

Gefördert durch:



Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages



TrendAuto
2030plus

Daten und KI in der Produktion bei der Vollmann Group

Handlungsfeld: Technologie

Entwicklungssystematik und kollaboratives Arbeiten

09.06.2026

10.06.2026

Seite 1

Arbeitskreis Entwicklungssystematik und kollaboratives Arbeiten

Workshop: Daten und KI in der Produktion bei der Vollmann Group

Gefördert durch:



Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages



TrendAuto
2030plus



Quelle: geralt auf pixabay.com

Begrüßung und Zielsetzung

Agenda

09:00 Uhr

Onboarding

- I. Begrüßung, Vorstellung der Teilnehmer und Zielsetzung des Workshops
- II. Vorstellung der Vollmann Group und Führung durch die Produktion
- III. Challenge der Vollmann Group als Grundlage für die Lösung durch die KI

10:30 Uhr



Pause 


10:45 Uhr

Der Prozess: Von Daten zum KI-System

- I. Akquise und Analyse der Daten
- II. Auswahl und Training des Modells
- III. Modellanalyse und Fine-Tuning
- IV. Einsatz und Weiterentwicklung

11:45 Uhr



Pause 

12:00 Uhr

Reflektion

- I. Lessons Learned: Zusammenfassung und Beurteilung
- II. Vergleich: Good Practices und die akademische Perspektive – Sinnvolle Synthese
- III. Feedbackrunde und Abschluss

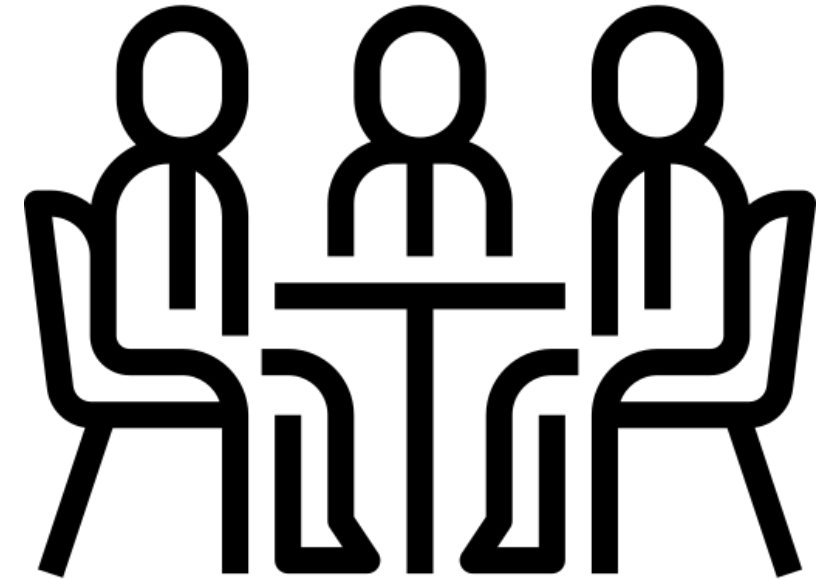
13:00 Uhr



Begrüßung und Zielsetzung

Ihre Vorstellung(en)

- *Ihr Name*
- *zugehöriges Unternehmen + Funktion*
- *Mit welchen Gedanken und welcher Stimmung bin ich heute Morgen in den Workshop gegangen?*
- *Meine Erwartungen an den Workshop*
- *Wie verwenden Sie KI im Alltag?*
- ...



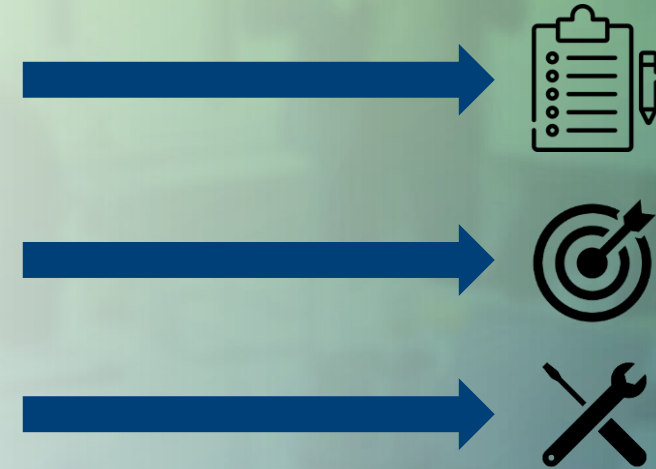
Innerhalb des Arbeitskreises können sich die Akteure kennenlernen und vernetzen. Durch das gemeinsame Bearbeiten der Themenstellungen entstehen neue Einblicke in die unterschiedlichen Arbeits- und Denkweisen.

Begrüßung und Zielsetzung

Zielsetzung des Workshops

Kernpunkte des heutigen Termins

- Beleuchtung der **gewählten Herausforderung** für das Unternehmen und der daraus resultierenden Perspektive
- Strukturierte **Darstellung des Ablaufs** für den Prozess hin von den Daten zum eigentlichen KI-System
- **Best Practices & Reflektion:** Vergleich zwischen akademischer und industrieller Perspektive sowie Verallgemeinerung der Vorgehensweise



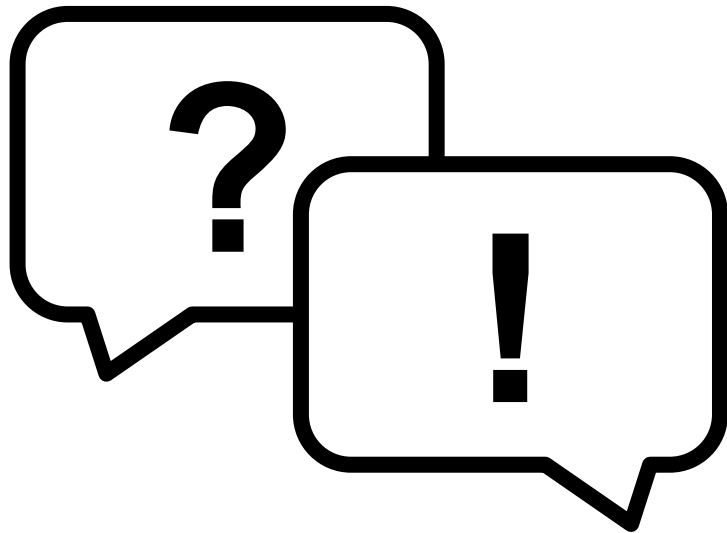
Dieser Workshop verbindet Unternehmensherausforderungen mit dem Prozess des Aufbaus von KI-Systemen und vereint akademische sowie industrielle Best Practices.

Vorstellung der Vollmann Group



Challenge der Vollmann Group

→ Grundlage für die KI-Lösung

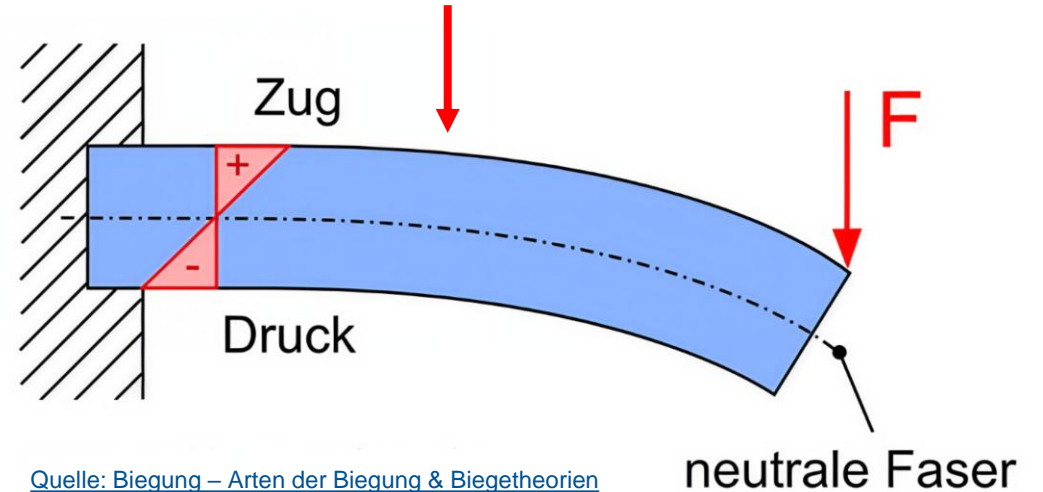


Challenge der Vollmann Group

Problemstellung

- **Phänomen:**
 - Erhöhte Rissbildung bei Umformprozessen entlang der gestreckten Faser
 - bei gleicher Stahlsorte, aber anderer Charge
 - bei gleichbleibenden Prozessparametern / Tools
- **Hypothese:**
 - Rissbildung ist auf wesentliche Unterschiede in den Legierungsbestandteilen zurückzuführen
- **Zusätzliche Schwierigkeit:**
 - Legierungsanteile sind nicht gesammelt bzw. tabellarisch erfasst
 - Prozessparameter nur beschränkt aufgezeichnet

Risse können bei **Biegung** parallel und senkrecht zur neutralen Faser entstehen



Challenge der Vollmann Group

Problemstellung

- **Phänomen:**
 - Erhöhte Rissbildung bei Umformprozessen entlang der gestreckten Faser
 - bei gleicher Stahlsorte, aber anderer Charge
 - bei gleichbleibenden Prozessparametern / Tools
- **Hypothese:**
 - Rissbildung ist auf wesentliche Unterschiede in den Legierungsbestandteilen zurückzuführen
- **Zusätzliche Schwierigkeit:**
 - Legierungsanteile sind nicht gesammelt bzw. tabellarisch erfasst
 - Prozessparameter nur beschränkt aufgezeichnet



Challenge der Vollmann Group

Identifikation von Materialanomalien

- **Status Quo:**
Freigabe von Stahlerzeugnissen orientiert sich an Zugversuch / Materialzeugnissen
- **Erkennungsdefizit:**
Material-Anomalien fallen manchmal zu spät auf
→ Ausschuss oder Reklamationen
- **Experten-Abhängigkeit:**
Bewertung der Legierungszusammensetzungen erfolgt erfahrungsbasiert (Kombination der Legierung)
- **Die Folge:**
Reaktive Fehlerbehebung, riskante Abhängigkeit von individuellem Wissen und Ausfallkosten



Technology
Arts Sciences
TH Köln



Vollmann
Group

Während der Problemdefinition sind ausführliche Gespräche zwischen Entwicklern und Partnerunternehmen wesentlich, um gemeinsam die Problemstellung zu verstehen und die Aufgabenstellung festzulegen.

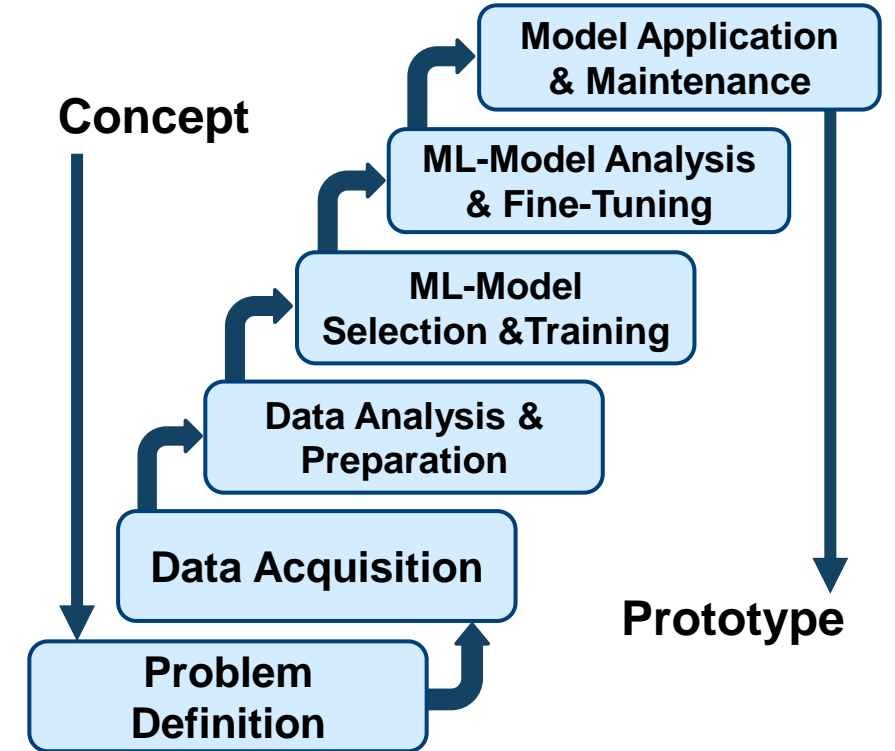


Wie gehen wir diese Herausforderung an?

