Mit Blockchain und Federated Learning zur unternehmensübergreifenden Predictive Maintenance

Marisa Mohr, Christian Becker

20. Oktober 2020

www.marisa-mohr.de

Marisa Mohr

Machine Learning Engineer @ inovex

- Team Künstliche Intelligenz
- IIot und Predictive Maintenance

Externe Doktorandin @ Uni Lübeck

- Institut für Informationssysteme, Prof. Dr. Ralf Möller
- Thema: Time Series Representation Learning

inovex BLOG

Christian Becker

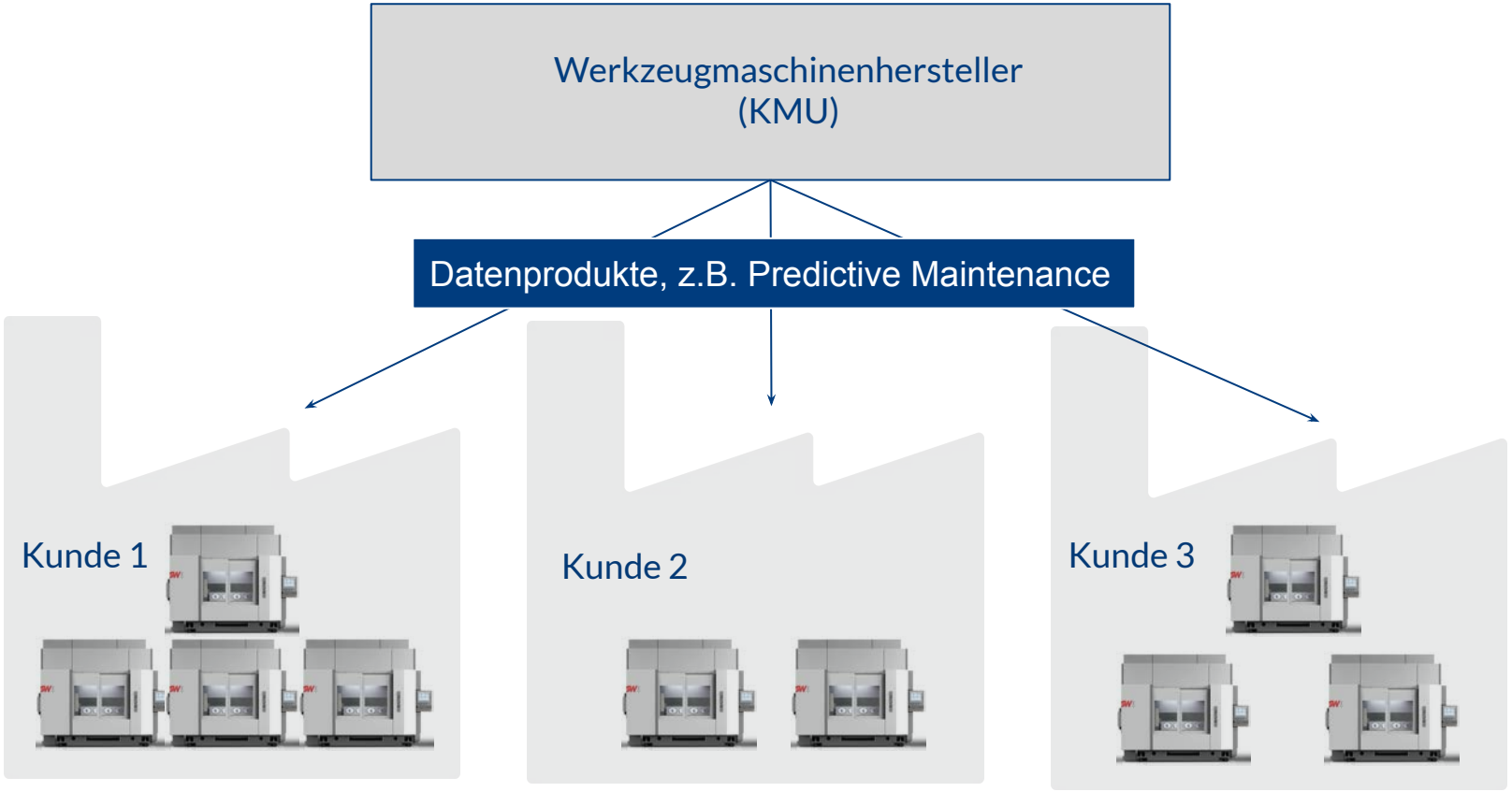
Werkstudent @ inovex

- Team Data Management and Analytics
- Bachelorarbeit “Evaluation of Federated Learning in Deep Learning”

