MLOX – MLOps in a Box

Schlussbericht

Zuwendungsempfänger: Dr. Nico Görnitz

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Forschung, Technologie und Raumfahrt unter dem Förderkennzeichen **16IS25S13** gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei dem Autor.

1. Kurze Darstellung der Aufgabenstellung und Motivation in der regulären Förderphase

Ziel des Projekts MLOX war es, eine offene und zugängliche Infrastruktur für Machine Learning Operations (MLOps) zu entwickeln, die insbesondere für Einzelpersonen und kleine Teams nutzbar ist. Während große Cloud-Anbieter umfassende Lösungen bereitstellen, sind diese häufig komplex zu bedienen, kostenintensiv und mit einer starken Abhängigkeit von proprietären Plattformen verbunden.

Vor diesem Hintergrund bestand die Motivation darin, eine souveräne Alternative zu schaffen, die auf Offenheit, Transparenz und Selbstbestimmung basiert. MLOX verfolgt den Ansatz, bestehende Open-Source-Komponenten zu einer integrierten und leicht nutzbaren Gesamtlösung zu verbinden, die auf eigener Infrastruktur betrieben werden kann und reproduzierbare sowie sichere ML-Systeme ermöglicht.

Die geplante Umsetzung orientierte sich an einem klar strukturierten Arbeitsplan mit aufeinander aufbauenden Meilensteinen. Zu Beginn stand die Analyse und Erweiterung möglicher Deployment-Strategien sowie die Entwicklung robuster Infrastrukturvorlagen für unterschiedliche Einsatzszenarien.

Darauf aufbauend sollte eine benutzerfreundliche Schnittstelle entstehen, die es erlaubt, MLOps-Komponenten einfach zu konfigurieren und zu betreiben. Parallel dazu war die Integration zentraler Dienste sowie grundlegender Monitoring-Funktionalitäten vorgesehen.

In einem weiteren Schritt sollte ein Python-Paket entwickelt werden, das die Anbindung von Nutzer-Code an die Infrastruktur vereinfacht und typische MLOps-Workflows unterstützt.

Abschließend waren umfassende Tests und eine strukturierte Dokumentation geplant, um die Funktionsfähigkeit des Prototyps unter realen Bedingungen sicherzustellen und eine breite Nutzbarkeit zu ermöglichen.

2. Beitrag des Projektes zu den Zielen der Förderinitiative „Software Sprint“

MLOX leistet einen direkten Beitrag zu den Themenfeldern Infrastruktur und Sicherheit. Im Bereich Infrastruktur adressiert das Projekt die Herausforderung, komplexe ML-Systeme für eine breitere Zielgruppe zugänglich zu machen. Durch die Integration etablierter Open-Source-Tools entsteht eine Plattform, die den Aufbau und Betrieb von MLOps-Systemen erheblich vereinfacht und gleichzeitig flexibel bleibt.

Im Bereich Sicherheit verfolgt MLOX einen bewusst selbstbestimmten Ansatz. Durch den Betrieb auf eigener Infrastruktur behalten Nutzer die Kontrolle über ihre Daten und Systeme. Ergänzend werden etablierte Mechanismen wie verschlüsselte Kommunikation, isolierte Netzwerke und Secret Management integriert.

Darüber hinaus trägt das Projekt zur Demokratisierung von KI-Infrastruktur bei, indem es insbesondere kleinen Teams und Forschungseinrichtungen ermöglicht, leistungsfähige Systeme unabhängig von großen Cloud-Anbietern zu betreiben.

3. Ausführliche Darstellung der Ergebnisse in der regulären Förderphase

In der regulären Förderphase wurde ein funktionsfähiger Prototyp von MLOX entwickelt, der zentrale Komponenten eines MLOps-Systems integriert und deren Zusammenspiel über eine einheitliche Konfiguration ermöglicht.

Die Umsetzung folgte dem im Antrag definierten Arbeitsplan und wurde schrittweise entlang der geplanten Meilensteine realisiert. Zunächst wurden verschiedene Deployment-Strategien analysiert und erweitert, insbesondere im Hinblick auf On-Premise- und hybride Szenarien sowie potenzielle Erweiterungen durch Kubernetes. Darauf aufbauend wurden Infrastrukturvorlagen entwickelt und validiert, um unterschiedliche Einsatzkontexte zuverlässig abzudecken.