Lösung der Challenge durch KI

Lösungsweg: Entwicklungssystematik für Künstliche Intelligenz

- **Kein einheitlich** definierter Standard
- verschiedenen Standardwerken zeigen jedoch **ähnliche** Vorgehensweise [1, 3, 4]
- Rechts → **Eigener Ablauf** auf Basis eigener Erfahrungen
- Iterationen innerhalb des Prozesses sind **möglich**
- Je nach **Projekt** weisen die Blöcke verschiedene Inhalte auf

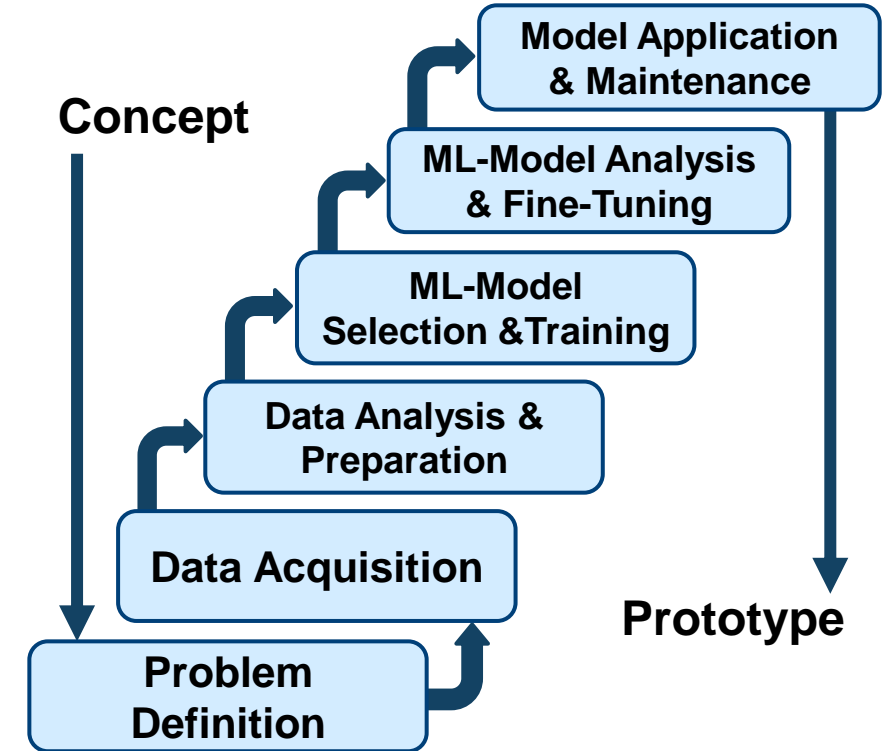


Als Entwicklungssystematik hat sich ein stufenartiger Prozess bewährt, der von der Problemfindung bis zur praktischen Anwendung reicht.

Lösung der Challenge durch KI

Lösungsweg: Entwicklungssystematik für Künstliche Intelligenz

Mehr dazu
in unserem KI-Leitfaden unter
trendauto2030plus.de



Als Entwicklungssystematik hat sich ein stufenartiger Prozess bewährt, der von der Problemfindung bis zur praktischen Anwendung reicht.

In welchen Unternehmensbereichen setzen Sie KI ein?

statista 

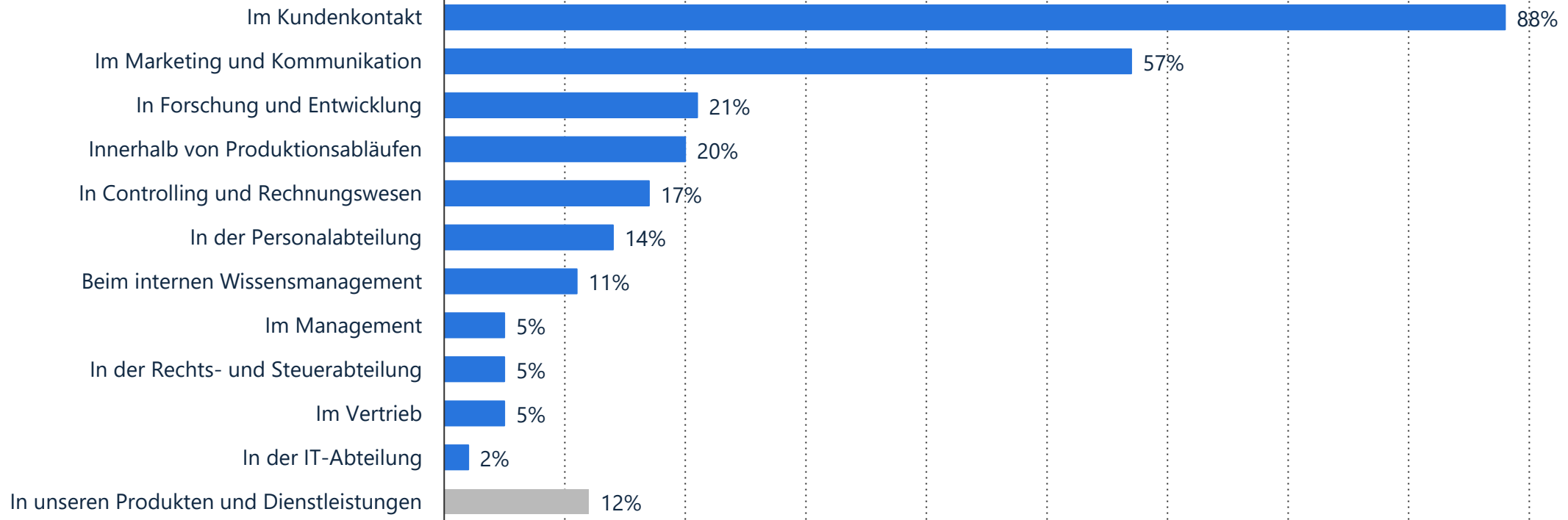
Lösung der Challenge durch KI

Umfrage: KI Einsatz

In welchen Unternehmensbereichen setzen Sie KI ein?

Befragte in Prozent

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%



Beschreibung: Eine Umfrage von Bitkom aus 2025 zeigt, dass Unternehmen KI schon jetzt vor allem im Kundenkontakt einsetzen (88 Prozent der Befragten). 57 Prozent der befragten Unternehmen nutzen KI im Marketing und Kommunikation. [Mehr](#)
Hinweis(e): Deutschland; KW 27 bis KW 32 2025; 215 Befragte; Unternehmen in Deutschland mit mindestens 20 Beschäftigten
Quelle(n): Bitkom

statista

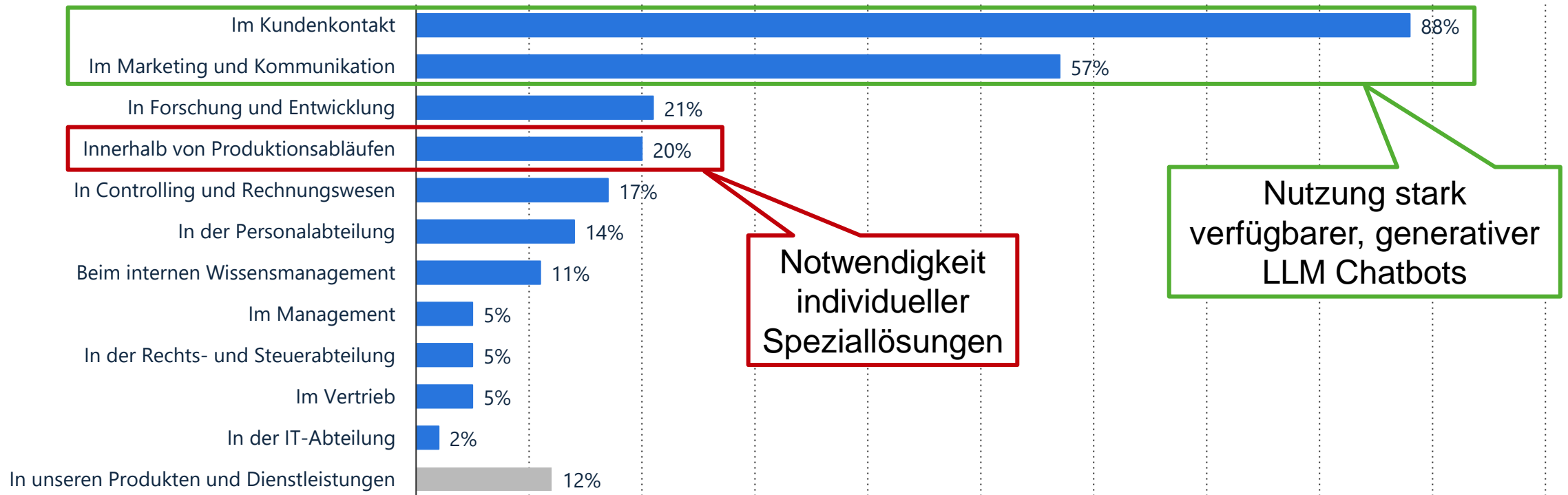
Lösung der Challenge durch KI

Umfrage: KI Einsatz

In welchen Unternehmensbereichen setzen Sie KI ein?

Befragte in Prozent

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%



Beschreibung: Eine Umfrage von Bitkom aus 2025 zeigt, dass Unternehmen KI schon jetzt vor allem im Kundenkontakt einsetzen (88 Prozent der Befragten). 57 Prozent der befragten Unternehmen nutzen KI im Marketing und Kommunikation. [Mehr](#)
Hinweis(e): Deutschland; KW 27 bis KW 32 2025; 215 Befragte; Unternehmen in Deutschland mit mindestens 20 Beschäftigten
Quelle(n): Bitkom

statista

Was sind die größten Hemmnisse beim Einsatz von KI in Ihrem Unternehmen?

statista 

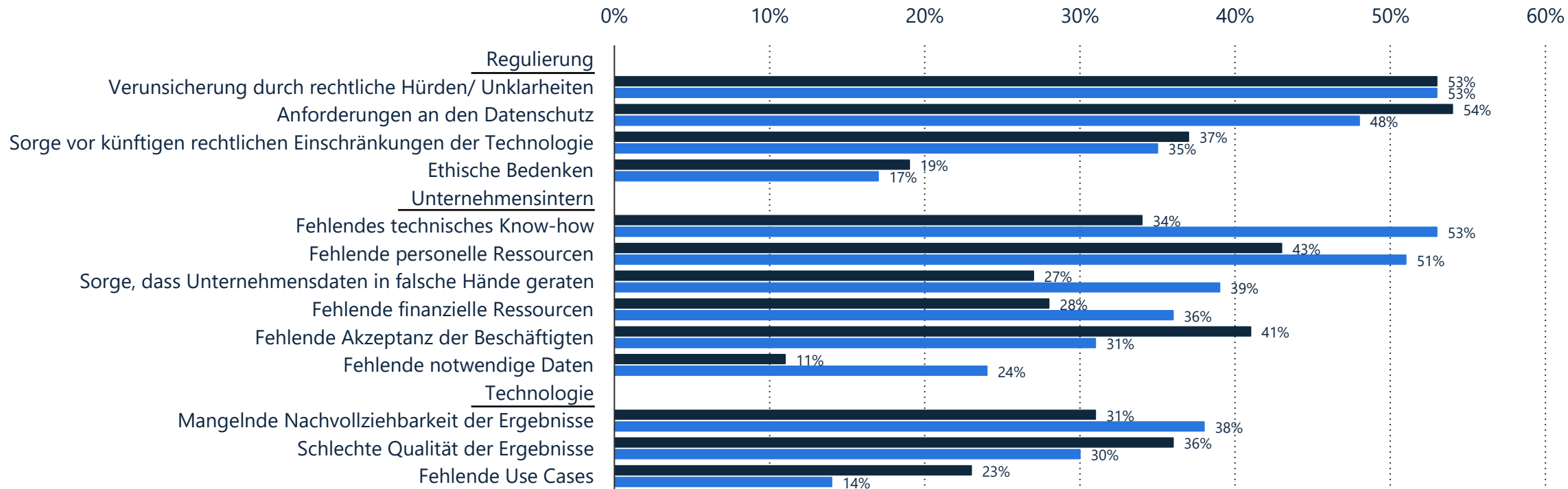
Lösung der Challenge durch KI

Umfrage: Hemmnisse des Einsatzes

Was sind die größten Hemmnisse beim Einsatz von KI in Ihrem Unternehmen?