Masterand @ Hochschule Karlsruhe

- Fakultät für Informatik und Wirtschaftsinformatik
- Vertiefung Maschinelles Lernen

Datenprodukte im IIoT



inovex

KOSMoS

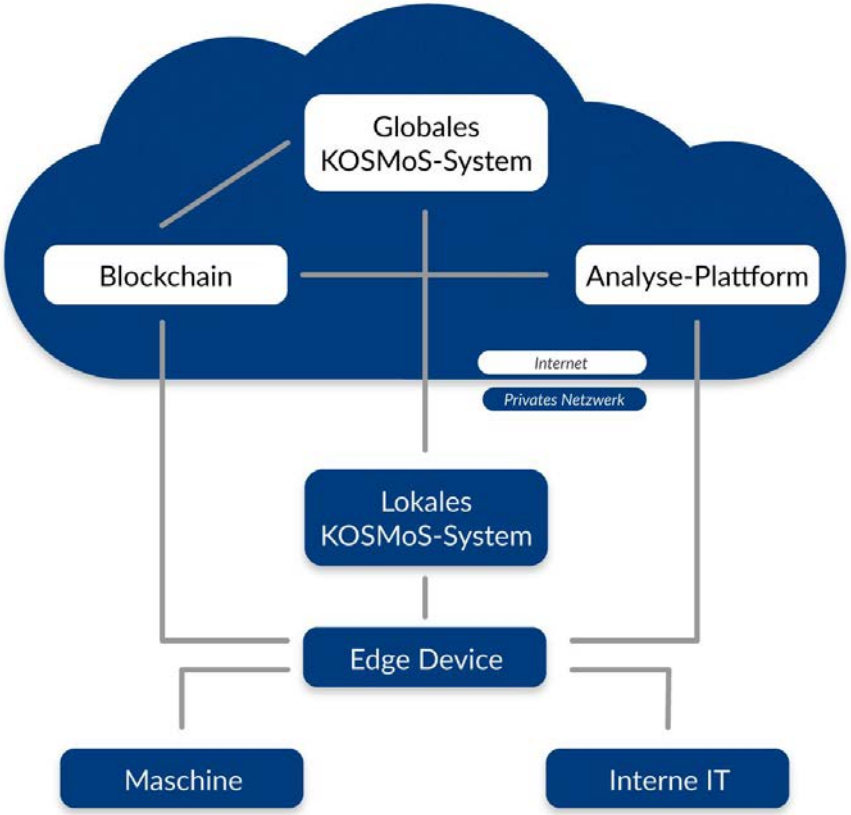
Bundesministerium für Bildung und Forschung

Kollaborative Smart Contracting Plattform für digitale Wertschöpfungsnetze

Maschinenhersteller	Enabler	Forschungsinstitute

- Framework für Systeme zum sicheren und semi-transparenten Austausch von produktions-beschreibenden Informationen zwischen kooperierenden Firmen
- Umsetzung von unternehmensübergreifenden datengetriebenen Geschäftsmodellen

KOSMoS Architektur



inovex

Herausforderung Predictive Maintenance

- **Verfügbare Trainingsdaten** bestimmen die Genauigkeit eines datengetriebenen Modells
- Daten müssen Ausfall-Muster enthalten
- **Relativ seltenes Auftreten von Maschinenausfällen**
- Absichtliche Degradierung von Maschinen zur Sammlung von Trainingsdaten ist unverantwortlich

Lösungsvorschlag

- Zusammenführen von Trainingsdaten über Unternehmen hinweg
- Einsatz von **Blockchain** und **Federated Learning** zur Wahrung der Privatsphäre von Daten

Collaborative Predictive Maintenance

Herausforderungen bei der Kombination von Datenquellen:

1. Datenmanipulation (z.B. zur Sabotage von Modellvorhersagen)
2. Transparente Dokumentation der Wartungen
3. Gemeinsames Training von Modellen für die vorausschauende Wartung ohne Offenlegung von Geschäftsinformationen

Was ist eine Blockchain?

