

MODIS

MLOps für verteilte Cyberphysische Systems

Programm / Ausschreibung	FinVn 22/23, Kooperationsstrukturen, Bridge Ausschreibung 2022	Status	abgeschlossen
Projektstart	01.01.2023	Projektende	31.12.2025
Zeitraum	2023 - 2025	Projektlaufzeit	36 Monate
Keywords	Software Engineering for Machine Learning (SE4ML), Machine Learning Operations (MLOps), Continuous Delivery (CD), Software-Architektur, Cyber-physische Systeme (CPS)		

Projektbeschreibung

Moderne cyber-physische Systeme (CPS) zeichnen sich durch einen hohen Grad an Verteilung aus, oft mit substantiellen Edge- und Cloud-Subsystemen. Diese Edge-/Cloud-Subsysteme werden häufig für Aspekte, die hohe Rechenleistungen benötigen eingesetzt, wie z.B. Machine Learning (ML). Insgesamt nimmt der ML-Anteil in CPS in den letzten Jahren deutlich zu. Machine Learning Operations (MLOps) bezeichnet die Disziplin, die sich mit der Bereitstellung von ML-Modellen durch automatisierte, wiederholbare und effiziente Arbeitsabläufe beschäftigt. MLOps ist für ML in CPS besonders interessant, um ML-Modelle effizienter ausliefern zu können, die Skalierung zu einer möglicherweise riesigen Menge von Modellen, Modellvarianten und Daten zu unterstützen und durch die Reproduzierbarkeit der Abläufe eine erhebliche Reduktion der möglichen Risiken zu erreichen.

MODIS zielt darauf ab, die Software-Engineering-Grundlagen für die nächste Generation von intelligenten cyber-physischen Systemen zu legen. Da bisher die software-technische Unterstützung und die Methodik der Softwarearchitektur für MLOps für verteilte CPS unzureichend ist, wird MODIS systematisch Muster und Smells in diesem Bereich analysieren und darauf basierend eine rigoros spezifizierte Modellierungsmethode und Metriken für MLOps in verteilten CPS entwickeln. Überdies soll ein neuartiges Architekturleitsystem zur semi-automatischen Anleitung des Engineerings von MLOps für verteilte CPS entworfen werden. Da die Überwachung der Qualität von MLOps für verteilte CPS bisher nur in begrenztem Umfang möglich ist, wird MODIS neuartige statische und dynamische Qualitätsprüfverfahren entwickeln und diese in aussagekräftigen Indikatoren zur Verfügung stellen. Dafür wird ein Pipeline-Quality-Insight-as-a-Service-Konzept entwickelt werden. Da eine Automatisierung von Aufgaben zur konsistenten und korrekten Realisierung von MLOps für verteilte CPS bisher nicht vorhanden ist, wird MODIS einen Low-Code-Ansatz für die Spezifikation der Pipeline-Architekturen vorschlagen, sowie neue systematische Pipeline-Architektur-Optimierungs- und -Generierungsansätze.

Mit diesen neuen Software-Engineering-Grundlagen für MLOps für verteilte CPS wird MODIS systematisches Engineering solcher Systeme, einhergehend mit deutlich verbesserter Qualität und geringeren Risiken in der Entwicklung und im Betrieb, ermöglichen. Es ist wichtig, heute diese Grundlagen zu erforschen, damit um die Jahre 2028-2030 eine nächste Generation von systematisch konstruierten, intelligenten cyber-physischen System großflächig zum Einsatz kommen kann.

Abstract

Modern Cyber-physical Systems (CPS) are characterized by a high degree of distribution, often with substantial edge and cloud subsystems. These edge/cloud subsystems are often used for aspects that require high computational power, such as Machine Learning (ML). Overall, the ML parts in CPS are increasing significantly in recent years. Machine Learning Operations (MLOps) refers to the discipline concerned with providing ML models through automated, repeatable, and efficient workflows. MLOps is of particular interest for ML in CPS in order to deliver ML models more efficiently, to support scaling to a potentially huge set of models, model variants, and data, and to achieve a significant reduction in potential risks through reproducibility of operations.

MODIS aims to lay the software engineering foundation for the next generation of intelligent cyber-physical systems. Since software engineering support and software architecture methodology for MLOps for distributed CPS is insufficient to date, MODIS will systematically analyze patterns and smells in this area and develop a rigorously specified modeling methodology and metrics for MLOps in distributed CPS based on them. Moreover, a novel architecture guidance system will be developed to semi-automatically guide the engineering of MLOps for distributed CPS. Since monitoring the quality of MLOps for distributed CPS has been limited so far, MODIS will develop novel static and dynamic quality checking techniques and make them available in meaningful indicators. For this purpose, a Pipeline-Quality-Insight-as-a-Service concept will be developed. Since the automation of tasks for the consistent and correct realization of MLOps for distributed CPS does not exist yet, MODIS will propose a low-code approach for pipeline architecture specification, as well as new systematic pipeline architecture optimization and generation approaches.

With these new software engineering foundations for MLOps for distributed CPS, MODIS will enable systematic engineering of such systems, accompanied by significantly improved quality and reduced risks in development and operations. It is important to explore these foundations today so that the next generation of systematically engineered intelligent cyber-physical systems can be deployed on a large scale around 2028-2030.

Projektkoordinator

- Universität Wien Forschungsgruppe Software Architecture (SWA)

Projektpartner

- Siemens Aktiengesellschaft Österreich