Befragte in Prozent

■ Unternehmen, die KI nutzen ■ Alle Unternehmen



Beschreibung: Eine Bitkom-Umfrage aus dem Jahr 2025 zeigt, dass vor allem fehlendes technisches Know-how sowie die Verunsicherung durch rechtliche Hürden und Unklarheiten dafür verantwortlich sind, dass KI nicht von Unternehmen eingesetzt wird (53 Prozent). Außerdem werden Unternehmen insbesondere durch personelle Ressourcen und Anforderungen an den Datenschutz abgeschreckt. [Mehr](#)

Hinweis(e): Deutschland; KW 27 bis KW 32 2025; 604*; * Unternehmen, die generative KI nutzen (n=215) Mehrfachnennungen waren möglich. [Mehr](#)

Quelle(n): Bitkom

statista

Lösung der Challenge durch KI

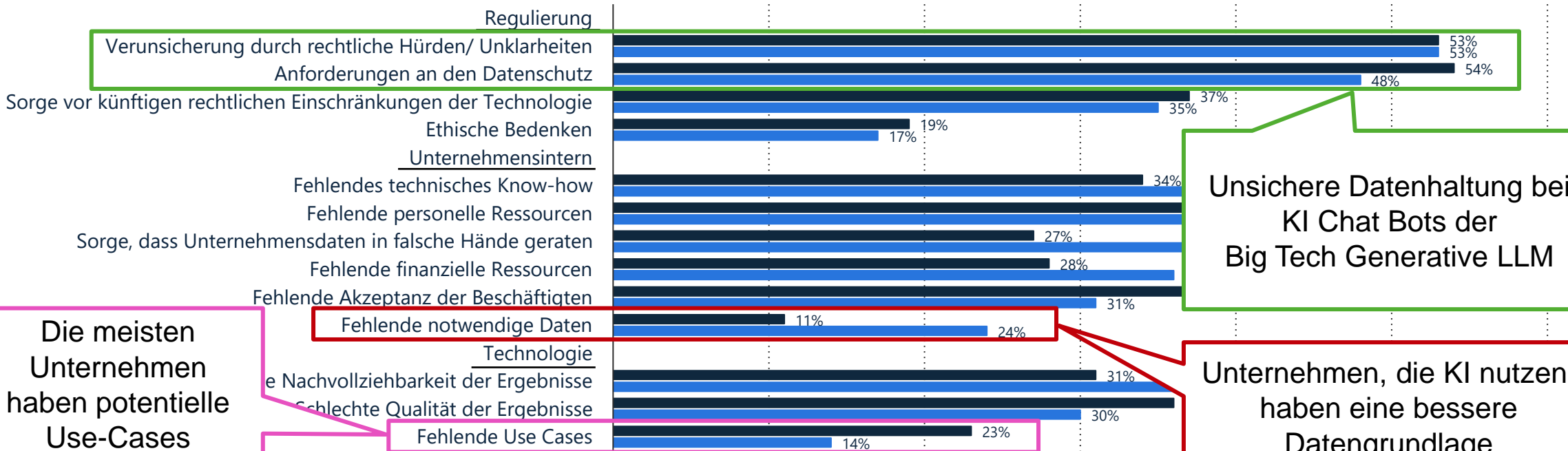
Umfrage: Hemmnisse des Einsatzes

Was sind die größten Hemmnisse beim Einsatz von KI in Ihrem Unternehmen?

Befragte in Prozent

■ Unternehmen, die KI nutzen ■ Alle Unternehmen

0% 10% 20% 30% 40% 50% 60%



Die meisten Unternehmen haben potentielle Use-Cases

Fehlende notwendige Daten

Unsichere Datenhaltung bei KI Chat Bots der Big Tech Generative LLM

Unternehmen, die KI nutzen, haben eine bessere Datengrundlage

Beschreibung: Eine Bitkom-Umfrage aus dem Jahr 2025 zeigt, dass vor allem fehlendes technisches Know-how sowie die Verunsicherung durch rechtliche Hürden und Unklarheiten dafür verantwortlich sind, dass KI nicht von Unternehmen eingesetzt wird (53 Prozent). Außerdem werden Unternehmen insbesondere durch personelle Ressourcen und Anforderungen an den Datenschutz abgeschreckt. [Mehr](#)
Hinweis(e): Deutschland; KW 27 bis KW 32 2025; 604*; * Unternehmen, die generative KI nutzen (n=215) Mehrfachnennungen waren möglich. [Mehr](#)
Quelle(n): Bitkom



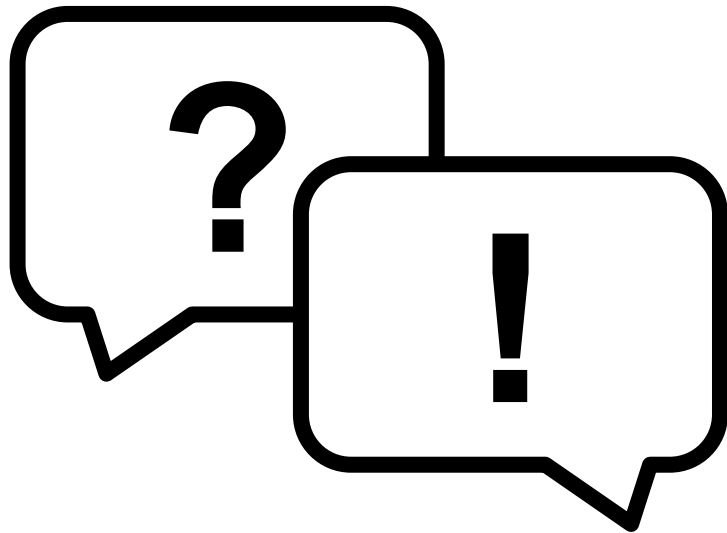
PAUSE



Quelle: geralt auf pixabay.com

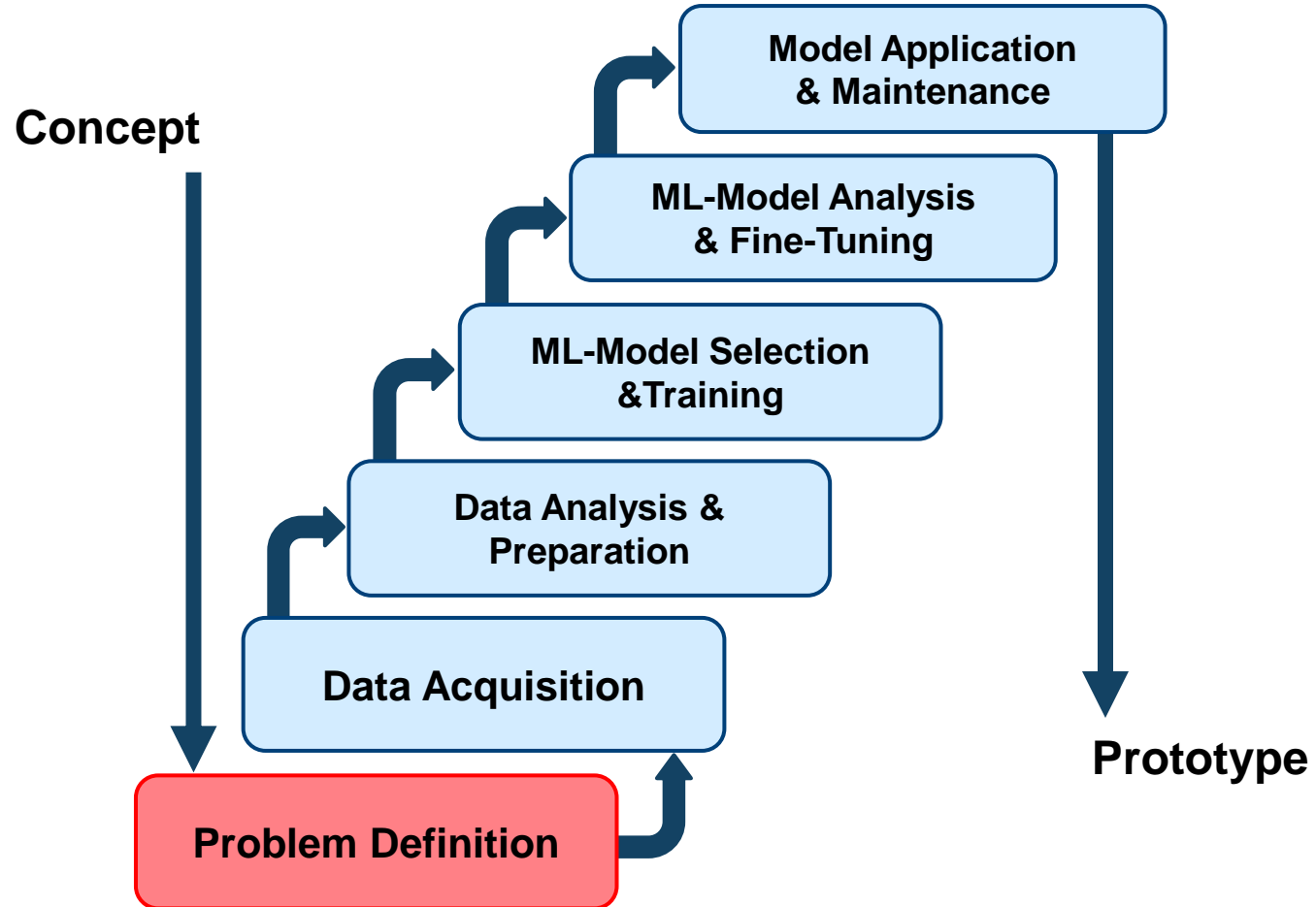
Workshop

Von den Daten zum KI System



Workshop – Daten und KI in der Produktion

Problem Definition



Problem Definition

KI-gestützte Früherkennung von rissgefährdeten Materialien

WARUM: Herausforderung der Rissbildung bei unterschiedlichen Stahl-Chargen

Ziel: Abweichungen von Legierungszusammensetzungen im Materialeingang identifizieren

Herausforderung: Innerhalb der Güte S650MC treten schwankende Legierungsgehalte in erheblichem Maße auf

Mehrwert: Die Nutzung einer KI-basierten Lösung kann die Prozessplanung vereinfachen und -sicherheit erhöhen

WIE: Einführung einer KI-gestützten Anomalieerkennung

Datennutzung: Analyse von Werkszeugnissen nach DIN EN 10204/3.1, anhand erhöhten Ausschusses klassifiziert

Modellparameter: Alle oder ausgewählte mechanische Kennwerte und chemische Bestandteile

Klassifizierungsziel: Klassifizierung von Werkszeugnissen und zugehörigen Materialchargen

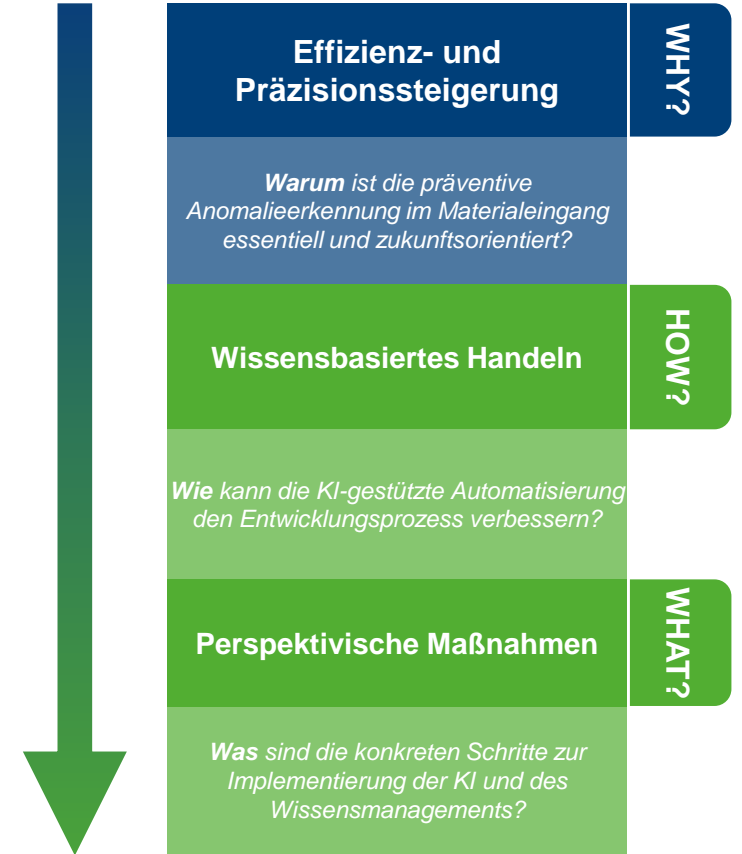
Wissensmanagement: Konsolidierung und Systematisierung des Auswerteprozesses, inklusive Materialgrenzwerte

WAS: Implementierung und Weiterentwicklung

Implementierung: Entwicklung eines Prototyps, der mittels KI eine Klassifizierung anhand der Modellparameter prädiziert

Wissensarchivierung: Erste strukturierte Zusammenführung des vorhandenen Wissens in zentral zugängliches System.