DATARELLA

- **Dezentrales Register an Daten**, die chronologisch aufeinander aufbauen und durch einen **Konsensus-Mechanismus** abgesichert sind
- Jeder kann durch das Erstellen eines public/private key pairs teilnehmen
- Es gibt **keine zentrale Partei,** die die Blockchain verwaltet, jeder kann einen eigenen Node betreiben oder mit der Blockchain interagieren
- Daten in der Blockchain sind **unveränderlich**
- Daten in der Blockchain sind **transparent**

HYPERLEDGER

inovex

Blockchain: Schutz gegen Manipulation

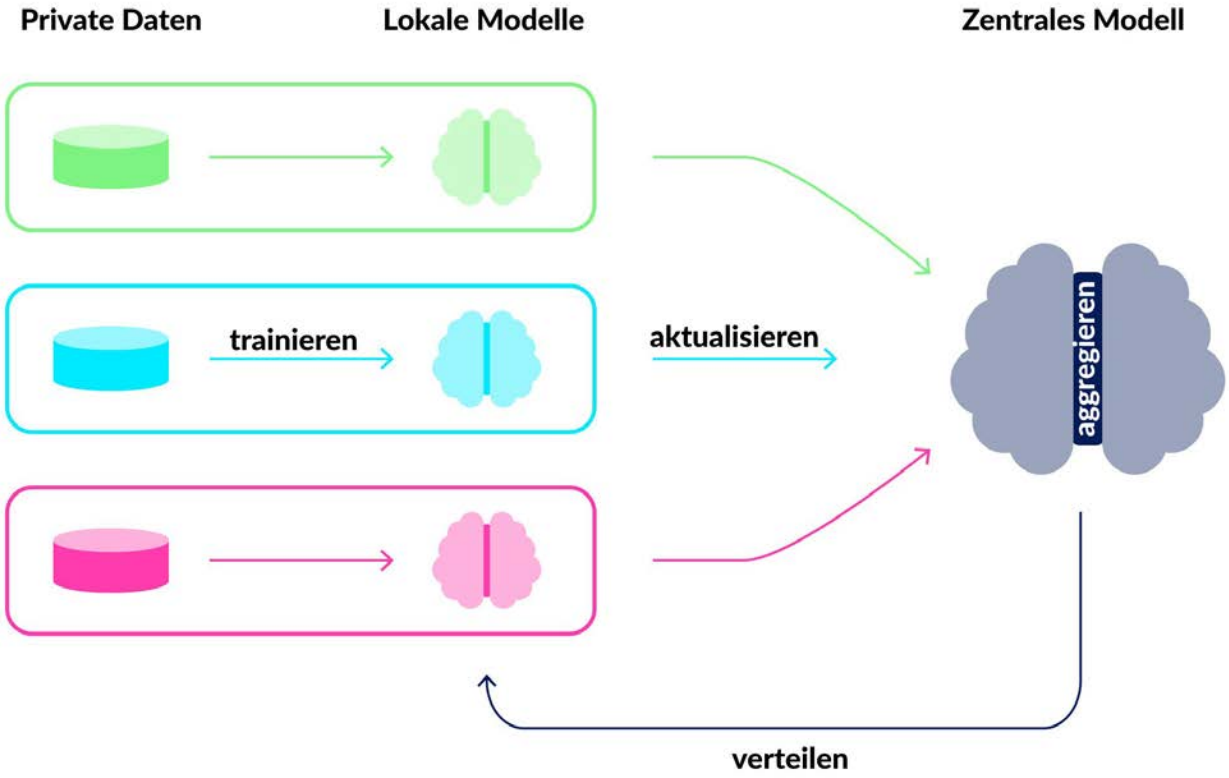
- Daten dort signieren, wo sie aufgezeichnet werden
- Kommerzielle Systeme sind i.d.R. zwischen verschiedenen Herstellern nicht kompatibel
- Installation eines Mikrocontrollers zwischen Sensor und nachgeschaltetem System (Korb et al., ISW Stuttgart)
- Implementierung einer **Blockchain** als gemeinsame Datenbasis, auf der geeignete Signaturverfahren eingesetzt werden können

Quelle: Korb, T.; Michel, D.; Riedel, O.; Lechler, A.: Securing the Data Flow for Blockchain Technology in a Production Environment. IFAC-PapersOnLine, 52(10):125–130, January 2019.