Im nächsten Schritt wurde eine Benutzeroberfläche konzipiert und umgesetzt, die es ermöglicht, komplexe MLOps-Stacks über einfache Interaktionen zu konfigurieren und zu betreiben. Parallel dazu wurden erste Monitoring-Funktionalitäten auf Basis von OpenTelemetry integriert, um Einblicke in den Zustand und die Nutzung der Infrastruktur zu ermöglichen.

Ein weiterer Schwerpunkt lag auf der Integration zentraler MLOps-Komponenten. Hierbei wurden mehr als zwanzig etablierte Open-Source-Werkzeuge in die Plattform eingebunden und über ein gemeinsames Orchestrierungskonzept zugänglich gemacht. Die zugrunde liegende Architektur folgt dabei einem service-orientierten Ansatz, bei dem einzelne Komponenten flexibel kombiniert und betrieben werden können.

Im weiteren Verlauf wurde ein Python-Paket entwickelt, das die Anbindung von Nutzer-Code an die bereitgestellte Infrastruktur vereinfacht und über PyPI zugänglich gemacht wurde. Ergänzend dazu wurden Beispiel-Workflows implementiert, die typische MLOps-Anwendungsfälle abbilden und einen schnellen Einstieg ermöglichen.

Ein zentraler Bestandteil der Projektarbeit war die frühzeitige Validierung in realen Anwendungsszenarien. In diesem Kontext wurde ein Pilotprojekt mit der Airtime Software AG initiiert, in dem MLOX zur Umsetzung produktionsnaher Machine-Learning-Anwendungen eingesetzt wurde. Dieses Pilotprojekt wurde über die reguläre Förderphase hinaus fortgeführt und diente als kontinuierlicher Real-World-Test für die entwickelte Infrastruktur.

Die Zusammenarbeit mit einem realen Anwendungspartner erwies sich als besonders wertvoll. Sie führte zu konkreten Verbesserungen in der Benutzerfreundlichkeit, insbesondere bei Deployments und Schnittstellen, und lenkte den Fokus stärker auf praxisrelevante Anforderungen statt auf rein technische Erweiterungen.

Abschließend wurde der Prototyp umfassend getestet und dokumentiert. Neben umfangreichen Unit-Tests wurden auch komplexe Integrationstests entwickelt, um das Zusammenspiel der einzelnen Komponenten unter realistischen Bedingungen zu überprüfen. Die Dokumentation wurde parallel dazu gezielt ausgebaut und umfasst unter anderem ein strukturiertes öffentliches GitHub-Repository, eine generierte Dokumentationswebsite, eine ausführliche README sowie Richtlinien für Beiträge und Zusammenarbeit. Ergänzt wird dies durch weiterführende technische Dokumentationen, beispielsweise zur Architektur.

Begleitet wurde die reguläre Förderphase durch die Open Knowledge Foundation, deren kontinuierliches Feedback einen wichtigen Beitrag zur Strukturierung und Fokussierung des Projekts leistete. Insbesondere wurde frühzeitig deutlich, dass neben der technischen Entwicklung auch Aspekte wie Dokumentation, Zugänglichkeit und Community-Aufbau entscheidend für den nachhaltigen Erfolg eines Open-Source-Projekts sind.

Ergänzend dazu fanden im Rahmen des Programms gezielte Coaching-Formate statt, unter anderem in den Bereichen Community Building, Kommunikation und strategische Produktentwicklung. Diese Impulse halfen dabei, die Zielgruppen von MLOX klarer zu definieren, die Außendarstellung zu schärfen und erste Ansätze für eine langfristige Weiterentwicklung über die reine Prototypenphase hinaus zu entwickeln.

Insgesamt trug die begleitende Unterstützung wesentlich dazu bei, das Projekt nicht nur als technische Lösung, sondern als anschlussfähiges Open-Source-Vorhaben mit klarer Perspektive für Nutzung, Community und Weiterentwicklung zu positionieren.

4. Kurze Darstellung der Aufgabenstellung und Motivation in der Second-Stage-Förderung

Die Second Stage verfolgte das Ziel, den in der regulären Förderphase entwickelten Prototypen in eine nachhaltige und breit nutzbare Open-Source-Lösung zu überführen. Während in der ersten Phase die technische Umsetzung im Vordergrund stand, lag der Fokus nun auf Stabilität, Nutzerzugang, Community-Aufbau sowie der strategischen Weiterentwicklung des Projekts.

Ausgangspunkt war die Erkenntnis, dass der langfristige Erfolg von MLOX nicht allein von der technischen Qualität abhängt, sondern maßgeblich durch eine aktive Nutzer- und Entwicklergemeinschaft sowie durch reale Anwendungsszenarien bestimmt wird. Ziel war es daher, MLOX von einem funktionierenden Prototypen zu einer anschlussfähigen Plattform mit wachsender Nutzung und klarer Perspektive weiterzuentwickeln.