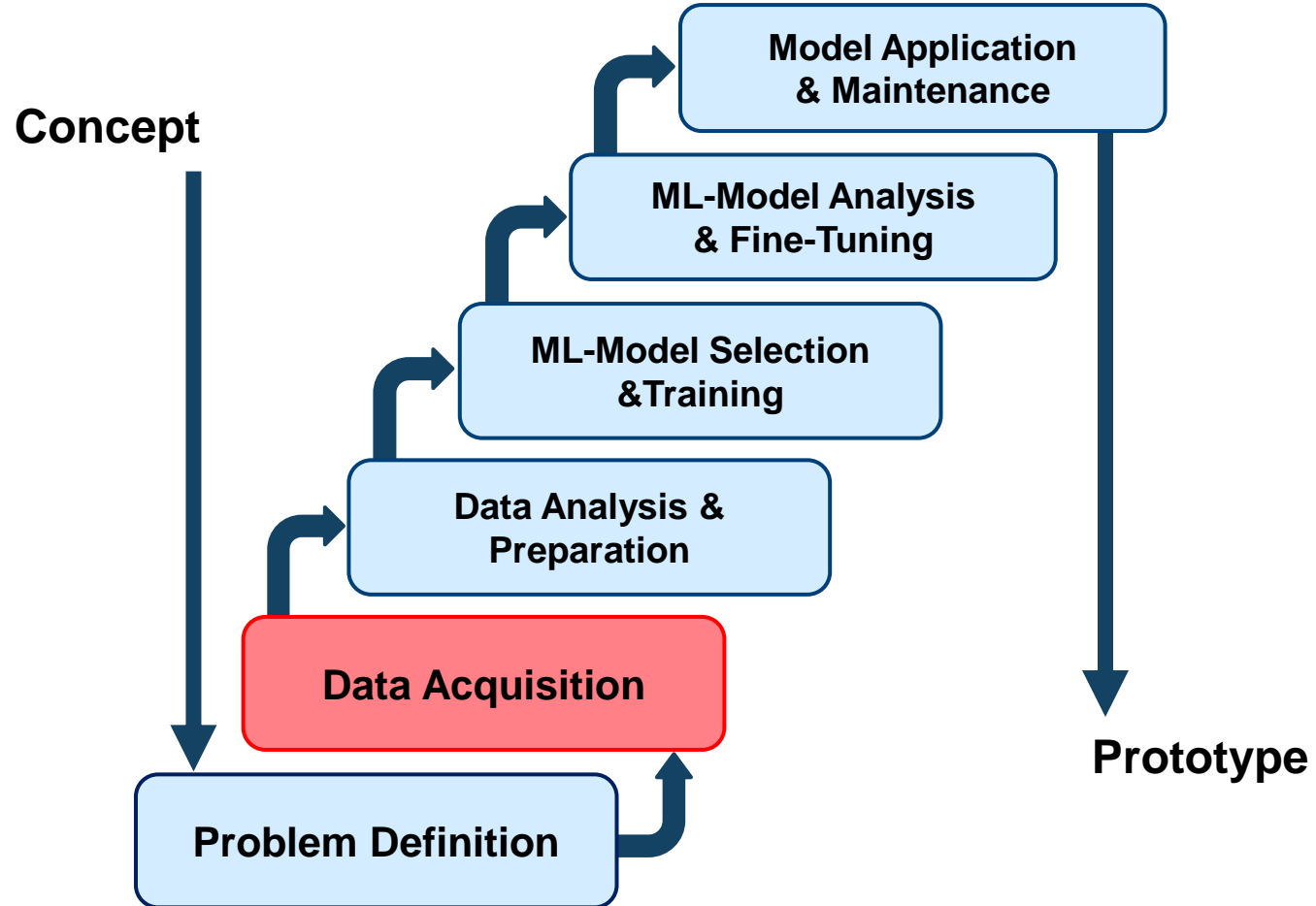
Langfristiges Ziel: Steigerung der Prozesssicherheit und –planbarkeit, Verbesserung der Produktqualität durch präzise und datenbasierte Entscheidungen.



Das Projekt optimiert die Prozessplanung mittels moderner KI-Methoden und konsolidiert dabei Expertenwissen.

Workshop – Daten und KI in der Produktion

Data Acquisition



Data Acquisition

Identifikation der Datenanforderungen

Welche Daten liegen zur Problemlösung vor?

- Systematische Bereitstellung von Materialzertifikaten gemäß DIN EN 10204/3.1 in digitaler **PDF-Form** für jede Materiallieferung
- Erfolgte Klassifizierung einzelner Chargen auf Basis dokumentierter Abweichungen und erhöhter **Ausschussraten**.

Welche werden benötigt?

- Ein strukturierter Datensatz aus den Materialzertifikaten samt Klassifizierung
- Zuordnung der Produktionschargen zu den Materialchargen muss sichergestellt werden

Güte : S650MC gebeizt, geölt
Kundenartikelnr. : WB 28600 Lnr. 70911
Abmessung in mm : 3,00 x 149,50
Dickentoleranz in mm: +0,00/-0,20
Liefergewicht kg : 3.986
Ringanzahl : 4

Charge : **1304832**
W-Schmelze : 55-2371
Werkscoilnr. : 193033

Mechanische Werte Prüfnorm: DIN EN ISO 6892-1 Prüfverfahren B

Name	Wert	Kundenvorschrift	DIN EN 10149-2
Streckgrenze (Re)	726,00 MPa		min. 650,00 MPa
Zugfestigkeit (Rm)	778,00 MPa		700,00 - 880,00 MPa
Bruchdehnung A	21,00 %		min. 12,00 %

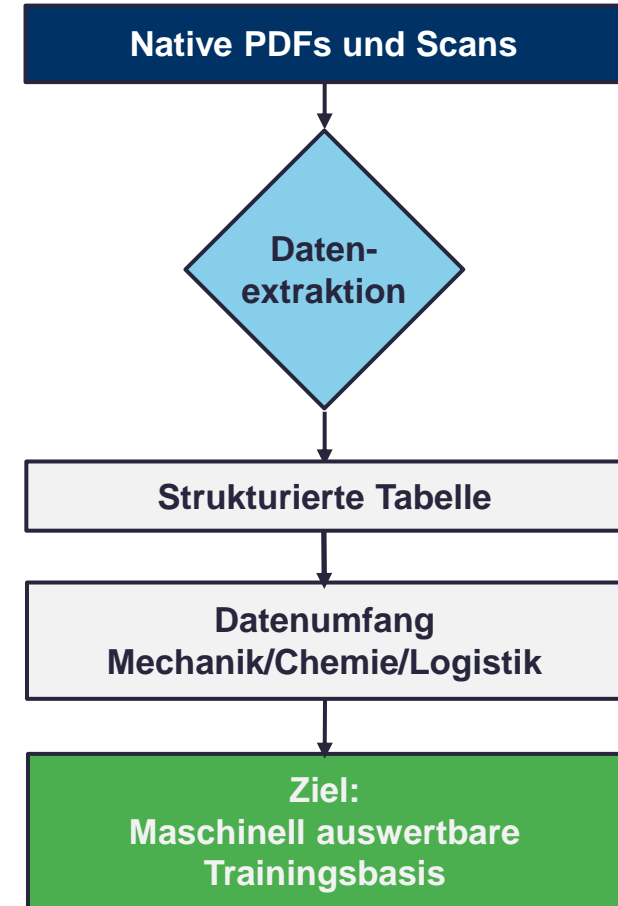
Chemische Analyse

Element	Analysewert %	Kundenvorschrift	DIN EN 10149-2
Kohlenstoff C	0,0730		max. 0,1200
Silizium Si	0,0200		max. 0,6000
Mangan Mn	1,3300		max. 2,0000
Phosphor P	0,0110		max. 0,0250
Schwefel S	0,0020		max. 0,0150
Aluminium Al	0,0460		min. 0,0150
Chrom Cr	0,0300		
Vanadium V	0,0110		max. 0,2000
Niob Nb	0,0440		max. 0,0900
Titan Ti	0,0870		max. 0,2200
Kupfer Cu	0,0100		
Nickel Ni	0,0400		
Stickstoff N	0,0050		
Bor B	0,0002		max. 0,0050

Data Acquisition

Datenerhebung

- **Heterogener Ist-Zustand:**
Die Materialzertifikate liegen als PDFs vor. Native PDFs von **Zulieferer 1** ermöglichen eine direkte Extraktion, während PDF-Scans von **Zulieferer 2** eine Bilderkennung erfordern.
- **Strukturierter Soll-Zustand:**
Das Ziel ist eine **Excel-Tabelle** mit einem Datensatz pro Zeile. Diese Struktur bildet die Basis für das KI-Training.
- **Spezifikation der Datenfelder:**
Der **Ziel-Datensatz** umfasst mechanische Kennwerte, chemische Bestandteile, logistische Daten und die Klassifizierung.



