Organisation for Economic Co-operation and Development

Unclassified **English - Or. English**

Third OECD/NEA EGMUP Task Force on Artificial Intelligence and Machine Learning Benchmark Meeting (TF-AIML-3)

30 June 2025
Department of Engineering of the University of Cambridge, Trumpington Street, CB2 1PZ, Cambridge, United Kingdom

The workshop will be held in Cambridge, UK, as part of the Benchmarks Workshops of the Working Party on Scientific Issues and Uncertainty of Reactor Systems (WPRS). The WPRS Benchmarks Workshops event with three parallel tracks combines benchmarking efforts in common areas such as neutronics, thermal-hydraulics, and multi-physics modelling and uncertainty analysis.

wprs@oecd-nea.org

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Third OECD/NEA EGMUP Task Force on Artificial Intelligence and Machine Learning Benchmark Meeting (TF-AIML-3)

Sponsorship

The third OECD/NEA EGMUP Task Force on Artificial Intelligence and Machine Learning Benchmark Meeting (TF-AIML-3) will be held on 30 June 2025 in Cambridge, UK, as part of the Benchmarks Workshops Event of the Working Party on Scientific Issues and Uncertainty of Reactor Systems (WPRS). This event with three parallel tracks combines benchmarking efforts in common areas such as neutronics, thermal-hydraulics, and multi-physics modelling and uncertainty analysis. The full schedule for the incoming WPRS Workshops, SINUS-3 school and CTF-11 Meeting and Training is given on the public meeting web page: https://www.oecd-nea.org/jcms/pl_98752/wprs-benchmarks-workshops-2025

Executive Summary

AI/ML have gained tremendous interest across all branches of nuclear engineering, driven by the increasing computing power, advancements in AI/ML algorithms, and the availability of user-friendly ML libraries. While it is widely acknowledged that AI/ML has great potential to enhance data-driven scientific computing problems in nuclear engineering, its potential has not been fully realized due to several challenges. First, there are no guidelines on how to apply application-agnostic AI/ML algorithms to high-consequence nuclear systems. Second, there are no AI/ML benchmarks based on realistic nuclear reactor measurement data or numerical simulation data to compare and validate various ML algorithms. Third, there are significant gaps in the trustworthiness assessment of AI/ML models for nuclear applications. In May 2022, the OECD/NEA Working Party on Scientific Issues and Uncertainty Analysis of Reactor Systems (WPRS) endorsed at its annual Benchmark Workshops the creation of a task force on ``Artificial Intelligence and Machine Learning for Scientific Computing in Nuclear Engineering" under the auspices of its Expert Group on Reactor System Multiphysics (EGMUP).

The goal of this task force is to develop benchmark exercises that aim at evaluating the performance of AI/ML in multi-physics M&S of nuclear reactor systems. It will provide recommendations to the WPRS and the nuclear community on the scientific development needs (data, methods, and benchmark exercises) for trustworthy AI/ML applications in nuclear scientific computing problems. This benchmark will design and execute benchmark exercises on both single physics (reactor physics, thermal-hydraulics and fuel performance) and multi-physics coupled simulation problems. The exercises will target specific key challenges of each computational domain through interactions with the corresponding WPRS expert groups. A broad spectrum of AI/ML sub-domains will be explored, including supervised learning, unsupervised learning, reinforcement learning, deep generative learning, and probabilistic ML. In each benchmark exercise, the participants will use self-chosen AI/ML algorithms to perform tasks specified in the exercises. The development of the benchmark specifications will be conducted in two phases.

1. Phase I of the benchmark will focus on tasks related to regression, classification, dimensionality reduction and anomaly detection, and
2. Phase II of the benchmark will focus on tasks related to VVUQ, transfer learning, deep generative learning and deep reinforcement learning.

The first benchmark from Phase 1 (CHF) was successfully executed and the main results and conclusions will be discussed. The finalized Phase 1 specifications for the PUR-1 benchmark and the last updates for

the SAFARI-1 benchmark will be presented and discussed. Participants' results and studies will be presented. In depth discussion on the Phase 2 benchmark specifications development will be included.

The information about the Task Force on AI/ML is provided at:

https://www.oecd-nea.org/jcms/pl_77779/task-force-on-artificial-intelligence-and-machine-learning-for-scientific-computing-in-nuclear-engineering

Scope and Technical Content of the Meeting

The topics to be addressed at the workshop include:

- Review and discussion of databases that will be used for evaluating the performance of AI/ML techniques in the benchmark exercises,
- Updates from the subgroup leads on the benchmark exercise development,
- Updates from the subgroup leads on the remaining benchmark specification drafts for Phase I,
- Feedback and concerns of task force participants for the exercises and specifications,
- Presentation and discussion of results from the Phase I CHF benchmark execution.
- Discussion for the upcoming Phase II benchmark specifications,
- Discuss other relevant activities, and
- Defining a work plan, schedule, and goals for the next year of the Task Force.

The proposed meeting program is attached as Annex 1.

Organization of the Meeting

The meeting is organized around the discussion of the benchmark exercises, benchmark specification preparation, TF participants' concerns, Phase I execution, and other benchmark-related activities. The subgroup leads are requested to present their progress on benchmark exercise development and benchmark specification preparation.

Participation in the Meeting

Participation to the meeting is not restricted.

Organization and Program Committee of the Meeting

An Organization and Program Committee has been nominated to make the necessary arrangements for the TF-AIML-3 benchmark meeting and to draw up the final program, etc.

The members of the Program Committee are:

Xu Wu – *Chair, Co-Cordinator*
North Carolina State University, USA

Gregory Delipei - *Co-Chair, Co-Coordinator*
North Carolina State University, USA

NEA Technical Secretariat: **Oliver Buss**
OECD/Nuclear Energy Agency, France

Proposed Program of the Meeting

The proposed program was drawn up by the Program Committee and is enclosed as Annex 1.

Language of the Benchmark Workshop

The official language of the workshop is English.

Proceedings of the Meeting

The program committee and the session Chairs will prepare a summary report on the main results of the meeting. Presentations will be available free of charge to the benchmark participants to download from participants' restricted area of the benchmark website after the workshop.

Contacts and Registrations

Link to the registration page for the WPRS-related workshops and meetings and the overall program: https://www.oecd-nea.org/jcms/pl_98752/wprs-benchmarks-workshops-2025

Workshop Location

The meeting place/venue for the meetings/workshops during the week of 30 June to 4 July 2025 is the Department of Engineering of the University of Cambridge:

Trumpington Street, CB2 1PZ, Cambridge, United Kingdom.

Annex 1:

The Third OECD/NEA EGMUP Task Force on Artificial Intelligence and Machine Learning Benchmark Meeting (TF-AIML-3)

Host Organization

Hosted by Cambridge Nuclear Energy Centre, University of Cambridge,
Cambridge, United Kingdom

June 30, 2025

PROPOSED PROGRAM

TA01-10: Session codes

TA01. Introduction and opening remarks

TA02. Overview of Task Force activities and current status

TA03. Updates from the subgroup leads on the benchmark exercise development

TA04. Updates from the subgroup leads on the benchmark specification drafts for Phase I

TA05. Feedback and concerns of Task Force participants for the exercises and specifications,

TA06. Discussion on the results and conclusions from the CHF benchmark Phase 1 execution.

TA07. Discussion on the Phase 2 benchmark specifications development.

TA08. Discussion on the potential interaction of this Task Force with other relevant activities

TA09. Action items, schedule, and goals for the next workshop (TF-AIMI-3)

TA10. Conclusions and closing remarks