## Anforderungen der Regulatorik für KI-Innovationen systematisiert sicherstellen

# AI Compliance mit MLOps

Sonderdruck aus
BI-SPEKTRUM 5/2022

**Die Mensch-Maschine-Interaktion wird in Zukunft stärker aus einer kollaborativen Perspektive betrachtet werden. Dazu gehört die organisatorische Kompetenz zur stabilen Bereitstellung von KI-Systemen für die eigenen Prozesse oder in neuen Produkten und Services wie auch um die Einhaltung von komplexen Regularien und Normen. In diesem Artikel wird dargelegt, welche Qualitätsaspekte eine MLOps-Architektur strukturell erfüllen kann.**

Ein Beitrag von
Robert Kasseck,
Lilian Do Khac und
Oleg Smolanko

Die Fähigkeiten und somit Einsatzszenarien von Künstlicher Intelligenz (KI) entwickeln sich kontinuierlich weiter und holen sich selbst dynamisch Schlag um Schlag ein. Die Managementebene steht dem Druck eines Wandels zu einer datengetriebenen Organisation gegenüber. Die letzten Jahrzehnte der industriellen Revolution waren geprägt von Spezialisierung auf einzelne Aufgaben und damit einhergehend die Aufgabenteilung. So wurde auch in der Dritten Industriellen Revolution die Aufgabenteilung mit der Einführung von Maschinen weitergetrieben.

Es ist also nicht verwunderlich, dass in den letzten 50 Jahren die Mensch-Maschine-Interaktion schwerpunktmäßig aus einer rein technischen Sicht getrieben wurde. Nun spielen mit Einzug der Vierten und Fünften Industriellen Revolution verstärkt kollaborative Fähigkeiten mit Maschinen zusammen eine Rolle. Dazu gehören die organisatorische Kompetenz zur stabilen Bereitstellung von KI-Systemen für die eigenen Prozesse oder in neuen Produkten und Services wie auch die Einhaltung von komplexen Regularien und Normen. Wer sich hier nicht gut aufstellt, sowohl technologisch als auch auf der personellen und organisatorischen Ebene, wird in naher Zukunft einen massiven Wettbewerbsnachteil erfahren.

Der Weg zu einem datengetriebenen Unternehmen ist mehrschichtig und umfasst alle Systemebenen (siehe Abbildung 1). Auf der Makro-Ebene stellt sich die Frage, wie dieser Wandel gestaltet


Bild: Shutterstock

werden kann. Denn die BI-Systeme aus den vergangenen Jahrzehnten bauten auf silohaften Unternehmensstrukturen auf, und so verhält es sich zumeist auch mit der Datenlandschaft. Historisch bedingt sind IT-Abteilungen, also diejenigen, die für die Implementierung und Wartung von Software zuständig (im Sinne von Responsible), jedoch nicht verantwortlich (im Sinne von Accountable) sind, der Finanzabteilung untergeordnet. Der Ursprung liegt zum großen Teil in der Notwendigkeit, Transparenzanforderungen für und über das Management zu erfüllen, um das Unternehmen steuern zu können. Diese Strukturen müssen aufgebrochen werden und Kompetenzen hinsichtlich datengetriebener Produkte breiter in die Organisation getragen werden.

Auf einer Meso-Ebene bilden sich eine Vielzahl weiterer Herausforderungen heraus, die auf aktuelle gesetzliche Anforderungen (beispielsweise EU-KI-VO), Normen und Standards (wie VDE SPEC 90012 oder dem amerikanischen AI Risk Management Framework), aber auch den Anspruch nach Qualität und Vertrauen (zum Beispiel durch Bestehen von diversen Zertifizierungen wie dem von CertAI) zurückzuführen sind. Somit soll über die Meso-Ebene eine nachhaltige Wettbewerbsfähigkeit durchgesetzt werden. Zur Unterstützung der Meso-Ebene bieten sich die Einführung und Umsetzung von MLOps-Tools und Prozessen auf der Mikro-Ebene an.