Data Acquisition

Datenerhebung

Güte : S650MC gebeizt, geölt
 Kundenartikelnr. : WB 28600 Lnr. 70911
 Abmessung in mm : 3,00 x 149,50
 Dickentoleranz in mm: +0,00/-0,20
 Liefergewicht kg : 3.986
 Ringanzahl : 4

Charge : 1304832
 W-Schmelze : 55-2371
 Werkscoilnr. : 193033

Mechanische Werte Prüfnorm: DIN EN ISO 6892-1 Prüfverfahren B

Name	Wert	Kundenvorschrift	DIN EN 10149-2
Streckgrenze (Re)	726,00 MPa		min. 650,00 MPa
Zugfestigkeit (Rm)	778,00 MPa		700,00 - 880,00 MPa
Bruchdehnung A	21,00 %		min. 12,00 %

Chemische Analyse

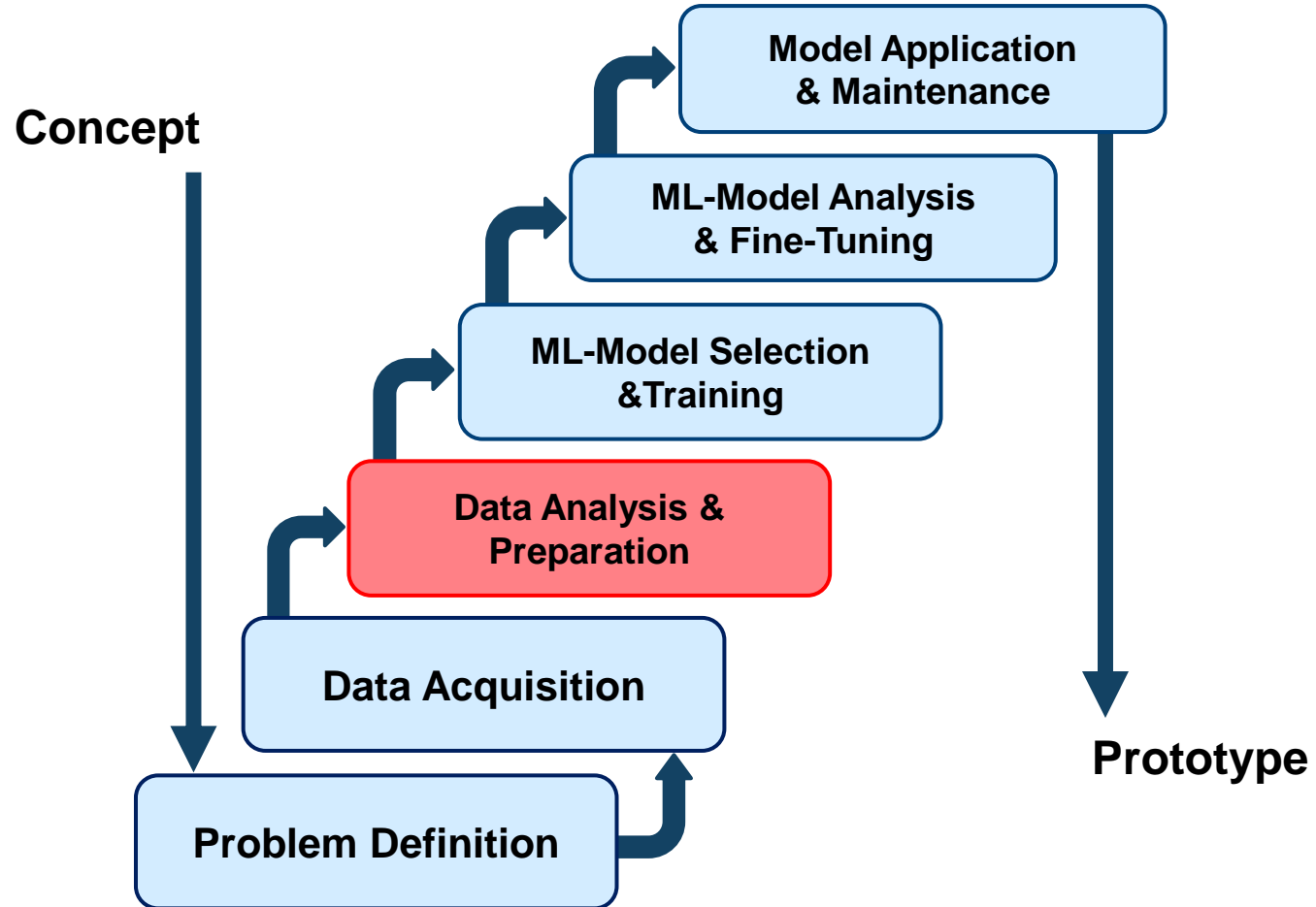
Element	Analysewert %	Kundenvorschrift	DIN EN 10149-2
Kohlenstoff	C 0,0730		max. 0,1200
Silizium	Si 0,0200		max. 0,6000
Mangan	Mn 1,3300		max. 2,0000
Phosphor	P 0,0110		max. 0,0250
Schwefel	S 0,0020		max. 0,0150
Aluminium	Al 0,0460		min. 0,0150
Chrom	Cr 0,0300		
Vanadium	V 0,0110		max. 0,2000
Niob	Nb 0,0440		max. 0,0900
Titan	Ti 0,0870		max. 0,2200
Kupfer	Cu 0,0100		
Nickel	Ni 0,0400		
Stickstoff	N 0,0050		
Bor	B 0,0002		max. 0,0050



Versandtag	Güte	Charge	W-Schmelze	Werkscoilnr.	Liefergewicht kg	Abmessung mm	Streckgrenze (Re)	Zugfestigkeit (Rm)
17.12.24	S650MC	g1304832	55-2371	193033	3.986	3,00 x 149,50	726.00	778.00
17.12.24	S650MC	g1304833	55-2370	193036	4.920	3,00 x 149,50	733.00	781.00
17.12.24	S650MC	g1304834	55-2371	193038	4.950	3,00 x 149,50	737.00	785.00
17.12.24	S650MC	g1304835	55-2370	193040	4.924	3,00 x 149,50	716.00	759.00
17.12.24	S650MC	g1306268	55-2371	193037	4.898	3,00 x 149,50	744.00	776.00
16.01.25	S650MC	g1304858	55-2370	193039	4.902	3,00 x 149,50	732.00	775.00
16.01.25	S650MC	g1304833	55-2370	193036	2.438	3,00 x 149,50	733.00	781.00
16.01.25	S650MC	g1304832	55-2371	193033	1.972	3,00 x 149,50	726.00	778.00
16.01.25	S650MC	g1314754	26-5526	203094A	11.998	2,50 x 195,00	736.00	791.00
16.01.25	S650MC	g1306267	55-2371	193035	2.414	3,00 x 149,50	755.00	780.00
10.02.25	S650MC	g1280739	26-3091	157635	1.216	3,00 x 149,50	747.00	801.00
10.02.25	S650MC	g1299263	26-5306	181417	4.862	3,00 x 149,50	727.00	785.00
10.02.25	S650MC	g1306269	55-2371	193034	2.366	3,00 x 149,50	756.00	787.00
10.02.25	S650MC	g1314753	26-6424	204494	17.250	3,00 x 149,50	703.00	763.00
12.02.25	S650MC	g1314752	26-6425	206789	16.974	3,00 x 149,50	714.00	773.00
17.02.25	S650MC	g1318002	26-6425	216226	9.862	3,00 x 149,50	723.00	775.00
17.02.25	S650MC	g1318002	26-6425	216226	7.388	3,00 x 149,50	723.00	775.00
17.02.25	S650MC	g1318004	26-6425	216225	17.164	3,00 x 149,50	727.00	785.00
27.02.25	S650MC	g1320147	027074	356432500	23.912	3,00 x 149,50	803.20	855.40
05.03.25	S650MC	g1321636	027791	434376500	8.302	3,00 x 149,50	738.00	800.00
05.03.25	S650MC	g1321636	027791	434376500	4.452	3,00 x 149,50	738.00	800.00
12.03.25	S650MC	g1321636	027791	434376500	8.302	3,00 x 149,50	738.00	800.00
12.03.25	S650MC	g1321636	027791	434376500	10.106	3,00 x 149,50	738.00	800.00
24.03.25	S650MC	g1321636	027791	434376500	4.122	3,00 x 149,50	738.00	800.00
24.03.25	S650MC	g1325249	26-6425	216224	4.794	3,00 x 149,50	727.00	778.00
26.03.25	S650MC	g1325249	26-6425	216224	12.384	3,00 x 149,50	727.00	778.00
08.04.25	S650MC	g1332252	51081	868652734	10.404	3,00 x 149,50	787.00	835.00
08.04.25	S650MC	g1332838	51081	868658586	9.360	3,00 x 149,50	764.00	841.00
08.04.25	S650MC	g1332838	51081	868658586	4.608	3,00 x 149,50	764.00	841.00

Workshop – Daten und KI in der Produktion

Data Analysis & Preparation



Differenzierte Datenqualität

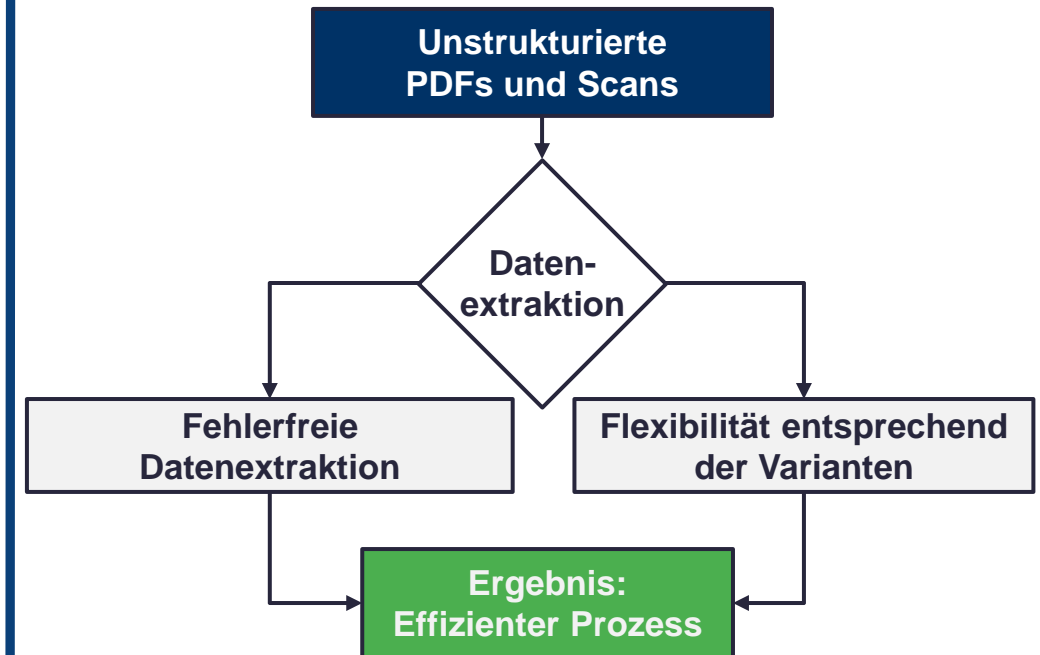
- **Native Dokumente (Zulieferer 1):**
Primärdaten mit Textlayer. Ermöglichen eine **direkte digitale** Extraktion und **Filterung** der Zielparameter.
- **Digitalisierte Scans (Zulieferer 2):**
Mittlere Datenqualität ohne Textlayer. **Erfordern** Bilderkennung sowie Vorverarbeitung zur **Korrektur** von Scan-Verzerrungen.

Skalierung des Datenumfangs

- **Validierter Kern:**
Initial **34 sicher klassifizierte** Datensätze von Zulieferer 1 (14 n.O.K. / 20 O.K.).
- **Erweiterter Datensatz:**
Durch Chargen-Matching und PDF-Auswertung **Steigerung** auf insgesamt **93** Zeugnisse (27 n.O.K. / 66 O.K.)

Ausgangssituation in vielen Produktionsunternehmen

- **Unstrukturierte Datenbasis:**
Werkstoffzeugnisse liegen in der Praxis als PDF oder Scan vor
- **Heterogene Layouts:**
Mehrere Zulieferer → **variierende** Dokumentenstrukturen
- **Prozessineffizienz:**
Notwendige Datenübernahme erfordert **händische** Prozesse und hohen **Zeitaufwand**
- **Fehleranfälligkeit:**
Manuelle Dateneingaben **generieren** Fehlerquellen und **gefährden** die Datenintegrität
- **Systemische Isolation:**
Qualitätsinformationen in zentralen Zielsystemen wie ERP oder MES **nicht** unmittelbar **verwertbar**

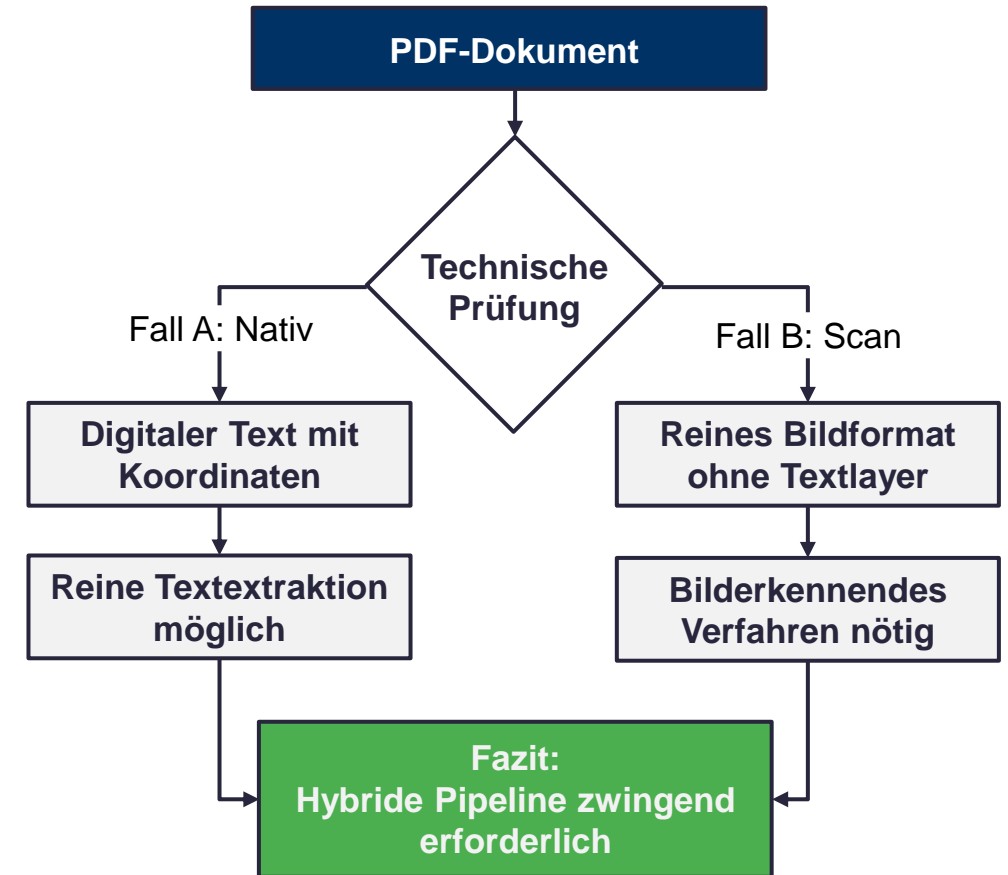


Kernfrage: Wie kann man aus unstrukturierten PDFs automatisiert strukturierte Daten erzeugen?

