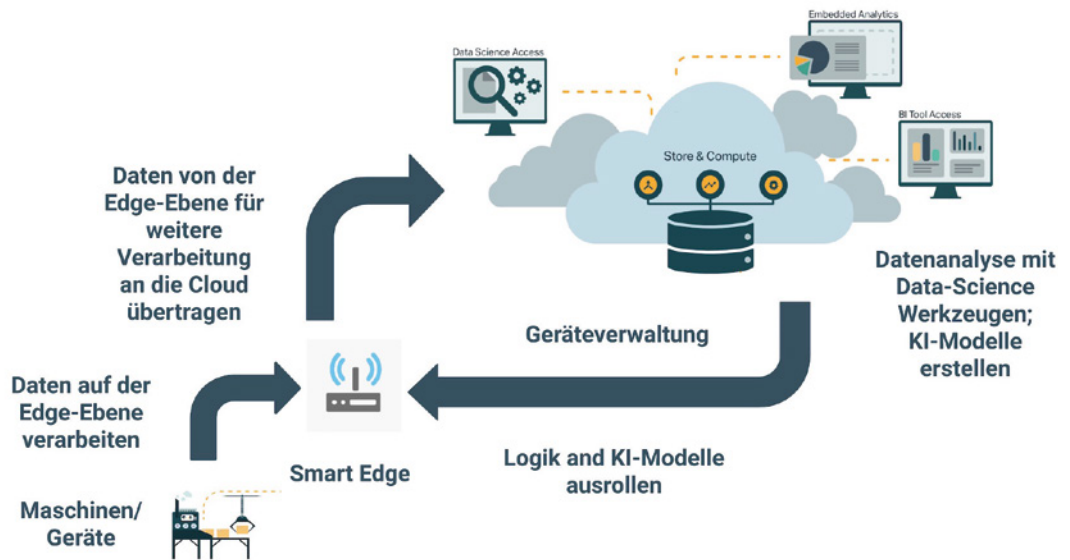


Abb. 1: Das Zusammenspiel von Cloud und Edge



Die Schließung der Lücke zwischen Engineering und Data Science

Industrielle IoT-Plattformen als Treiber des Umbruchs

Ein Beitrag von
Marko Petzold,
Alex Ortner und
Zornitsa Dimitrova

Der verstärkte Einsatz von Edge-Computing in IIoT-Szenarien erfordert heute sowohl Kompetenzen im Bereich Engineering als auch in den Bereichen Datenverarbeitung und Algorithmen. Eine Herausforderung für industrielle IoT-Projekte besteht daher darin, Engineering (Industrial Computing, PLC) und Data Science (Daten, KI-Modelle, Statistik, Cloud-IT) zusammenzubringen. Der vorliegende Artikel stellt einen ganzheitlichen Ansatz für Ingenieure und Data Scientists vor, um gemeinsam an integrierten IoT-Projekten zu arbeiten.

Im Mittelpunkt des Industrial Internet of Things (IIoT) steht die Prämisse, dass eine ausreichende Datenmenge zu neuen Erkenntnissen über Prozesse und Systeme sowie zu Optimierungsstrategien führen kann [Per19]. Diese Erkenntnisse können beispielsweise zur Entscheidungsfindung und für neue Produkte und Dienstleistungen genutzt werden [Set18] sowie zu internen Einsparungen und neuen Umsatzquellen führen.

Die Minimierung der Zeitspanne zwischen Datenerfassung, Datenbereinigung, Analyse und der Entscheidungsfindung, die zu schlüssigen Maßnahmen führt, ist ein wesentlicher Faktor bei der Umsetzung einer IIoT-Lösung [See19]. Denn nur so lässt sich feststellen, ob und welche IIoT-Varianten sich wirklich auszahlen und einen schnelleren ROI erzielen [Sch19].

Mit Blick auf den aktuellen Wandel hin zum Edge-Computing als Wertschöpfungsträger in IIoT-Anwendungsszenarien [Che18] müssen Ingenieure beim Bau von technischen Geräten immer stärker komplexe Logik, wie der Anwendung von KI-Modellen, direkt an oder nahe der Maschine an der Edge ausführen. Die Herausforderungen hier sind einerseits die Entwicklung von KI-Modellen für Edge-Geräte und andererseits die Datenaggregation zur Entwicklung der KI-Modelle. An dieser Stelle kommt die Notwendigkeit einer Vereinigung

der Kompetenzen von Ingenieuren und Data Scientists ins Spiel.