Blockchain: Transparente Dokumentation

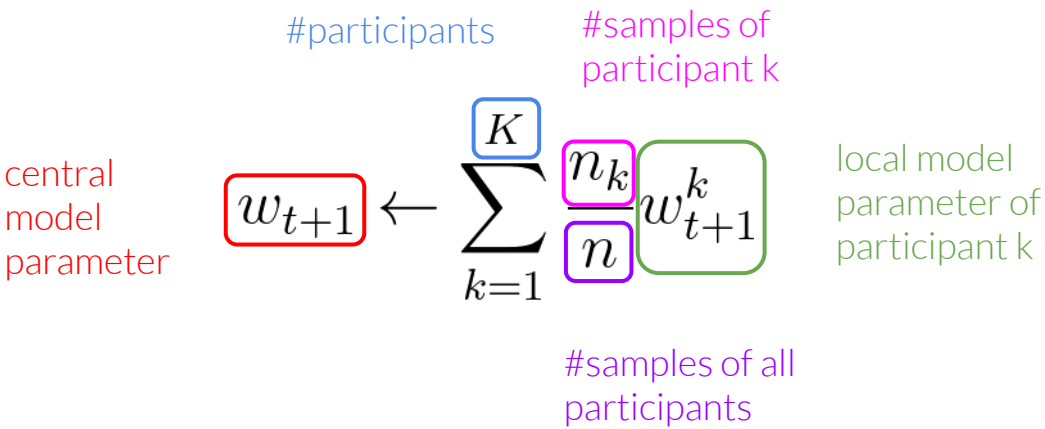
- Konsensus-Mechnismus: Alle Teilnehmer sind synchronisiert und validieren Transaktionen
- Transparentes Wartungsprotokoll auf der Grundlage von Fehlermeldungen und Wartungs-Einträgen
- Das Tupel eines Eintrags wird kryptographisch gehasht und der Hash wird in die Blockchain geschrieben
- Echtheit der Einträge, unberechtigte Manipulation und Vertrauen

Federated Learning mit privaten Daten



inovex

Aggregierung in Federated Learning



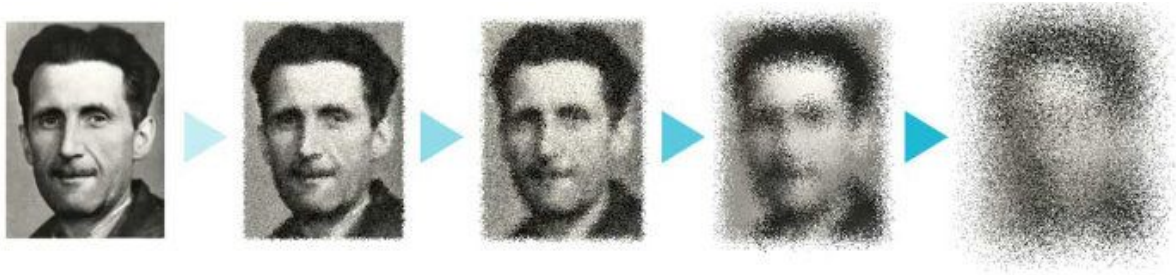
Quelle: Brendan McMahan, H., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data

Privacy in Machine Learning

- Trainingsdaten können aus dem Modell rekonstruiert werden: Verbreiten von geheimen Betriebs- und Produktionsdaten
- Verwendung von **Differential Privacy** im Training verhindert Rekonstruktion

Differential Privacy

- Differential Privacy verhindert, dass Trainingsdaten aus den Modellen generiert werden können
- Verrauscht Gradienten im lokalen Optimierungsprozess
- Beeinflusst durch das Rauschen eventuell die Qualität des Modells
- Begrenzt die Anzahl der möglichen Trainings-Epochen