Data Analysis & Preparation

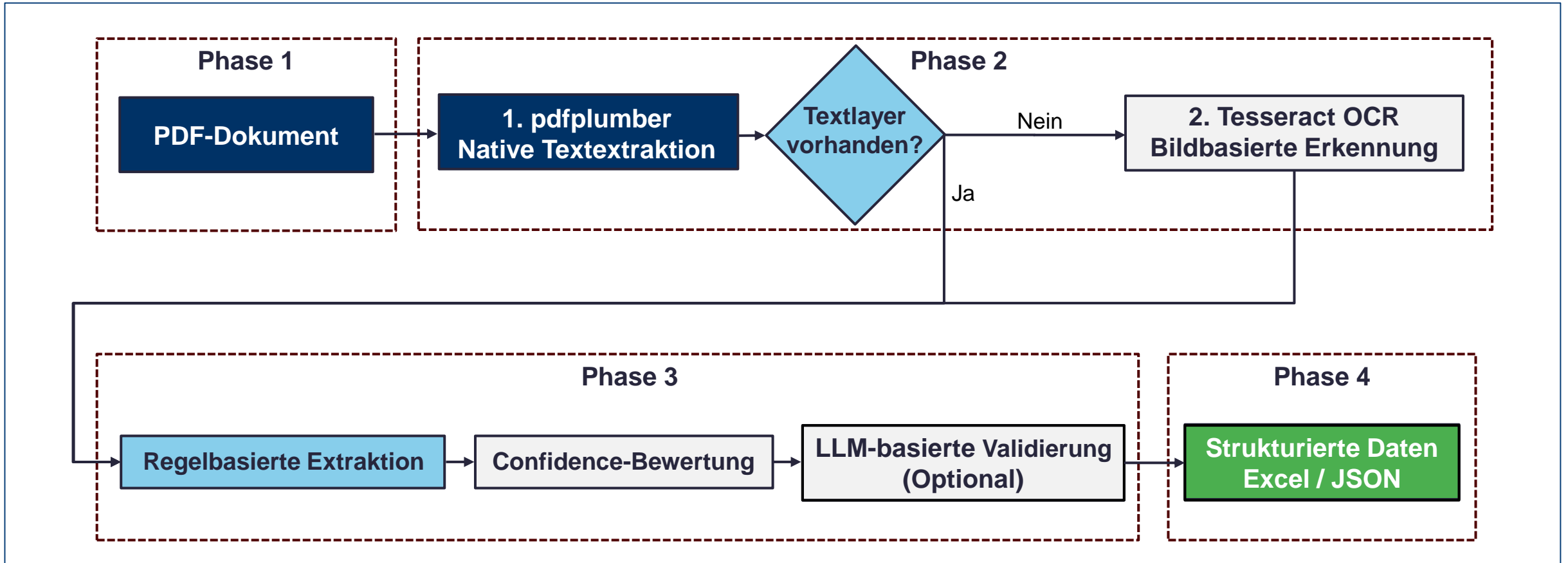
Die zentrale Herausforderung

- **Das Grundproblem:**
Ein Dokument im PDF-Format ist **nicht automatisch** maschinenlesbar oder digital auswertbar.
- **Fall A (Natives Text-PDF):**
Das Dokument enthält originären, digitalen Text inklusive Positionskordinaten. Dies ermöglicht eine **unkomplizierte, direkte** Datenextraktion.
- **Fall B (Gescanntes PDF):**
Das Dokument ist technisch eine reine Bilddatei (Pixelgrafik) ohne echten Text. Hier ist der Einsatz von **Optical Character Recognition (OCR)** zwingend erforderlich.
- **Wichtige Erkenntnis:**
Wegen der technischen Variierung der PDF-Qualität, erfordert eine verlässliche Verarbeitung immer eine **Kombination** mehrerer Technologien (Textextraktion + OCR).



Data Analysis & Preparation

Architektur der Pipeline



Wichtigstes Architekturprinzip: Der Erfolg basiert nicht „nur auf KI“, sondern auf der gezielten Kombination klassischer, deterministischer Methoden mit modernen KI-Verfahren.

Native Datenextraktion mit pdfplumber

- **Kernaufgabe:**
Das Modul liest Textinformationen direkt und deterministisch aus der digitalen Dateiebene des PDFs aus.
- **Technische Fähigkeiten:**
Zu den Hauptfunktionen zählen die Extraktion von eingebettetem Text, das Auslesen exakter Wortkoordinaten sowie die gezielte Analyse von Seiten- und Tabellenstrukturen.
- **Systemische Vorteile:**
Liegt ein natives Text-PDF vor, arbeitet das Tool schnell und hochpräzise. Da keine Bilderkennung stattfindet, werden OCR-Fehler vollständig vermieden und wertvolle Strukturinformationen bleiben erhalten.
- **Praktisches Anwendungsbeispiel:**
Der Algorithmus erkennt selbstständig den Beginn einer Tabelle, identifiziert spezifische Spalten (z. B. „Prüfwert“) und ordnet die dazugehörigen Messwerte rechts daneben relational korrekt zu.

Optische Zeichenerkennung (OCR) mit Tesseract

- **Definition & Tool:**
Optical Character Recognition wandelt Bildpixel in maschinenlesbaren Text um. Einsatz der leistungsstarken Open-Source-Engine **Tesseract**.
- **Leistungsumfang:**
Erkennt Buchstaben und Zahlen in Scans, PDFs und Fotos.
- **Datenstruktur:**
Liefert neben dem Textinhalt auch die **exakten** Positionsdaten der Zeichen.
- **Nutzen:**
Ermöglicht die **automatisierte** Auswertung von Dokumenten **ohne** digitalen Textlayer.

Güte	:	S650MC gebeizt, geölt
Kundenartikelnr.	:	R7000179000 ZTNr.: 7211
Abmessung in mm	:	3,00 x 149,50
Dickentoleranz in mm	:	+0,10/-0,20
Liefergewicht kg	:	9.876
Ringanzahl	:	10
Mechanische Werte		Prüfnorm: DIN EN ISO 6892-1
Name		Wert
Streckgrenze (Re)		718,00 MPa
Zugfestigkeit (Rm)		783,00 MPa
Bruchdehnung A		25,50 %

Streckgrenze (Re) 718,00 Mpa

Key-Learnings – Die Stärken der hybriden Architektur

- **Limitation von pdfplumber:**
Agiert exzellent bei **nativen** Digital-PDFs, bleibt jedoch bei **bildbasierten** Scans ohne verwertbares Ergebnis.
- **Limitation reiner OCR:**
Ermöglicht zwar die Bilderkennung, ist jedoch **fehleranfälliger** und führt oft zum **Verlust** komplexer Tabellenstrukturen.

Werkzeug	Stärke
pdfplumber	Struktur & Präzision
OCR	Verarbeitung von Scans
Kombination	Robustheit im realen Industrieinsatz

Integratives Architektur-Fazit: Der maximale Digitalisierungserfolg resultiert aus der intelligenten Orchestrierung spezialisierter Werkzeuge statt aus isolierten KI-Einzellösungen.

Data Analysis & Preparation

Regelbasierte Datenextraktion

Nach der Texterkennung erfolgt die Überführung in strukturierte Daten. Hierbei setzt die Pipeline konsequent auf einen **regelbasierten Ansatz**.

Vorteile des regelbasierten Vorgehens

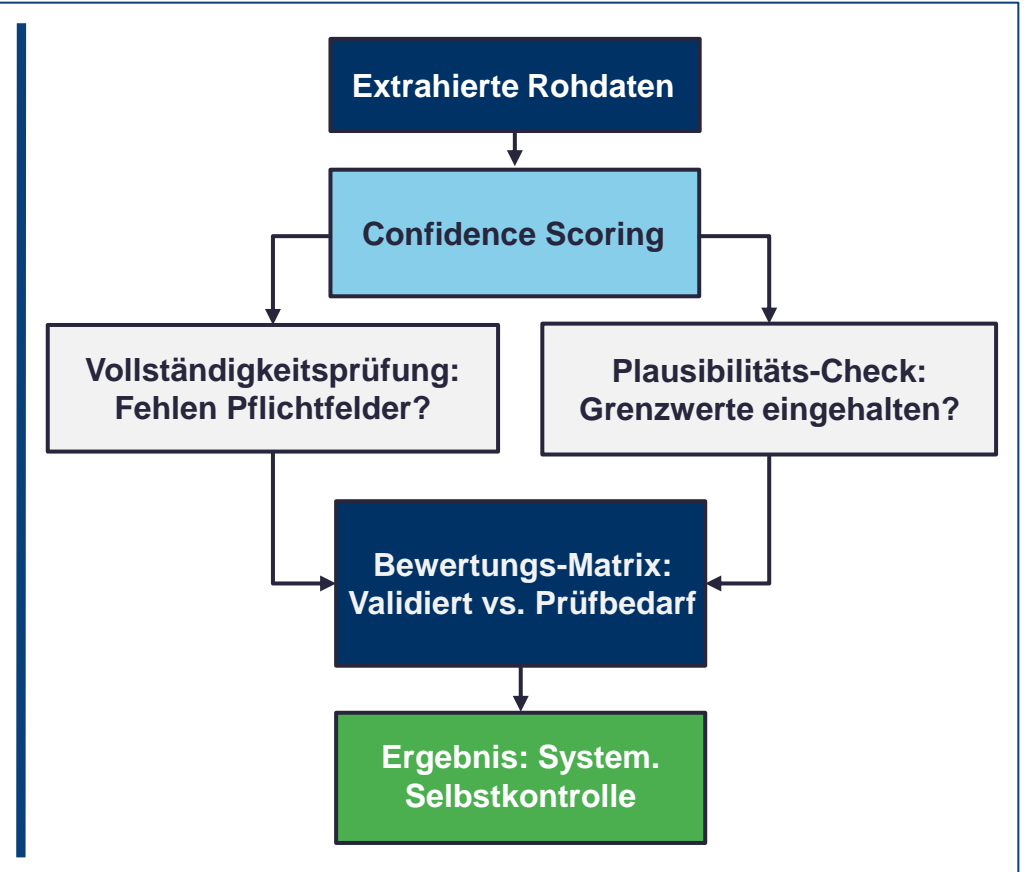
- **Deterministisch & Reproduzierbar:**
Identische Inputs führen immer zu identischen Ergebnissen.
- **Validierbar:**
Die Extraktionslogik ist für Qualitätsprüfer jederzeit nachvollziehbar und transparent.
- **Stabilität:**
Bei bekannten Layouts bietet dieser Ansatz eine deutlich höhere Zuverlässigkeit als reine KI-Modelle.