Zur technischen Unterstützung der Zusammenarbeit zwischen Ingenieuren und Data Scientists bietet sich eine Software-as-a-Service-(SaaS-)Lösung an, sodass gemeinsam an IIoT-Entwicklungsszenarien gearbeitet werden kann. Eine solche Lösung wird hier vorgestellt.

Der Wandel zum Edge-Computing

Edge-Computing ist Teil einer dezentralen Topologie, bei der die Informationsverarbeitung nahe an den Quellen der Daten angesiedelt ist [Gar19], die dann zu Informationen gewandelt werden. Laut Brian Burke, Research VP bei Gartner, wird Edge-Computing zu einem dominierenden Ansatz in praktisch allen Branchen und Anwendungsfällen [Gar19]. Ein wichtiger Treiber hinter diesem Trend ist die Notwendigkeit, massive Datenmengen aus verschiedenen heterogenen Quellen mit unterschiedlicher Syntax und Semantik schnell vor Ort auszuwerten und so auf der Edge-Ebene zu reduzieren, sodass sie zur weiteren Verwendung an die Cloud übertragen werden können.

Bei umfangreichen Datenmengen erscheint es sinnvoll, die Daten nahe am Ort ihrer Entstehung zu verarbeiten. Ein Beispiel ist der IoT-Einsatz in

sehr sensorintensiven und damit datenintensiven Umgebungen. Dies ist auf unterschiedliche Herausforderungen zurückzuführen, darunter Bandbreite, Netzwerklatenz und Gesamtgeschwindigkeit [Cra20]. Edge-Computing wird besonders in IIoT-Anwendungen mit einer erfolgskritischen oder dezentralen Komponente relevant, weil die Geschwindigkeit der Datenverarbeitung vor Ort ein wesentliches Unterscheidungsmerkmal bei solchen Anforderungen ist: Die Ergebnisse aus den Daten werden schnell gebraucht, und es dauert zu lange, bis die Daten von der Edge in die Cloud und die dort ermittelten Ergebnisse nach Verarbeitung wieder an die Edge zurückgespielt werden.

Darüber hinaus ermöglicht insbesondere bei hochsensiblen oder proprietären Informationen die Verarbeitung von Daten an der Netzwerk-Edge eine stärkere Kontrolle über den Datenzugriff gegenüber der Cloud [Blo18]. Die Entscheidungsfindung kann an der Edge erfolgen, wobei alle Daten im Unternehmen verbleiben und nur nicht vertrauliche Informationen den Weg in die Cloud finden würden. Auch die Anonymisierung und Verschlüsselung von Unternehmensdaten kann an der Edge erfolgen, wodurch Unternehmen kritische Datenbestände vor möglichen Sicherheitslücken schützen können.

Diese Daten können sogar auf der Edge für die Analyse über mehrere Stufen hinweg vorbereitet werden, einschließlich Datenvalidierung, Datenbereinigung, Transformation, Indexierung, Aggregation und Speicherung. Abhängig von der Art der Daten und dem konkreten Geschäftsziel müssen Unternehmen auch eine geeignete Verarbeitungsmethode auswählen. Diese wiederum kann von der Batch-Verarbeitung bis hin zur Echtzeitverarbeitung reichen.

Während Edge-Computing einen Mehrwert durch die oben aufgeführten Punkte bietet, ermöglicht die Speicherung der Daten und deren Analyse in der Cloud eine tiefere und umfassendere Verarbeitung. Hier können zum Beispiel auch Daten aus unterschiedlichen, nichtlokalen Quellen für eine umfangreichere Datenmodellierung kombiniert werden, um tiefere Erkenntnisse zu generieren. Solche Daten sind nicht unmittelbar an der lokalen Edge verfügbar, können aber über die Cloud abgerufen werden. Darüber hinaus stehen in der Cloud viel größere Speicher und Rechenressourcen für anspruchsvolle Analysen und die Optimierung von KI-Modellen zur Verfügung [Zor19]. So ergibt sich ein Zusammenspiel zwischen den Stärken von Cloud und Edge (Abbildung 1).