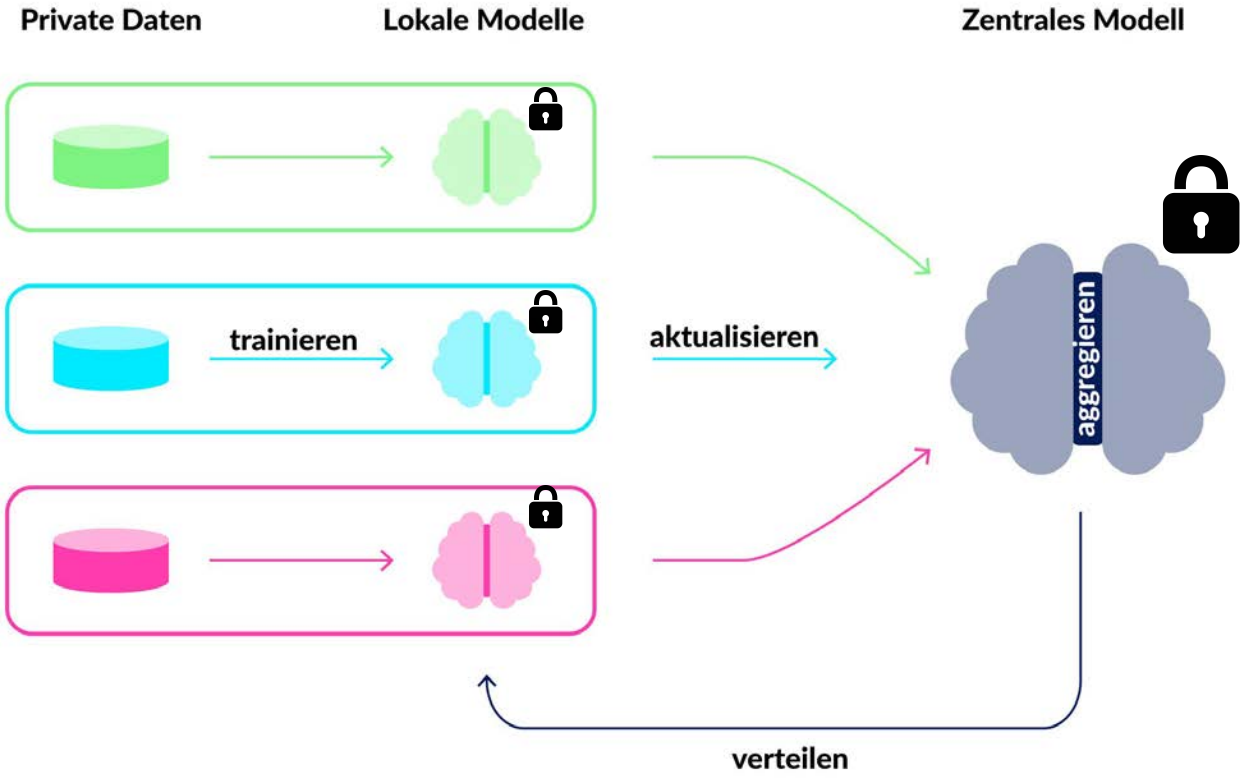
high utility
no privacy

high privacy
no utility

Quelle: https://aircloak.com/explaining-differential-privacy/

inovex

Federated Learning mit Differential Privacy



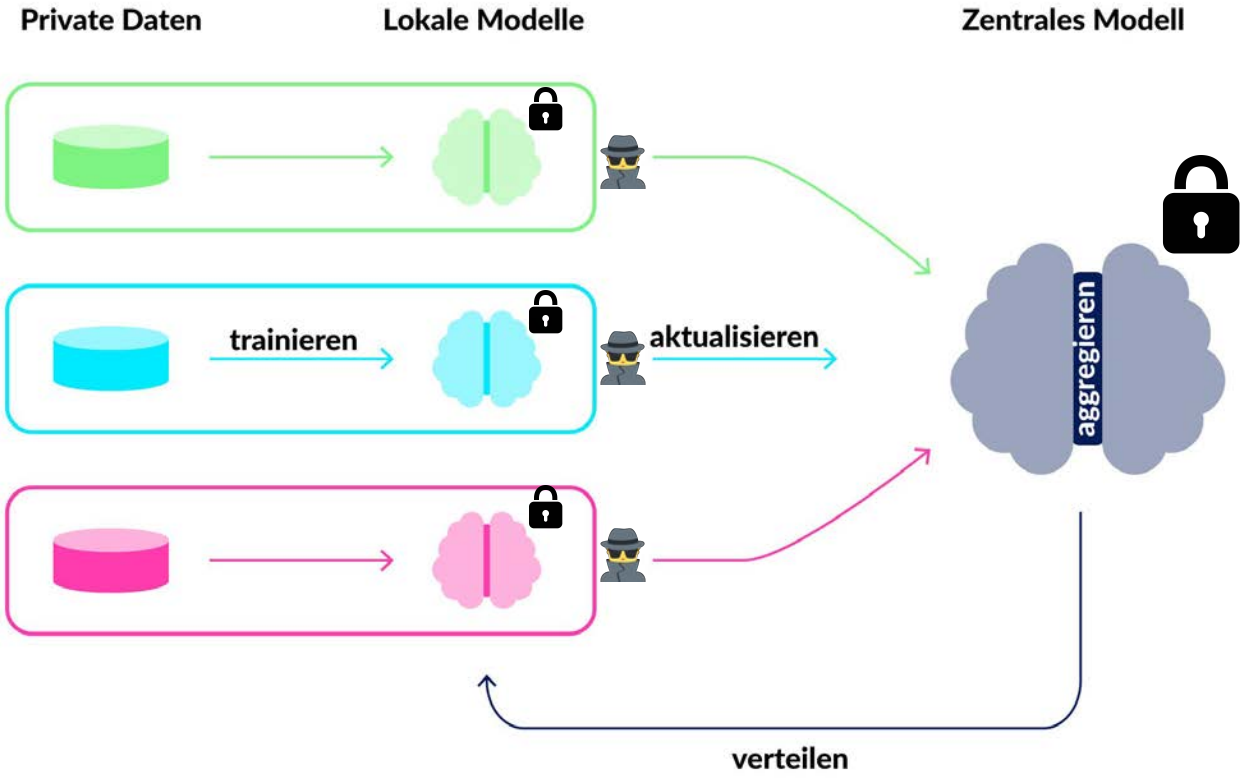
inovex

