaboration and Community Building ()

- Data-centric AI as a potential path of commonality to drive community collaboration
- The prospects of creating extensible and canonical forms of workflows for shared templates, benchmarking, or automated generation

Preliminary Agenda

Sunday – Aug 27, 2023

3:00pm-5:00pm	Arrival and Registration at Schloss Dagstuhl
6:00pm-8:00pm	Meet and Greet Dinner

Monday – Aug 28, 2023

9:00am-10:30am	Introductions and Seminar Goals - Participants introduce themselves
10:30am-11:00am	Break
11:00am-Noon	Administrative Items
Noon-2:00pm	Lunch
2:00pm-3:30pm	Short-talks: Current challenges on Integrating HPC, AI, and Workflows for Scientific Data Analysis - Fred Suter, "*How to express / handle the dynamic and cyclic nature of modern workflows?*" - Domenico Talia, "*ADAGE: Data-aware scheduling for large-scale distributed workflows*" - Rosa M Badia, "*Challenges considered in the framework of the eFlows4HPC project*"
3:30pm-4:00pm	Break
3:30pm-4:30pm	Breakouts session: current challenges
4:30pm-5:00pm	Reports from Breakout sessions
6:00pm-8:00pm	Dinner

Tuesday – Aug 29, 2023

9:00am-10:00am	Short-talks: Current challenges on Integrating HPC, AI, and Workflows for Scientific Data Analysis - Bertram Ludaescher, "*Trusting Computational Research: Reproducibility and Transparency*" - Ana Gainaru "*I/O patterns in AI workflows and their impact on the performance of HPC simulations*"
10:00am-10:30am	Break
10:30am-noon	Breakouts session: current challenges
Noon-2:00pm	Lunch
2:00pm-3:30pm	Breakouts session: current challenges
3:30pm-4:00pm	Break
4:00pm-5:00pm	Reports from Breakout sessions
6:00pm-8:00pm	Dinner

Wednesday – Aug 30, 2023

9:00am-10:00am	Short-talks: Future perspectives on HPC, AI, and Workflows - Paolo Missier, "*How can HPC and workflow technology best support the emerging Data-centric ML/AI paradigm?*" - Ilkay Altintas, "*From Workflows to Teamflows: Integrating Collaborative Science, AI, and Computing*" - Dejan Milojicic, "*Workflows for Heterogeneous Serverless Computing*"
10:00am-10:30am	Break
10:30am-noon	Breakouts session: future perspectives
Noon-1:30pm	Lunch
1:30pm-5:00pm	Group activity (touristic attraction)
6:00pm-8:00pm	Dinner

Thursday – Aug 31, 2023

9:00am-10:00am	Breakouts session: future perspectives
10:00am-10:30am	Break
10:30am-noon	Reports from Breakouts sessions
Noon-2:00pm	Lunch
2:00pm-3:30pm	Report writing activity: current challenges
3:30pm-4:00pm	Break
4:00pm-5:00pm	Report writing activity: future perspectives
6:00pm-8:00pm	Dinner

Friday – Sep 1, 2023

9:00am-10:00am	Finalize report writing activity
10:00am-10:30am	Break
10:30am-noon	Final thoughts and next steps
Noon-1:00pm	Lunch and departure

Proposed Topics for Discussion (from participants)

- How to efficiently and easily express/handle the dynamic and cyclic nature of AI+HPC workflows?
- Data-centric AI involves an iterative interplay between ML/AI modeling and data quality / data augmentation processes. These are at the same time compute-intensive but also interactive, including a strong "human in the loop" component.
- How can system infrastructure (HPC), AI techniques (for instance for generating synthetic data) and workflow modeling be combined to achieve effective data-centric management of AI models?
- Do we need different workflow engines for different types of problems? What properties of an engine are important for which types of problems?
- How should MLOps be merged with workflow engines?
- How can the community create common, generally agreed upon APIs between different components of workflow systems to allow interchanging of technologies (e.g.,

scheduling; resource allocation; resource estimation; file exchange; etc.). Why don't we already have them after dozens of years of research and development?
- How much should "big data technologies" remain accessible only for experts, and how much should they be opened for knowledgable lay persons (i.e., non computer science scientists)
- How should the communities react to the pressing needs of reducing carbon emissions / energy consumption?
- Are "ML-heavy" workflows in any way different from "normal workflows"? In which sense?
- A call for federated workflow engines - what are the challenges, what is the SOTA?
- How can the use of machine learning to improve machine learning and HPC utilization (or workflows) be promoted and expanded? I am particularly interested in looking for energy optimization where a combination of hardware and software yields answers in similar timeframes for much less energy usage.
- At a more general level, I am curious about what causes the seeming divide between HPC and 'data people' and how this can be bridged. Could data-centric AI as a specialty provide a path of commonality and if so, what are the research topics of interest that would drive community collaboration?
- Given an extensible set of existing AI+HPC workflows, can we reduce them to their canonical forms to share them as templates, or for benchmarking purposes or to use them for automated workflow generation? Can we formalize this problem (e.g. algebra)?
- What could be the role of synthetic data for designing, testing, debugging, validating AI+HPC workflows? Can we propose a framework?
- Is there a need for a solution of continuous benchmarking, maintenance, and versionning of AI+HPC workflows? What are the challenges?
- How to optimize a given end-to-end AI+HPC workflow?
- How to orchestrate the HIL in AI+HPC workflows? What are the quality metrics?
- Are there user-friendly, ready-to-use adaptive solutions to scale AI workflows up to AI+HPC workflows and inversely to scale them down depending on the needs and datasets' characteristics?
- Are there existing solutions that apply transfer learning to generate AI+HPC workflows from one domain to another domain? What are the main challenges?
- How LLMs can affect and be beneficial to the design/generation of AI+HPC workflows?
- How to "rewrite" some given AI+HPC workflows to make them more frugal and more sustainable?
- How to systematically consider uncertainty in every step of AI+HPC workflows?
- Seamless scalability and fluidity in deploying workflows
- Improving productivity of developing and deploying workflows
- Performance efficiency in running workflows