Während Edge-Computing in der Regel unter der Leitung von Ingenieurabteilungen steht, liegt die Verantwortung über Datenanalysen in der Cloud meist in der Hand der Data Scientists. So wurden zum Beispiel in einem Projekt bei der Continental AG die Edge-Geräte von den Ingenieuren zur Erkennung spezifischer Geräusche im Fahrzeug auf die Sensoren und Messparameter ausgelegt. Die Analyse der gesammelten Daten und die Modellierung der benötigten KI-Modelle zur Geräuscherkennung



DR. MARKO PETZOLD ist der Gründer von Record Evolution GmbH und verfügt über langjährige Erfahrung in der Konzeption und Implementierung von Projekten in den Bereichen Data Science, IoT und Data Warehousing in zahlreichen Unternehmen.

E-Mail: marko.petzold@record-evolution.de

DR. ALEX ORTNER ist Senior Data Scientist bei Record Evolution GmbH und arbeitet an Lösungen, um Data Science und IoT zusammenzubringen. Sein aktueller Fokus liegt auf der Architektur und Entwicklung einer End-to-End-Plattform für die gesamte IIoT-Wertschöpfungskette.

E-Mail: alex.ortner@record-evolution.de

DR. ZORNITSA DIMITROVA ist Technische Redakteurin bei Record Evolution GmbH und schreibt an der Schnittstelle zwischen Philosophie und Konzeptbildung im Bereich der IIoT-Lösungen. Ihre Interessen umfassen die Philosophie der Technik und die Nachhaltigkeitsforschung.

E-Mail: zornitsa.dimitrova@record-evolution.de

wurde von Data-Science-Experten in einer dediziert dafür ausgelegten Cloud-Umgebung durchgeführt.

Engineering und Data Science: Wie bringen wir diese Welten zusammen?

Der Verantwortungsbereich der Ingenieurabteilungen umfasst neben Edge-Computing auch Maschinen, elektronische Geräte und Sensoren, die an den Maschinen angebracht sind. Ingenieure sind auch in langsame Entwicklungszyklen eingebunden, die den Zugriff auf Maschinen ohne Beeinträchtigung der Funktion, das Sammeln von Daten aus physischen Geräten und die ständige, meist manuelle Extraktion großer Mengen an Rohdaten beinhalten.

Cloud Services, Datenverarbeitung und -analyse liegen in den Händen von Data Scientists. Hier sind agile Entwicklungen, schnelle Versuchs- und Fehlerszenarien, die Entwicklung von Algorithmen und die Datenanalyse im Allgemeinen das zentrale Thema.

Im klassischen Arbeitsprozess extrahiert die Ingenieurabteilung Daten aus wenigen industriellen Geräten und analysiert und bearbeitet diese Daten selbst an eigener spezieller Software auf ihren PCs (zum Beispiel mittels Labview, Matlab oder Excel). Dieses Vorgehen stößt aufgrund der großen Daten-

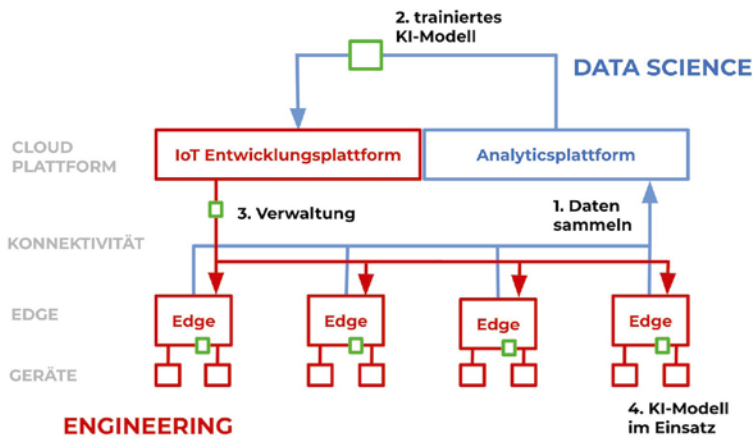


Abb. 2: Die Plattform für den IloT-Entwicklungszyklus

mengen und gestiegenen Anforderungen an die Datenauswertung heute an seine Grenzen.