Die geplante Vorgehensweise orientierte sich an mehreren ineinandergreifenden Meilensteinen. Ein zentraler Schwerpunkt lag auf der Skalierung der Community, insbesondere durch den gezielten Ausbau von Kommunikationskanälen wie Mastodon, LinkedIn und weiteren Plattformen sowie durch aktiven Outreach. Parallel dazu sollte die Sichtbarkeit des Projekts durch Fachveröffentlichungen, Konferenzbeiträge und praxisnahe Inhalte wie Tutorials und Demonstrationen erhöht werden.

Ein weiterer Fokus lag auf der Pilotierung und dem Ausbau realer Anwendungsszenarien. Das bestehende Startup-Pilotprojekt sollte weiterentwickelt und durch zusätzliche Partnerschaften ergänzt werden, um MLOX unter realen Bedingungen zu validieren und kontinuierlich zu verbessern.

Ergänzend dazu wurde der Aufbau nachhaltiger Strukturen adressiert, etwa durch die Vorbereitung organisatorischer Rahmenbedingungen für langfristigen Support und mögliche Dienstleistungen. Gleichzeitig war vorgesehen, die Dokumentation systematisch auszubauen und durch Best Practices sowie eine klar formulierte Roadmap die Weiterentwicklung des Projekts transparent zu machen.

Insgesamt zielte die Second Stage darauf ab, MLOX als skalierbare Open-Source-Plattform mit wachsender Nutzerbasis und klarer Zukunftsperspektive zu etablieren.

5. Ausführliche Darstellung der Ergebnisse in der Second-Stage-Förderung

In der Second Stage wurde MLOX sowohl technisch als auch strukturell deutlich weiterentwickelt. Die Plattform wurde stabilisiert, bestehende Komponenten verbessert und zusätzliche Integrationen ergänzt. Parallel dazu wurden die Entwicklerwerkzeuge, insbesondere CLI und TUI, erweitert, um den produktiven Einsatz zu erleichtern. Ein Demo-Modus ermöglicht inzwischen einen besonders einfachen Einstieg.

Ein zentraler Schwerpunkt lag auf dem Aufbau einer Community. Unter dem Namen „Busy Sloths“ wurden verschiedene Kanäle etabliert. Besonders hervorzuheben ist der Mastodon-Account „drbusysloth“, der als zentrales Medium für den Austausch mit der Community diente. Über diesen Kanal wurden kontinuierlich Einblicke in die Entwicklung von MLOX sowie Diskussionen zu MLOps und souveräner KI-Infrastruktur geführt.

Ergänzend dazu wurde ein Blogbeitrag im Rahmen des Prototype Funds veröffentlicht, der die gesellschaftliche Relevanz souveräner KI-Infrastrukturen adressiert und MLOX in diesen Kontext einordnet.

Parallel zur Community-Arbeit wurde die wissenschaftliche Verwertung vorangetrieben. Die Ergebnisse des Projekts wurden in der Veröffentlichung „MLOX: Open-Source MLOps for the Rest of Us“ zusammengefasst und auf der internationalen Konferenz CAIN 2026 im Rahmen der ICSE 2026 präsentiert. Diese Veröffentlichung trägt zur internationalen Sichtbarkeit des Projekts bei und positioniert MLOX im Kontext aktueller Forschung zu AI Engineering.

Darüber hinaus wurden gezielt Kooperationen mit wissenschaftlichen Partnern aufgebaut. Gespräche mit der Machine-Learning-Gruppe der TU Kaiserslautern sowie mit der Universität Bamberg im Bereich Security und Privacy eröffnen Perspektiven für den Einsatz von MLOX in der Lehre sowie für gemeinsame Forschungsaktivitäten. Ergänzend fanden Gespräche mit dem German Center for Open Source AI statt, um mögliche Integrationen und Nutzungsszenarien zu evaluieren.

Auch im Bereich Pilotierung und Partnerschaften wurde die Arbeit aus der ersten Phase weitergeführt. Das bestehende Startup-Pilotprojekt wurde ausgebaut und diente weiterhin als zentrale Referenz für reale Anwendungsfälle.

Insgesamt konnte MLOX in dieser Phase erfolgreich von einem Prototypen zu einer strukturierten, sichtbaren und anschlussfähigen Open-Source-Plattform weiterentwickelt werden.