## MLOps zur Unterstützung der Dokumentations- und Qualitätspflichten nach VDE SPEC 90012 V1.0 und der EU-KI-VO

Machine Learning Operations (MLOps) ist eine Weiterentwicklung von DevOps (siehe Abbildung 2). Das Ziel von DevOps war es, die Risikolücke zwischen der Entwicklung von IT-Artefakten und ihrer Bereitstellung und Übernahme in die Geschäftslinie für den Betrieb zu verringern. Die Kernprinzipien von DevOps sind Continuous Integration und Continuous Delivery (CI/CD), um Kontinuität und Reproduzierbarkeit sowie Automatisierung sicherzustellen. MLOps ist ein Framework, das klassisches Software Engineering, Machine Learning und Data Engineering kombiniert. Es fügt dem bestehenden DevOps-Paradigma neue Prinzipien wie ML-Reproduzierbarkeit, ML-Versionierung, kontinuierliches ML-Training und -Evaluierung, ML-Metadatenverfolgung und kontinuierliche Modellverbesserungen hinzu [KKH22]. In Abbildung 2 sind die klassischen funktionalen Anforderungen aus Sicht einer modernen Datenmanagement-Plattform in Dunkelgrau abgebildet. Diese sind ebenso wichtig für ML-Prozesse, denn sie umfassen die saubere, kontinuierliche und ef-

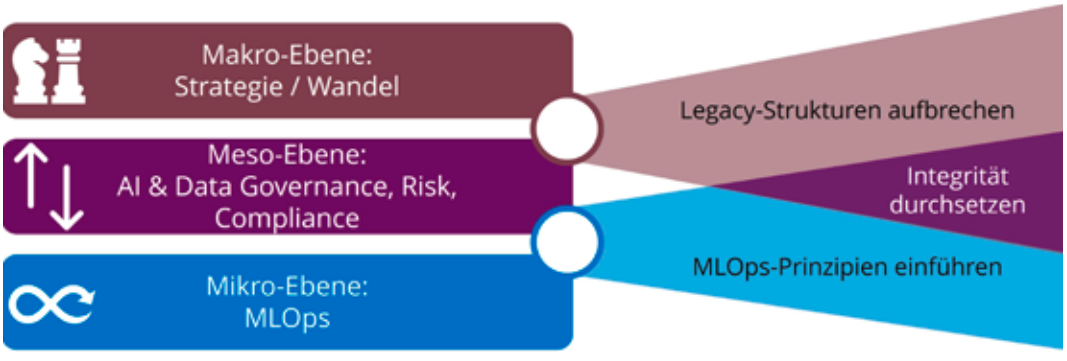


**Abb. 1:** Der Weg zu einem AI-Driven-Unternehmen ist mehrschichtig




**ROBERT KASSECK** ist Data Scientist im Bereich Data & Analytics. Er ist für die technische Planung und Realisierung von ML-basierten Geschäftsmodellen verantwortlich. Mit seiner Begeisterung für Datenanalyse und Machine Learning gestaltete er neue innovative Denkweisen mit und zeigte Optimierungen für eine datengetriebene Produktentwicklung auf.
**E-Mail: Robert.Kasseck@adesso.de**

**LILIAN DO KHAC** hat BWL mit der Vertiefung Operations Research and Management an der RWTH Aachen studiert. Sie beschäftigt sich mit der Konzeption und Implementierung von KI-Lösungen für die datengetriebene Entscheidungsunterstützung. In ihrer Promotion geht es um die bestmögliche Gestaltung von KI-Anwendungen.
**E-Mail: Lilian.Do-Khac@adesso.de**

**OLEG SMOLANKO** ist ein erfahrener Data Scientist mit mehr als sechs Jahren Berufserfahrung in verschiedenen Themengebieten. Seine Schwerpunkte liegen in der Analyse von Daten unter Anwendung verschiedener maschineller Verfahren. Unter anderem arbeitete er in zahlreichen Projekten mit dem Ziel, Vorhersagen zu erstellen sowie Business Insights zu generieren.
**E-Mail: Oleg.Smolanko@adesso.de**