Das Bedürfnis, umfangreiche Datenmengen von den Maschinen zu sammeln und diese Daten aufbereiten zu lassen, nimmt zu. Dies wird dann zur Aufgabe eines Data-Science-Teams. Wie in allen analyseorientierten Projekten ist das Extrahieren der Daten aus den Geräten ein aufwendiger Prozess. Die Engpässe für IloT-Projekte sind in der Regel nicht die Ideen oder die Algorithmen, sondern die Datenpipelines und die Datenheterogenität.

Die Daten, die aus der Ingenieurabteilung kommen, sind manchmal im falschen Format oder nicht so, wie sie im konkreten Fall benötigt werden. Die Ingenieure müssen möglicherweise auf den Edge-Computer, das heißt das Rechenggerät nahe der Quelle der Daten, zugreifen, um herauszufinden, wie sie die Daten auf eine andere Art und Weise extrahieren können, damit diese nutzbar sind. Es kann eine höhere Iterationsrate oder ein größeres Zeitfenster für die Erfassung erforderlich sein, sodass die Edge-Geräte möglicherweise komplett neu konfiguriert werden müssen, um die Daten auf eine andere Art und Weise zu sammeln. Ein an dieser Stelle umgesetzter wiederholter manueller Prozess ist schwer nachvollziehbar und kann zeitintensiv sein.

Ingenieure würden von einer Datenextraktionsmethode profitieren, die auf ihren Maschinen funktioniert und aus der Ferne über ein Netzwerk zu erstellen und anzupassen ist. Dazu müssen alle Edge-Geräte an den Maschinen über eine gemeinsame, mit Werkzeugen für Datenextraktion, Datenanalyse, Gerätemanagement und Anwendungsentwicklung ausgestattete Plattform verbunden sein.

Eine Lösung ist in diesem Fall eine SaaS-Plattform, die den gesamten IloT-Entwicklungszyklus von der Verarbeitung an der Maschine bis hin zu den Ergebnissen der Datenanalyse, das Gerätemanagement und das Ausrollen von Anwendungen abdeckt. Die Nutzer können sich einloggen, um eine Verbindung mit dem Rechenggerät an der Edge zu erstellen, Anwendungen bereitzustellen oder den gleichen Code oder die gleichen Konfigurationen auf einer Flotte von Geräten zu verwenden, die sich an unterschiedlichen, räumlich getrennten Standorten befinden.

Da die in der Plattform verknüpften Edge-Geräte bereits vernetzt sind, kann dasselbe Netzwerk auch dazu verwendet werden, Daten an eine Analyseplattform zu senden. Hier stehen dem Data Scientist die Daten aller Edge-Geräte konsolidiert zur weiteren Verarbeitung zur Verfügung. Diese lassen sich nun bereinigen, mit anderen Daten zusammenführen, modellieren und direkt für Analyse- und Visualisierungsaufgaben verwenden.

Auch Data Scientists können die Fernkonfiguration einer solchen Plattform nutzen, um die Modelle, die sie mit Hilfe der Daten erstellen, wieder auf die Edge-Geräte zu bringen. Dazu können sie dieselben Plattform-Werkzeuge nutzen, mit denen die Ingenieure zuvor ihre Datenextraktionslogik erstellt haben. Dank der holistischen Plattform können nun Ingenieure und Data Scientists gemeinsam sowohl auf den Geräten als auch an den Algorithmen arbeiten (Abbildung 2).

Die IloT-Plattform als digitales Herzstück eines industriellen Betriebs

Eine IloT-Plattform, die offen [PeH'18] und interoperabel [Set'18] ist und die Anforderungen von Ingenieuren und Data Scientists erfüllt, umfasst die Verwaltung von Geräten, die Entwicklung von Logik sowie etablierte Datenmanagement- und Datenanalysetechnologien. Im Detail benötigen sowohl Ingenieure als auch Data Scientists die Möglichkeit, per Fernzugriff Apps zu erstellen und als produktive Anwendung auf eine Flotte von Geräten in Echtzeit zu verteilen, sowie eine Gerätemanagement-Funktionalität, mit der sie sehen können, woher die Daten kommen und welches Gerät an einem bestimmten Ort eingesetzt wird (siehe Abbildung 2).