Secure Aggregation mit SMPC

- Secure Multi-Party Computation
- Ermöglicht das Rechnen mit verschlüsselten Zahlen!
- Beschränkt auf ganze Zahlen, Addition und Multiplikation

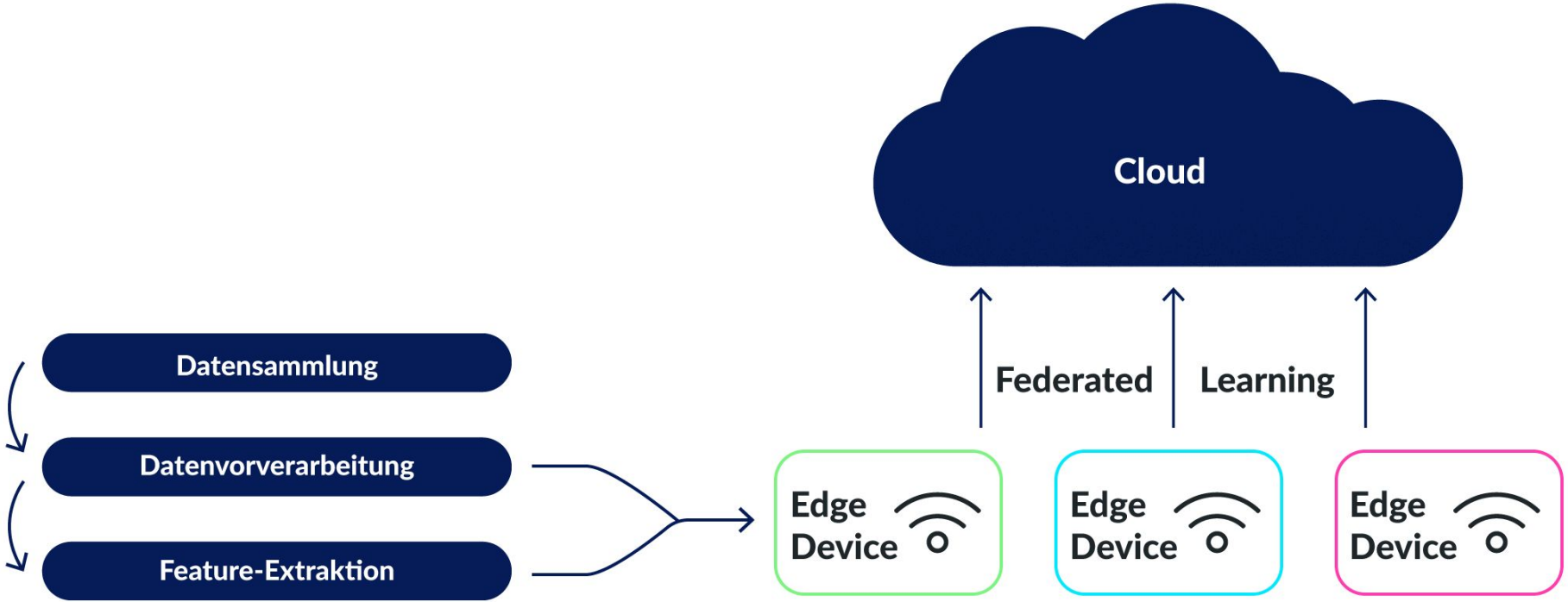
$$w_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^k$$