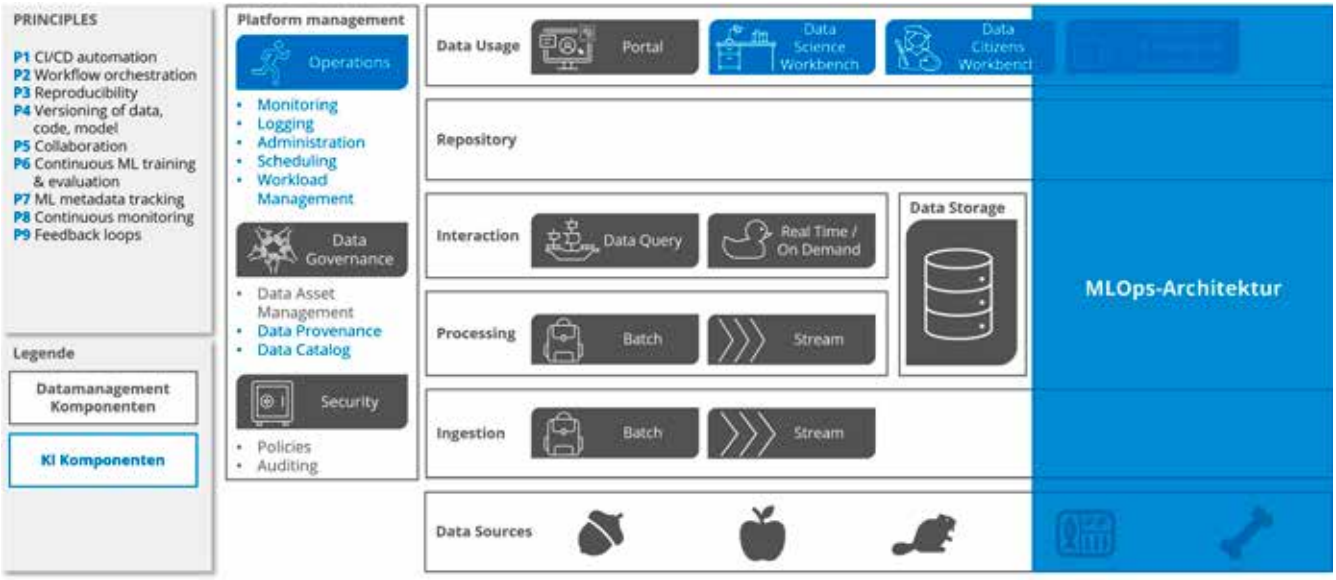


**Abb. 2:** MLOps-Architektur im Zusammenhang mit klassischer Datenmanagement-Architektur, funktionale Sicht

fektive Prozessierung von Daten. Insbesondere die blau hinterlegten Komponenten unterscheiden sich nicht hinsichtlich Datenmanagementprozessen und ML-Prozessen. Aus einer reinen Datenmanagement-Plattform-Sicht fällt die Prozedur rund um die Erstellung und Lieferung von ML-Anwendungen jedoch zu kurz aus (siehe Abbildung 2 „Data Science Workbench“ oder „Data Citizens Workbench“ und Vergleich Abbildung 3 rechts). Dieser Aspekt wird später detaillierter beleuchtet.

Die Einführung des EU-Gesetzes zur Künstlichen Intelligenz (EU-KI-VO) ist das erste konkrete Bestreben nach einer regulatorischen Manifestation weltweit. Sie fordert die Erfüllung vertrauenswürdiger KI-Anforderungen in Organisationen [EC21]. Die EU-KI-VO umfasst hinsichtlich Dimensionierung und Auswirkungen die DSGVO und geht hinsichtlich der Compliance-Anforderungen bei weitem über die DSGVO hinaus. Die EU-KI-VO wird die Durchsetzung von Rechenschaftspflicht und Verantwortung entlang des gesamten KI-Lebenszyklus fordern. Dies impliziert die Notwendigkeit, entsprechende Sicherungsprozesse, Rollen und Berichtsstrukturen als Ergänzung zu einer organisatorischen KI-Governance zu implementieren.

Die Erfüllung der EU-KI-VO ist besonders herausfordernd, da sie die Einhaltung ethischer Grundsätze verlangt, die im Wesentlichen unscharfe und latente Konzepte sind. Die VDE SPEC 90012 V1.0 stellt einen Rahmen dar, der speziell konzipiert wurde, um KI-Systeme an der Erfüllung vertrauenswürdiger KI-Prinzipien zu messen und damit zu bewerten. Sie bietet sich daher für eine prospektive Zertifizierung vertrauenswürdiger KI-Anwendungen an [VDE22].

Die Umsetzung vertrauenswürdiger KI-Anwendungen schließt den gesamten Datenbewirtschaftungszyklus mit ein. Das heißt, dass die oben genannten Notwendigkeiten aus der EU-KI-VO heraus zusammen mit der bestehenden Datenmanagement-Plattform bewerkstelligt werden müssen (siehe Abbildung 3 gelbe Posts). In Abbildung 3 sind einige Kriterien der VDE SPEC 90012 V1.0 der jeweiligen Stelle der funktionalen Gesamtbetrachtung von Datenmanagement und MLOps zugeordnet. Werteanforderungen wie Fairness (F), Transparenz (T) oder Zuverlässigkeit (Reliability = R) können direkt an der MLOps-Architektur verortet und erfüllt werden.

