

Seed Funding Global South 2023 – Project report

Green Machine Learning Operations

- **Chair at the TU Berlin:** Distributed and Operating Systems (DOS)
- **Partner country/countries:** North Macedonia
- **Partner institution (s):** University of Kiril and Metodij, Skopje, Faculty of Computer Science and Engineering
- **Sustainable Development Goals (SDGs):**



© UNITED NATIONS

SDG 12: Responsible Consumption and Production

With the unprecedented advances in machine learning, intelligent models in terms of online service offerings are an indispensable part of daily human lives. AIOps (Artificial Intelligence for IT Operations) has emerged as a key solution for supporting cloud operations by automating IT tasks. AIOps methods, operating as separate workloads on a dedicated part of the cloud, efficiently handle observability data and enhance overall computation. Despite being the key transformative technology that improves overall human well-being, the scale of machine learning adoption incurs high energy demands. For example, reports suggest that training a large language model consumes the energy equivalent to five car lifetimes. Given their wide adoption, machine learning models have non-negligible indirect contributions to the emission of greenhouse gases (e.g., CO₂). While machine learning model adoption is expected to grow further, to reduce its negative impact on the environment, it is also important to quantify the sustainability of their overall lifecycle (e.g., through CO₂ benchmarking) such that appropriate policies to protect the environment can be developed.

In this Global South initiative, we responded to the sustainable development goal "sustainable consumption and production" and we initiated a collaboration effort with Professor Gjorgji Madjarov from the University of Saint Cyril and Methodius in Skopje, North Macedonia. The collaboration aims to study and improve the indirect environmental impact of the current practices of executing AIOps workloads. Our primary goal is to understand the relationship between the existing scheduling policies in executing AIOps workloads and their sustainability. We explore the potential costs that arise from the time-criticality aspect of the AIOps workloads and any potential change in their rescheduling policy that considers their environmental impact. We focus first on exploring the current practices of when and where the AIOps workloads are scheduled. The goal in this regard is to understand the production practices when AIOps methods are executed. In the prevalent practice, they are considered part of regular workloads, and most commonly are manually retrained at regular intervals (e.g., once per day, or week).

Additional automatic strategies, such as concept drift detection, are employed but not frequently used in practice due to the additional error they introduce. Adequately, whenever the first strategy is adopted, the cloud providers can align the retraining schedules with periods of available green energy achieving zero-cost prevention, at no expense on performance. Concerning the latter strategy, AIOPs introduce greater associated costs, related to the constant need to run the concept-drift detection method. Our research emphasizes the need to quantify the carbon emission profiles of AIOPs workloads, considering factors such as geo-region and time of computation.

In the course of the study, we exchanged ideas and approaches with the worldwide leading AIOPs experts during two AIOPs workshop, which we organised in cooperation with Prof Gjorgji Madjarov. The first AIOPs workshop took place in Berlin in May 2023 and gathered 25 experts from all over the world. The second AIOPs workshop took place in December 2023 in the scope of the premier international conference on data mining (ICDM) in Shanghai, China. Seven talks on AIOPs underlined the importance of AIOPs and the need for sustainable development. To address these challenges further, we applied for one EU project and are currently planning further joint funding proposals. With the funding provided by TU Berlin' Seed Funding programme, we laid the ground for these activities by strengthening the collaboration through mutual exchange visits: Prof. Madjarov spent nearly two months in Berlin in 2023, while Prof. Kao and Dr. Bogatinovski spent around a month in Skopje, North Macedonia. The collaboration will continue with the next scheduled visit to Skopje in March 2024.