Federated Learning mit Secure Aggregation



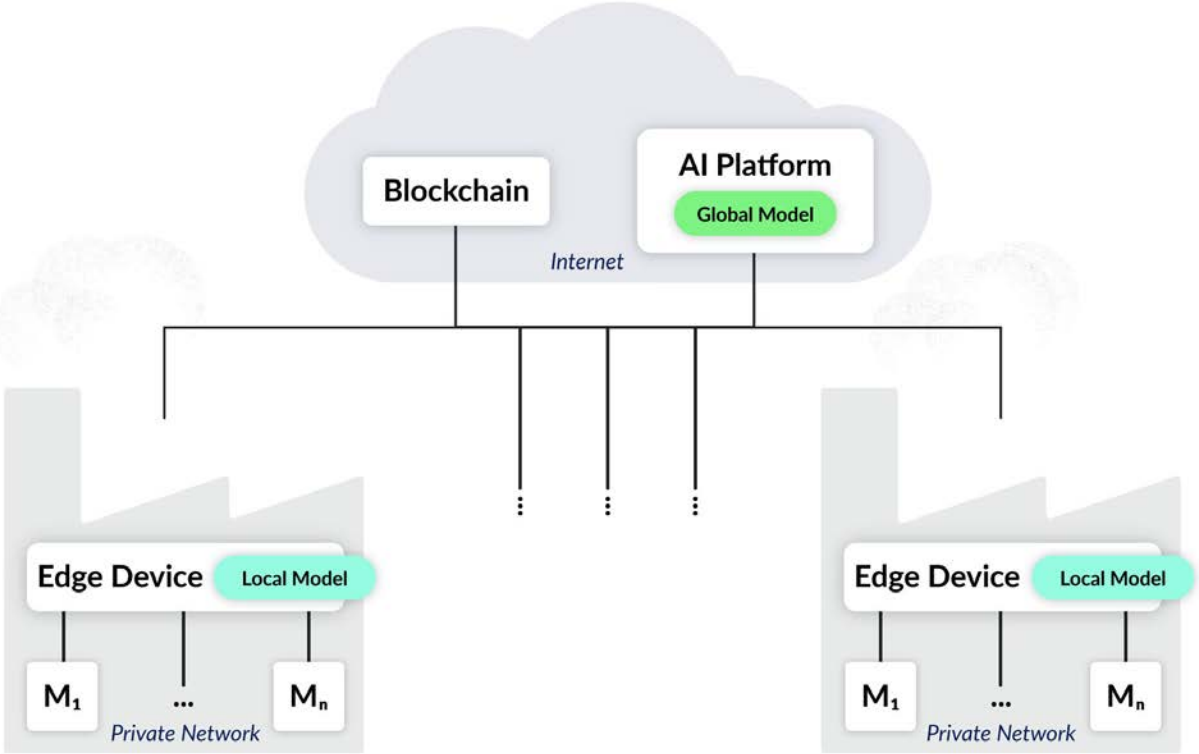
inovex

Lokale Datensammlung auf der Edge



inovex

Federated Learning in KOSMoS



inovex

Frameworks für Federated Learning

	TensorFlow Federated	PySyft	PaddleFL
Differential Privacy	✔️	🕐❓	✔️
SMPC	❌	✔️	✔️
Dokumentation & Ressourcen	++	+++	+

Keine zufriedenstellende Lösung ...

Siehe auch …

Wissenschaftliche Veröffentlichungen

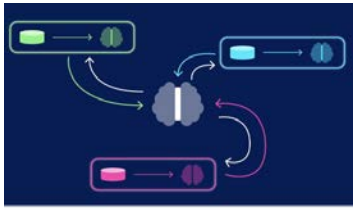
Marisa Mohr, Christian Becker, Ralf Möller, Matthias Richter: **Towards Collaborative Predictive Maintenance Leveraging Private Cross-Company Data**; in: Lecture Notes in Informatics, Band 307, Gesellschaft für Informatik e.V., Bonn, 2020.