Zentrales Learning: Für hochgradig strukturierte technische Dokumente sind klassische, logikbasierte Regeln oft performanter und sicherer als KI-Lösungen.

Data Analysis & Preparation

Confidence Scoring

- **Bei einer Confidence von z.B. 75%:**
Regelbasierte **Extraktion** wird gezielt zu den fehlenden / unplausiblen Werten wiederholt
- **Bei einer Confidence von z.B. 25%:**
Die fehlenden / unplausiblen Werte werden per LLM extrahiert, es erhält den extrahierten **Textlayer** mit Strukturinformationen und wird gezielt zu den fehlenden Werten befragt
- **Der Output des LLMs wird ebenfalls geprüft:**
Liegt die daraus erreichte **Confidence** unter der vorherigen, regelbasierten, wird die regelbasierte genutzt.



Zentrales Key-Learning: Eine verlässliche industrielle Automatisierung funktioniert niemals blind, sondern benötigt zwingend eine integrierte Selbstkontrolle.

Data Analysis & Preparation

Rolle des LLMs in der Pipeline

- **Zentrales Learning:**
Das LLM fungiert nicht als Hauptlogik.
Für industrielle Datenintegrität ist eine rein generative Extraktion aufgrund von **Inkonsistenzen** und **Halluzinationsrisiken** ungeeignet.
- **Limitationen generativer KI:**
LLMs reagieren **empfindlich** auf unstrukturierten OCR-Output und interpretieren Tabellenstrukturen oft **nicht deterministisch genug**.
- **Gezielte Einsatzgebiete:**
Das Modell dient ausschließlich als **Fallback-Ebene**, zur **finalen Validierung** und zur **intelligenten Ergänzung fehlender Datenfelder**.

Architektur-Entscheidung:

- **Primärsystem:**
Regelbasierte Extraktion (Sichert Stabilität & Präzision).
- **Supportsystem:**
LLM (Bietet Flexibilität & Validierung).

Data Analysis & Preparation

Hybride Validierung

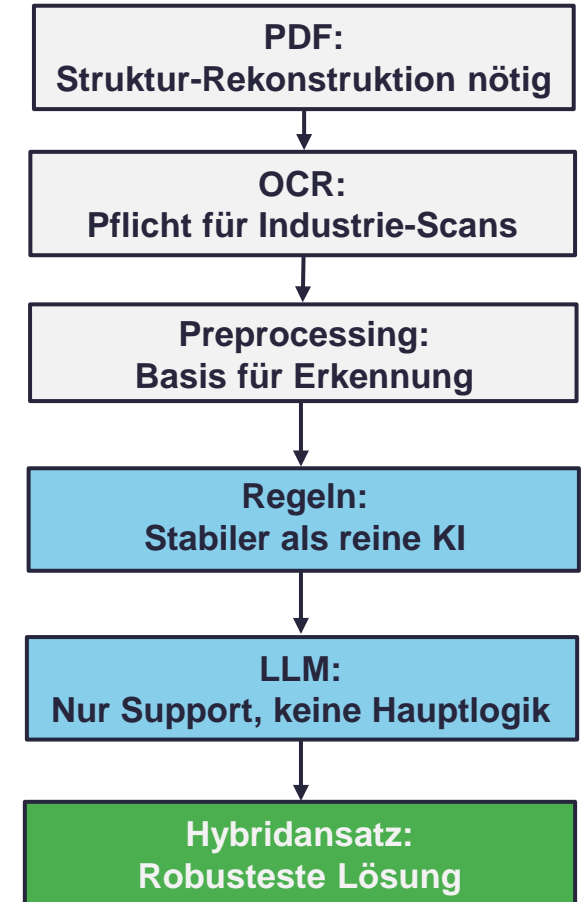
- **Dualer Ansatz:**
Die Kombination beider Welten **sichert** industrielle Robustheit und Zuverlässigkeit.
- **Klassische Logik (Stabilität):**
Übernimmt die **präzise Extraktion** von Tabellen, strikte Regelwerke und automatisierte Plausibilitätsprüfungen.
- **KI-Komponente (Flexibilität):**
Ideal für die **Interpretation** von Sonderfällen, unbekanntem Layouts und die Ergänzung fehlender Werte.
- **Synergie-Effekt:**
Stabilität durch deterministische Regeln **kombiniert** mit der Anpassungsfähigkeit generativer Modelle.

Klassische Logik	KI / LLM
Stabil	Flexibel
Deterministisch	Kontextfähig
Schnell	Intelligent
Nachvollziehbar	Kreativ

Data Analysis & Preparation

Technische Key-Learnings

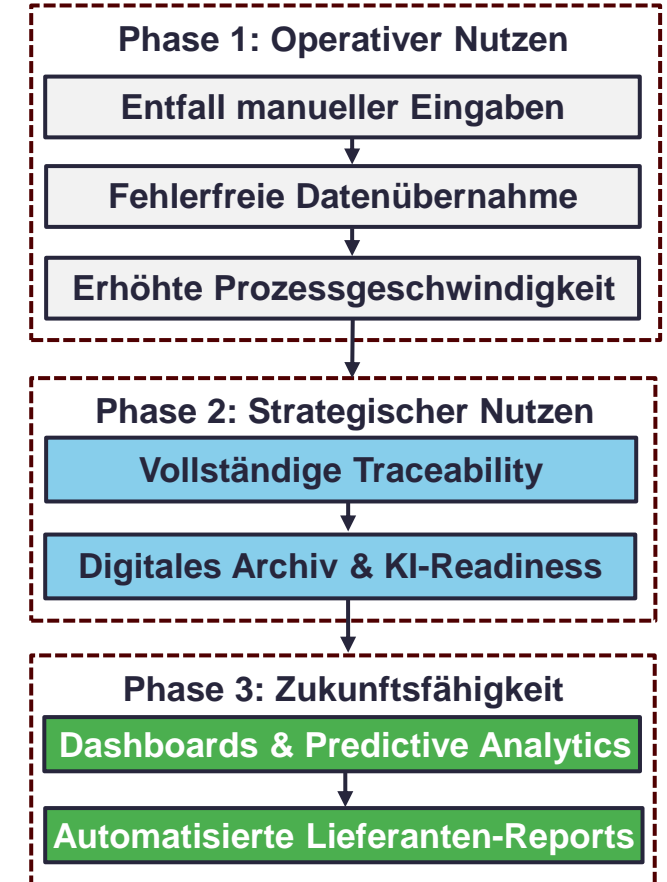
- **PDF-Hürde:**
PDFs sind Layout-Container, keine strukturierten Datenformate. Die Rekonstruktion ist **komplexer** als erwartet
- **OCR-Pflicht bei gescannten PDFs:**
Wenn Industriedokumente auch als Scan vorliegen, ist optische Texterkennung (OCR) **unverzichtbar**
- **Bild-Tuning:**
OCR-Qualität steht und fällt mit der Bildvorverarbeitung (Schärfe, Ausrichtung)
- **Logik vor KI:**
Deterministische Regeln sind bei Tabellen und technischen Daten **präziser und performanter** als KI
- **LLM-Rolle:**
LLMs sind exzellente **Assistenten** zur Validierung, aber riskante **Solisten** (Halluzinationsgefahr)
- **Hybrid-Erfolg:**
Die Kombination aus regelbasiertem Vorgehen und KI bietet die **höchste Robustheit**



Data Analysis & Preparation

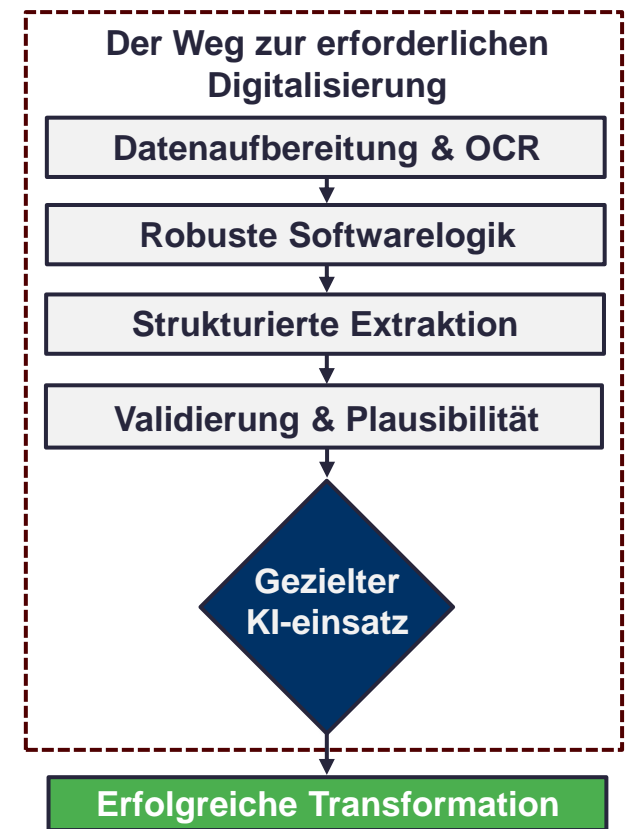
Nutzen in der Praxis

- **Operative Exzellenz:**
Massive Reduktion manueller Aufwände und Übertragungsfehler bei gleichzeitiger Beschleunigung der Dokumentenverarbeitung.
- **Strategische Befähigung:**
Aufbau einer lückenlosen **Rückverfolgbarkeit** und die **Digitalisierung** historischer Bestände als Basis für künftige KI-Anwendungen.
- **Langfristige Wertschöpfung:**
Transformation von Dokumenten in valide Datensätze für Dashboards, Lieferantenvergleiche und Predictive Analytics.
- **Datenverfügbarkeit:**
Strukturierte Informationen stehen abteilungsübergreifend in Echtzeit für **automatisierte** Reports und Qualitätsanalysen zur Verfügung



Fazit

- **Keine isolierte KI-Lösung:**
Erfolgreiche Digitalisierung in der Produktion ist kein Ergebnis eines einzelnen Werkzeugs, sondern das Resultat einer abgestimmten **Prozesskette**.
- **Das Fundament:**
Erst die Kombination aus **Datenaufbereitung** (OCR/Preprocessing) und robuster **Softwarelogik** schafft die notwendige Stabilität.
- **Struktur vor Interpretation:**
Regelbasierte **Extraktion** und konsequente **Validierung** sichern die industrielle Präzision, bevor generative Elemente unterstützen.

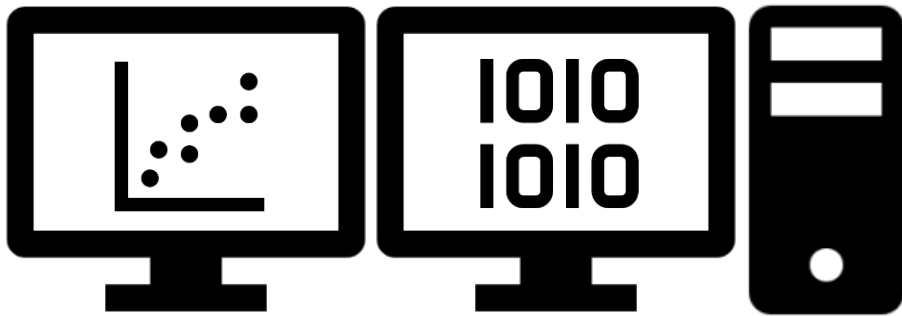


Abschlussbotschaft: KI entfaltet ihre maximale Wirkung nicht einzeln, sondern als gezielt integrierter Baustein innerhalb eines stabilen, validierten Gesamtprozesses.

Tips zur Erstellung von Python-Pipelines via Vibe-Coding

1. **Das Ziel klar definieren:** „Ich möchte eine fertige, direkt ausführbare Python-Pipeline die XY macht“
 2. **Die Umgebung klar definieren:** „Ich arbeite dabei mit Anaconda“ (Empfehlung, alternativ auch Docker)
 3. **Datenformate definieren:** „Ich will, dass du Tabelleninhalte aus PDFs mit Textlayer zeilenweise in eine Excel extrahierst“
 4. **Bei Fehlern:** Fehlermeldungen aus der Konsole kopieren und der KI schicken.
 5. **Bei Installationsproblemen der Packages:** Ebenfalls KI fragen und ggfs. Umgebung nochmal einbringen. Hier gibt es teils unterschiedliche Befehle zur Installation
- ✓ **Wieso Anaconda oder ein Docker-Container?** Virtuelle Umgebungen die sich als „Spielplatz“ zur Erprobung gut eignen. Dadurch werden z.B. Versionskonflikte installierter Packages vermieden.