Die vereinten Fähigkeiten von Geräteverwaltung, Edge-Anwendungsentwicklung und einer Cloud-Data-Science-Plattform für Analytics bilden eine End-to-End-Lösung, die für ein breites Spektrum von IloT-Szenarien anwendbar ist. Auf diese Weise besteht ein wiederholbarer Zyklus von Datenerfassung, Analyse, Modellentwicklung und Deployment in der Edge, der als IloT-Entwicklungszyklus bezeichnet werden kann. Ausgehend von der Datenerfassung von Sensoren und smarten Geräten, dem Sammeln, Vorverarbeiten und Aggregieren der Daten am Gateway ist der nächste Schritt die Übertragung der Daten an eine integrierbare Cloud-Data-Science-Plattform, um weitergehende Analysen durchzuführen oder KI-Modelle zu trainieren. Sobald die Nutzer ihr KI-Modell in der Cloud trainiert haben, können sie die Plattform nutzen, um die Logik und die Modelle wieder an die Edge zu bringen und auf einer Vielzahl von Geräten bereitzustellen (siehe Abbildung 2).

Fazit

Es genügt nicht, nur einzelne Aspekte eines IloT-Implementierungsprojekts zu realisieren. Um die Herausforderungen bei der Implementierung von IloT-Lösungen erfolgreich zu bewältigen, benötigt

man einen umfangreichen holistischen Ansatz, mit dem sowohl die Bedürfnisse der Ingenieurabteilungen als auch die Arbeit der Data Scientists abgedeckt werden. Dies bedeutet, dass es nicht reicht, eine IIoT-Lösung aufzubauen, die nur Gerätemanagement, nur Anwendungsentwicklung oder nur Datenanalyse bietet. Eine erfolgreiche Zusammenarbeit von Ingenieur- und Data-Science-Bereichen beinhaltet alle drei Aspekte der industriellen IIoT-Lösungen.

Eine IIoT-Plattform kann kritische Herausforderungen bewältigen, indem sie eine Kombination von Funktionen wie die Verwaltung von IIoT-Endgeräten und deren Konnektivität sowie Funktionen zur Erstellung und Verwaltung von Edge-Anwendungen bietet. Darüber hinaus wird eine IIoT-Plattform, die Engineering und Data Science zusammenbringt, diese Fähigkeiten durch die Aufnahme, Transformation und Analyse von IIoT-Daten ergänzen.

Literatur

- [Blo18] Bloom, G. et al.: Design Patterns for the Industrial Internet of Things. In: 14th IEEE International Workshop on Factory Communication Systems (WFCS), 2018, S. 1-10
- [Che18] Chen, B. et al.: Edge Computing in IIoT-Based Manufacturing. In: IEEE Communications Magazine 56 (9), 2018, S. 103-109
- [Cra20] Craclunescu, M. et al.: IIoT Gateway for Edge Computing Applications. In: Borangiu, T. et al. (Hrsg.): Service-oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future. Springer 2020
- [Gar19] Gartner: Gartner Identifies the Top Strategic Technology Trends for 2020. <https://www.gartner.com/en/newsroom/press-releases>, abgerufen am 26.1.2020
- [PeH18] Petrik, D. / Herzwurm, G.: Platform Ecosystems for the Industrial Internet of Things. In: Software-Intensive Workshop., 2018, S. 57-71
- [Per19] Perwej, Y. et al.: The Internet of Things (IIoT) and its Application Domains. In: International Journal of Computer Applications 182(49), 2019, S. 36-49
- [Sch19] Schermuly, L. et al.: Developing an Industrial IIoT Platform. 14. Int. Tagung Wirtschaftsinformatik, 2019
- [See19] Seetharaman, A. et al.: Customer Expectations from Industrial Internet of Things (IIoT). In: Journal of Manufacturing Technology Management 30 (8), 2019, S. 1161-1178
- [Set18] Setzke, D. et al.: Platforms for the Industrial Internet of Things: Enhancing Business Models through Interoperability. In: Zelm, M. et al. (Hrsg.): Enterprise Interoperability. Wiley 2018, S. 65-71
- [Zor19] Zornio, P.: Cloud vs. Edge Computing - What's the Difference? <https://www.forbes.com>, abgerufen am 2.2.2020

c/o DATA SCIENCE

SAVE THE DATE
HAMBURG, 28 APRIL 2020



c/o data science is designed for YOU - exchange and discuss your ideas and challenges with other data scientists & data geeks.

Head over to our website to check out our program:

CO-DATASCIENCE.DE

zum Inhalt