6. Zielgruppe, Nutzen und mögliche Weiterentwicklungen

MLOX richtet sich insbesondere an Startups, kleine und mittlere Unternehmen, Freelancer sowie Forschungseinrichtungen. Diese profitieren von einer deutlich reduzierten Einstiegshürde in den Aufbau und Betrieb von ML-Infrastruktur sowie von einer hohen Transparenz und Kontrolle über ihre Systeme.

Durch die Open-Source-Verfügbarkeit entsteht zusätzlicher Mehrwert durch Wiederverwendbarkeit, Erweiterbarkeit und gemeinschaftliche Weiterentwicklung. Die Kombination aus realen Pilotprojekten, wissenschaftlicher Veröffentlichung und aktiver Community-Arbeit führt zu einer engen Verzahnung von Forschung, Praxis und Open Source.

Darüber hinaus kann MLOX als grundlegender Baustein für weitere Open-Source-Initiativen und Projekte dienen. Insbesondere im Kontext von Initiativen wie dem German Center for Open Source AI (GC.OS) ergeben sich Potenziale, MLOX als Infrastruktur- oder Integrationsschicht für andere Projekte einzusetzen und so zur Stärkung eines offenen und interoperablen AI-Ökosystems beizutragen.

Für die Weiterentwicklung sind insbesondere der Ausbau der Kubernetes-Integration, die Einführung von Multi-User-Funktionalitäten sowie eine stärkere Automatisierung und Ausbau des Plugin Systems vorgesehen. Eine zentrale Rolle wird dabei das Thema Sicherheit und Datenschutz spielen. Darüber hinaus bestehen Perspektiven für nachhaltige Strukturen, etwa durch die Gründung einer Organisation zur Unterstützung von Betrieb, Support und Beratung.

7. Persönliche und fachliche Weiterentwicklung

Die Arbeit an MLOX hat sowohl fachlich als auch persönlich zu einer deutlichen Weiterentwicklung beigetragen. Neben der Vertiefung im Bereich MLOps-Infrastruktur und Systemdesign konnten insbesondere Kompetenzen im Bereich Open Source, Community Building und Produktentwicklung ausgebaut werden. Darüber hinaus wurden erste Schritte in Richtung unternehmerischer Verwertung unternommen.

8. Kurze Darstellung der Arbeiten, die zu keiner Lösung geführt haben

Im Verlauf des Projekts wurden verschiedene Ansätze erprobt, die sich als nicht zielführend erwiesen. Dazu zählen insbesondere komplexere Architekturvarianten sowie alternative Integrationen, die nicht ausreichend nutzerzentriert waren. Stattdessen wurde der Fokus bewusst auf Stabilität, Einfachheit und praktische Nutzbarkeit gelegt.

9. Präsentationsmöglichkeiten für mögliche Nutzer*innen

Webseite: https://mlox.org
GitHub Repository: https://github.com/BusySloths/mlox
Community: https://github.com/BusySloths
Mastodon: https://mastodon.social/@drbusysloth
Einführungsvideo: https://youtu.be/GufZagzIKzY
Dokumentation: https://busysloths.github.io/mlox/docs/mlox.html
Weiterführende Dokumentation: https://github.com/BusySloths/mlox/tree/main/docs
Presentationsfolien: https://www.slideshare.net/slideshow/mlox-open-source-mlops-for-the-rest-of-us/286942159
Wissenschaftliche Veröffentlichung: CAIN / ICSE 2026 (https://doi.org/10.1145/3793653.3793795)

10. Kurze Erläuterung zur Einhaltung der Arbeits- und Kostenplanung

Das Projekt wurde im Wesentlichen im Rahmen der ursprünglichen Planung umgesetzt. Anpassungen ergaben sich vor allem durch den höheren Aufwand bei Integration und Stabilisierung sowie durch die zunehmende Bedeutung des Community-Aufbaus in der Second Stage. Diese Anpassungen waren notwendig, um die Nachhaltigkeit des Projekts sicherzustellen.

11. Kurze Darstellung von etwaigen Ergebnissen bei anderen Stellen

Parallel zum Projekt ist eine zunehmende Dynamik im Bereich MLOps und AI-Infrastruktur zu beobachten. Insbesondere Cloud-basierte Lösungen bestätigen den Bedarf, verdeutlichen jedoch gleichzeitig Herausforderungen wie Komplexität, Kosten und Abhängigkeiten. Vor diesem Hintergrund zeigt sich die Relevanz offener, integrierter Ansätze wie MLOX.

Dietramszell, den 15.04.2026