Das Kriterium F1 – Fairness kann erzielt werden durch den Ansatz „by Design“ und im Rahmen von dedizierten Modelltests, die standardisiert in der MLOps-Prozessierung und Architektur verankert werden können. Weiterhin kann das Kriterium T3 – Interpretierbarkeit (Intelligibility) durch standardisierte Prozesse in der MLOps-Prozedur verankert werden, indem nicht nur ein Modell für das zugrunde liegende Problem trainiert wird, sondern – sofern das Modell nicht intrinsisch erklärbar ist – auch ein zweites Modell, um die Interpretierbarkeit des Anwendungsmodells zu steigern. Weitere Kriterien betreffen Anforderungen an die Betriebsqualität von ML-Anwendungen (siehe R2 und T2). Diese können im Plattform-Management zusammen mit den sonstigen Komponenten für den operativen Bereich erfüllt werden.

Die bestehenden Abläufe für BI-Prozesse müssen lediglich um die ML-spezifischen Anforderungen erweitert werden. Andere Kriterien sind jedoch eher ein traditionelles Datenmanagement-Thema, wie die Dokumentationsanforderungen an Datensets (siehe T1) und der Schutz von personenbezogenen Daten (siehe P2). Andere Aspekte wie die ökologische Nachhaltigkeit (siehe F3) oder das Einwilligungsverfahren (siehe P3) können nicht dediziert zugeordnet werden, betreffen allerdings beide Strukturen gleichermaßen. Mit einer sauber aufgesetzten MLOps und entsprechend explizit umgesetzten Kriterien können an den zuvor beschriebenen funktionalen Stellen somit herausragende ML-Entwicklungsqualität und Auditergebnisse erzielt werden und so zu einer hohen strukturellen KI-Compliance führen.