Christian Becker, Marisa Mohr: **Federated Machine Learning: über Unternehmensgrenzen hinaus aus Produktionsdaten lernen**, in: atp magazin - 05 2020, S.18-20, ISSN 2190-4111. Vulkan-Verlag GmbH, 2020.

Tobias Bux, Marisa Mohr: **Blockchain-Lösungen für den produktionstechnischen Mittelstand**, in: Prof. Dr.-Ing. Thomas Bauernhansl (Hrsg.): WT WERKSTATTSTECHNIK, Band 111, Nr. 4. VDI Fachmedien GmbH & Co., 2020.

Siehe auch ...

inovex BLOG



Federated Learning: Frameworks for Decentralized Private Training – Part 2

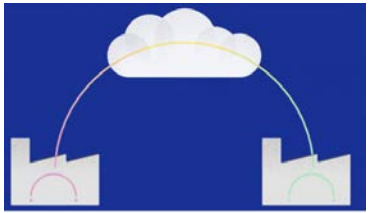
This blogpost evaluates three different Federated Learning frameworks and the concepts they use to achieve a collaborative training.

Federated Learning: A Guide to Collaborative Training with Decentralized Sensitive Data – Part 1

This blog post explains how Federated Learning works and what privacy techniques are necessary to ensure that sensitive data is protected.

TimescaleDB vs. influxDB: Zeitreihendatenbanken für das IIoT

In diesem Artikel diskutieren wir unsere Erfahrungen mit den zwei populären Open-Source-Zeitreihen-Datenbanken TimescaleDB und InfluxDB.



Blockchain-Lösungen für den produktionstechnischen Mittelstand

Die Digitalisierung findet Einzug in den deutschen produktionstechnischen Mittelstand. Bisherige Ansätze beschränken sich auf die Optimierungen firmeninterner Prozesse.

Quelle: https://www.inovex.de/blog/

inovex

Vielen Dank

Marisa Mohr

marisa-mohr.de

@ marisa.mohr@inovex.de

@ mohr@ifis.uni-luebeck.de

Christian Becker

@ christian.becker@inovex.de

www.kosmos-bmbf.de

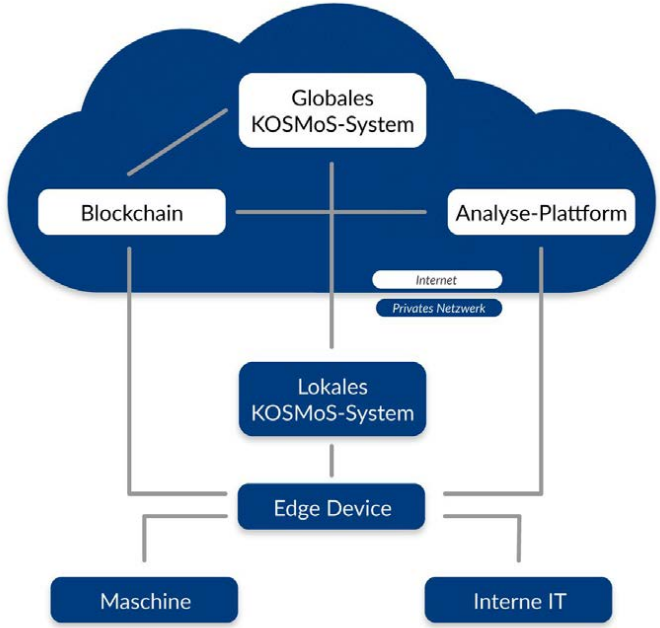
inovex

KOSMoS

Bundesministerium
für Bildung
und Forschung

Kollaborative Smart Contracting Plattform für digitale Wertschöpfungsnetze

- Framework für Systeme zum sicheren und semi-transparenten Austausch von produktions-beschreibenden Informationen zwischen kooperierenden Firmen
- Umsetzung von unternehmensübergreifenden datengetriebenen Geschäftsmodellen, wie z.B. Predictive Maintenance



Schütte · SW Technology People · ASYS GROUP · Hochschule Furtwangen University HFU · Universität Stuttgart ISW · Frankfurt School Blockchain Center · DATARELLA · ONDIDS Apps und Webs für Ihre Prozesse · inovex