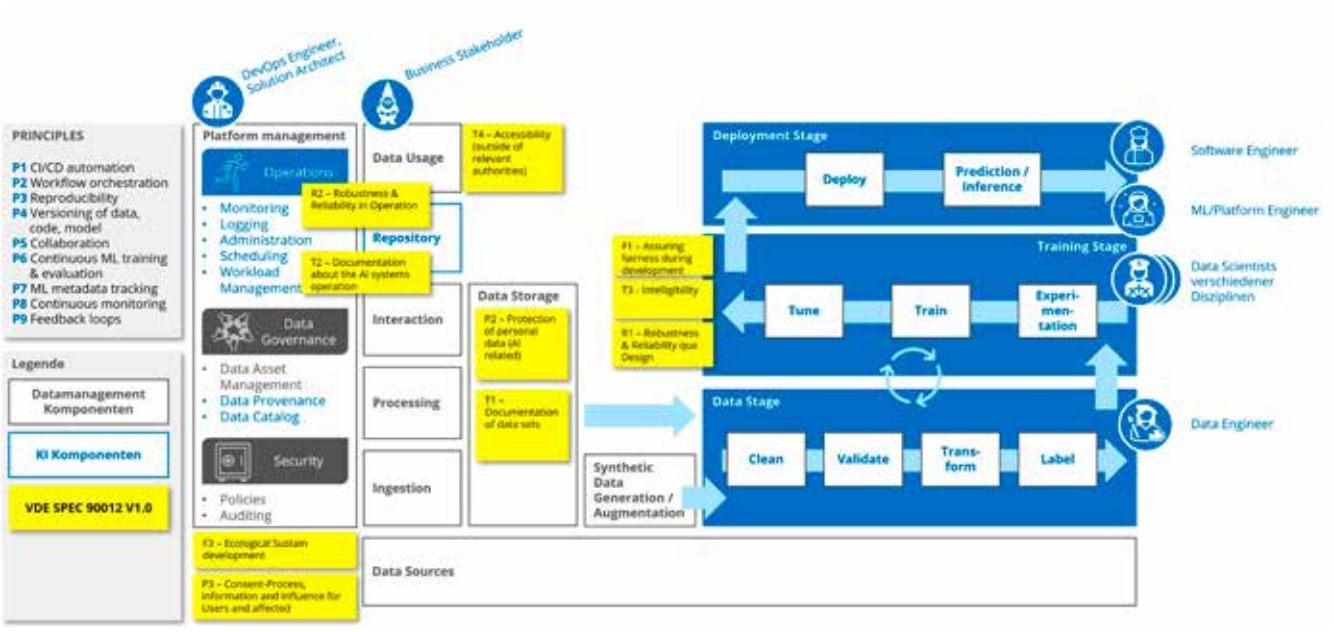


**Abb. 3:** Grobe Übersicht über die funktionalen Bestandteile einer MLOps-Architektur (in Anlehnung an [Jef22])

## Praxisbeispiel einer MLOps-Implementierung

Im Folgenden wird anhand einer Beispielimplementierung ein Praxisbeispiel dargestellt und erläutert, wie die jeweiligen Rollen für die Entwicklung von ML-Anwendungen integriert sind.

Bei der Umsetzung einer KI-Architektur, in der ein vollständiger MLOps-Prozess implementiert werden kann, muss zuerst das Anwendungsgebiet der KI bestimmt werden. Dies kann grundlegend in drei Bereiche gegliedert werden (siehe Abbildung 4): KI für Analytik, KI für Produkte und KI als Produkt. KI für Analytik wird während der Datenaufbereitung für Business-Reports oder für Ad-hoc-Analysen angewendet. Hier sind in der Regel einige bis alle Prozesse zur Auslieferung von Softwareartefakten und der Zeitpunkt der Anwendung/Vorhersage durch bestehende Anwendungen vorgegeben. Für den Einsatz von KI in Produkten gelten wiederum andere Anforderungen: Zum einen ist das Produkt der Erzeuger und Lieferant der Daten, mit denen die KI entwickelt wird, aber auch Konsument der entwickelten KI. Hier besteht die Herausforderung, diese zweiseitige Beziehung zu warten, Schnittstellen zu definieren und unterschiedliche Entwicklungsmethoden – DevOps versus MLOps – zu verwalten. Im dritten Anwendungsbereich von KI als Produkt besteht die Herausforderung darin, die Qualität der KI aus dem Labor in der operativen Umgebung sicherzustellen. Anders als bei den anderen Bereichen kann schwankende Qualität, die durch sich ändernde neue Daten entsteht, nicht durch die eigentliche Anwendung korrigiert werden.

Dieser Vergleich der drei Bereiche ist bei weitem nicht vollständig und dient hier lediglich zur Veranschaulichung der unterschiedlichen Herausforderungen, denen KI selbst ohne den Aspekt Regulatorik unterliegt.

In einem Projekt wurde die Herausforderung in Angriff genommen, mehrere KI-Komponenten für ein Produkt zu erstellen, reproduzierbar, erklärbar und nach dem MLOps-Prozess zu entwickeln und


**Abb. 4:** Verschiedene Arten von KI

zu überwachen. Als zusätzliche Herausforderung war man konfrontiert mit einem On-Premises-Deployment bei unterschiedlichen Kunden, sodass kein direkter Feedback Loop implementiert werden konnte. Glücklicherweise ist ein Datenaustausch über Drittsysteme möglich, was ein indirektes Überwachen der Modelle ermöglicht.

Die Architektur wurde aus vier grundlegenden Komponenten aufgebaut:

- Einem Object Store, in dem drei Bereiche zur Datenhaltung konfiguriert sind.
- Die Landing Zone ist Empfänger aller Rohdaten (historisiert) aus den Drittsystemen.
- In der Interim Zone werden alle aggregierten, bereinigten und kuratierten Daten in der letzten validen Version für Analysten und die Rückverfolgung der Erstellung des Datensets vorgehalten.
- In der letzten Zone werden Datensets historisiert, die in Experimenten sowie automatisierten Trainingsläufen genutzt werden und eine neue Modellversion erzeugt hatten.

Durch dieses Vorgehen kann im Zusammenhang mit einer Quellcode-Versionierung eine grundlegende Nachverfolgbarkeit der Datenabstammung sichergestellt werden. Um mit diesen Daten für Analysen, Experimente, Aufbereitung und automatisiertes Training zu interagieren, wurde Kubeflow auf einem Kubernetes-Cluster implementiert. Die Standard-Kubeflow-Umgebung wird durch MLFlow ergänzt, was eine bessere Verwaltung der Verfolgung von Experimenten und Modellversionen ermöglicht.

Grundlegend lässt sich die Arbeitsweise mit dem System wie folgt zusammenfassen. In den integrierten Notebook- oder CodeServer-Instanzen kann ein Data Scientist mit den direkt im Dateisystem eingebundenen Daten interagieren und experimentieren. Seine Erkenntnisse und seine Implementierung werden in dem gemeinsamen Git Repository hinterlegt. Die Arbeit basiert an dieser Stelle auf der GitFlow-Vorgehensweise. Nach der Experimentierphase des Data Scientist standardisiert ein ML-Engineer unterschiedliche Schritte. Dazu zählen Bereinigung, Feature-Erstellung, Training und Validierung in einer Python-Bibliothek. Diese Standards kann der Data Scientist auch in zukünftigen Experimenten wiederverwenden, wodurch die Entwicklung beschleunigt wird.

Nach der Standardisierung werden alle Bestandteile von Daten-Aufbereitung und Model-Training in Kubeflow-Pipelines automatisiert und neue Modelle in MLFlow nachverfolgbar hinterlegt. In diesen Pipelines erfolgt auch die nachgelagerte Überwachung der Modelle. Alle Bestandteile der Standardisierung und Datenaufbereitung werden durch automatisierte Unit-Tests und automatische Code-Überprüfungen qualitätsgesichert. Für die letzte Komponente in der Architektur wurde ein FastAPI-Service entwickelt, durch den die Modelle für die eigentliche Anwendung bereitgestellt werden. Auch hier werden alle genannten Schritte zur Qualitätssicherung angewendet.

Daneben besteht auch die Anforderung, die Vorhersagen der Modelle erklärbar und für den Anwender deutbar zu machen. Maximal transparent wird die Erklärung erst, wenn man sich möglichst nahe an dem verwendeten Modelltyp orientiert. Zum aktuellen Projektstatus werden die einflussreichsten Faktoren benannt, anhand derer ein Modell eine Entscheidung getroffen hat. Die Entwicklung dieser Erklärung durchläuft denselben Zyklus wie die Modellentwicklung und wird mit dem Service an die Anwendung ausgeliefert.

## Zusammenfassung

MLOps ist eine Weiterentwicklung von DevOps. Das Ziel von DevOps war es, die Risikolücke zwischen der Entwicklung von IT-Artefakten und ihrer Bereitstellung und Übernahme in den Betrieb zu verringern. Die MLOps-Architektur ist in der Lage, die Werteanforderungen Fairness, Transparenz und Zuverlässigkeit zu erfüllen. Dabei müssen Aspekte der VDE SPEC 90012 erfüllt werden, was mittels einer MLOps-Architektur möglich ist. Der Weg zu einem datengetriebenen Unternehmen ist allerdings mehrschichtig und kann nicht nur über eine MLOps-Architektur gestemmt werden. Um jedoch zukünftig nachhaltig wettbewerbsfähig zu bleiben, ist die Einführung von MLOps-Strukturen mehr als nur ein Nice-to-have.

## Literatur

**[DIN19]** DIN SPEC 92001-1: Artificial Intelligence – Life Cycle Processes and Quality Requirements – Part 1: Quality Meta Model

**[EC21]** Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. COM (2021) 206 final, 21.4.2021, https://ec.europa.eu/newsroom/dae/redirection/document/75788, abgerufen am 4.11.2022

**[Jef22]** Jeffries, D.: Why we started the AIIA and what it means for the rapid evolution of the canonical stack of machine learning. 7.1.2022, https://ai-infrastructure.org/why-we-started-the-aiia-and-what-it-means-for-the-rapid-evolution-of-the-canonical-stack-of-machine-learning/, abgerufen am 4.11.2022

**[KKH22]** Kreuzberger, D. / Kühl, N. / Hirschl, S.: Machine Learning Operations (MLOps): Overview, Definition, and Architecture. 2022, arXiv preprint arXiv:2205.02302

**[VDE22]** VDE SPEC 90012 V1.0: VCIO based description of systems for AI trustworthiness characterization, 25.4.2022, https://www.vde.com/resource/blob/2176686/a24b13db01773747e6b7bba4ce20ea60/vde-spec-vcio-based-description-of-systems-for-ai-trustworthiness-characterisation-data.pdf, abgerufen am 4.11.2022