Code Exkurs



Data Analysis & Preparation

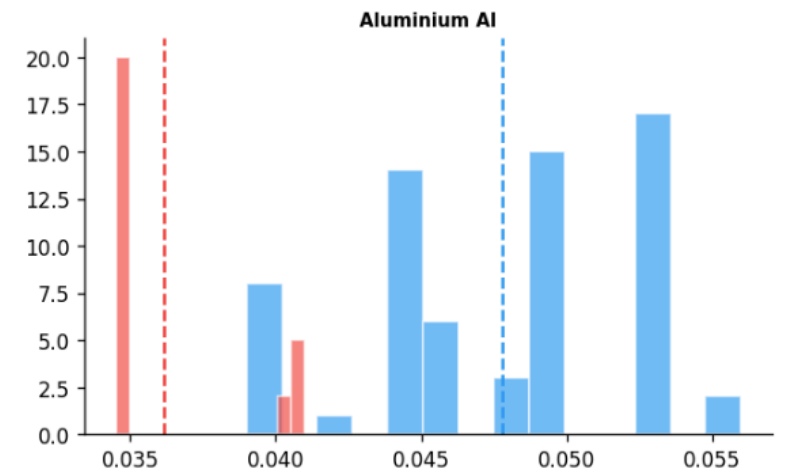
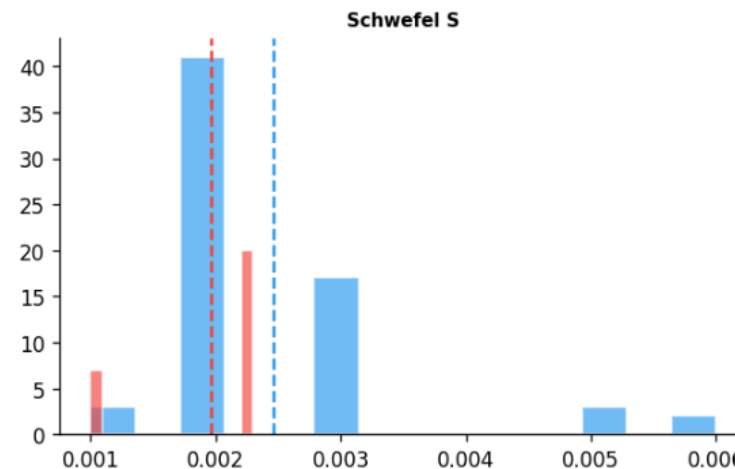
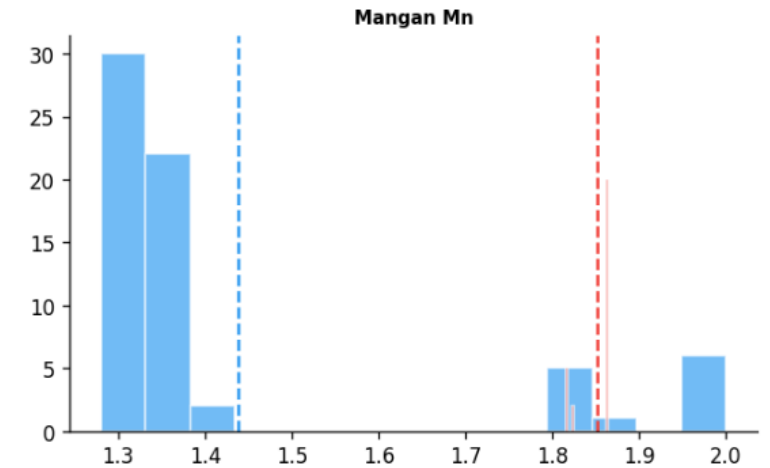
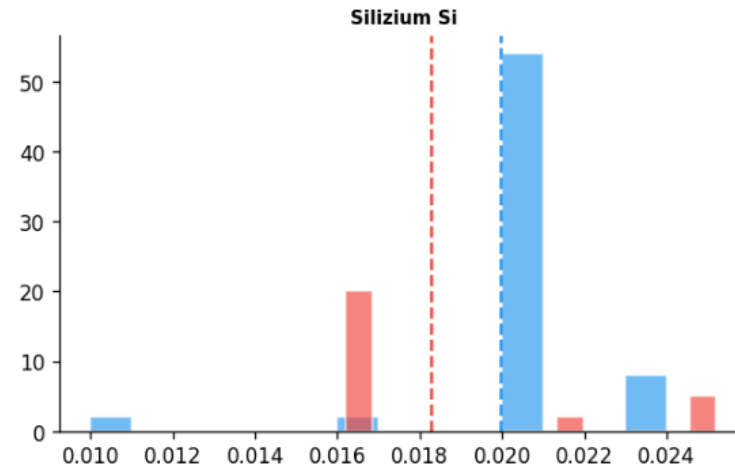
Anleitung – Start in Google Colab

- **Schritt 1: Vorbereitung**
 - Melden Sie sich in Ihrem **Google-Konto** (bei Bedarf bitte ein kostenloses Konto erstellen).
- **Schritt 2: Datei Download**
 - Laden Sie die Python-Notebook-Datei (**.ipynb**) sowie die **Excel-Datendatei** auf Ihren lokalen Rechner herunter.
- **Schritt 3: Notebook in Colab öffnen**
 - Rufen Sie den folgenden Link auf:
https://drive.google.com/drive/folders/1efNbHEBAqIOuPQHqoQjY0qD7Z_mHpdNe?usp=drive_link
 - Navigieren Sie im Menü zu: **Datei** → **Notebook hochladen**.
 - Klicken Sie auf **Durchsuchen** und wählen Sie die heruntergeladene **.ipynb-Datei** aus.

Data Analysis & Preparation

Datenanalyse im Colab Worksheet

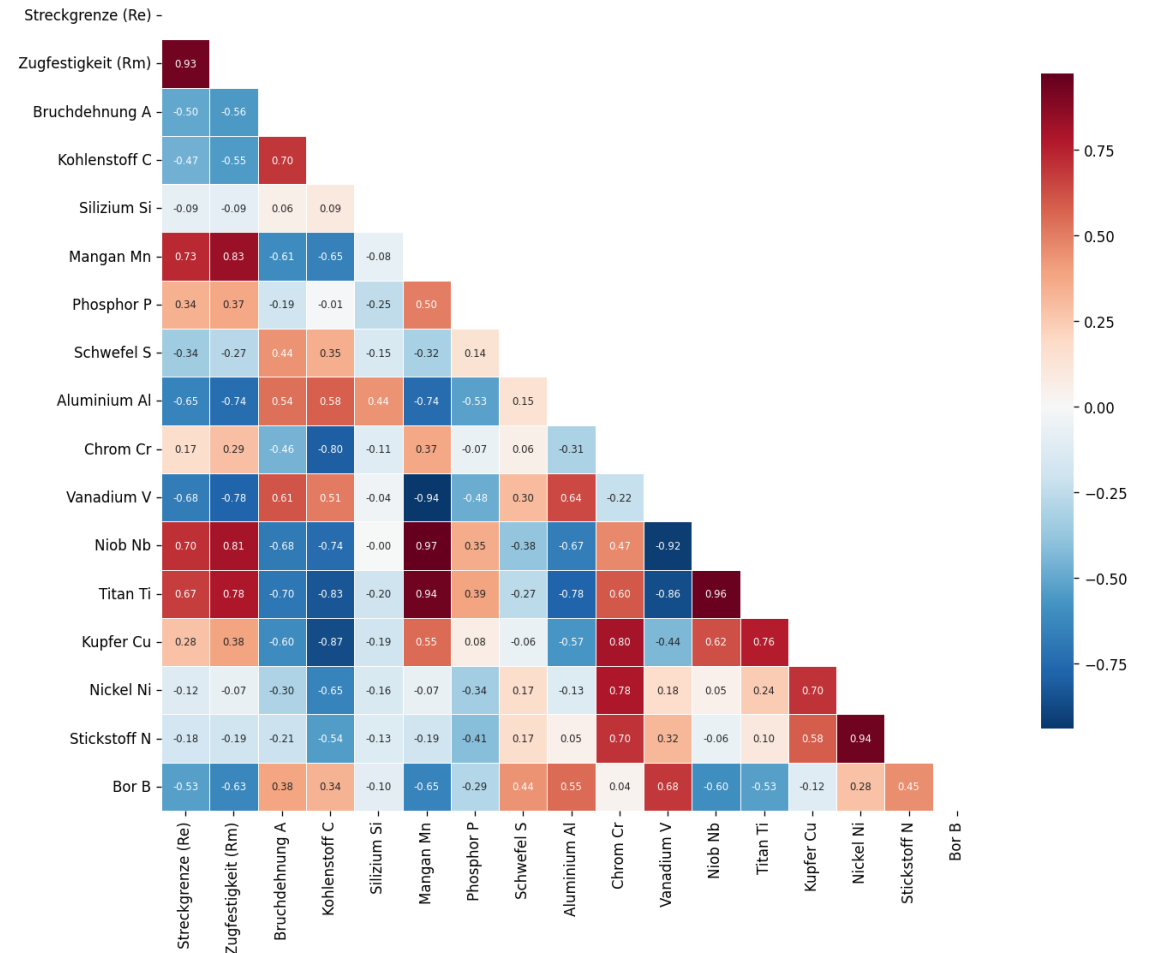
- Die statistische Auswertung zeigt **Auffälligkeiten**
- Sie ermöglicht die **Identifizierung** von relevanten Merkmalen
- Sie **bestätigt** vorher aufgestellte Vermutungen



Data Analysis & Preparation

Datenanalyse im Colab Worksheet

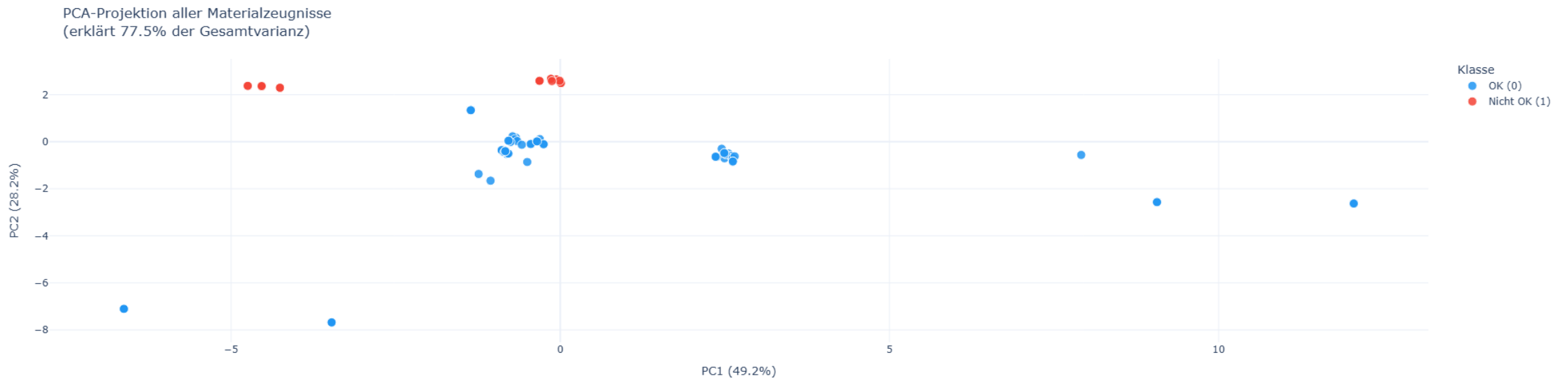
- Die Korrelationsmatrix zeigt Korrelationen der einzelnen **Datenpunkte zueinander**
- Korrelierende Eigenschaften können ggf. **zusammengefasst** werden



Data Analysis & Preparation

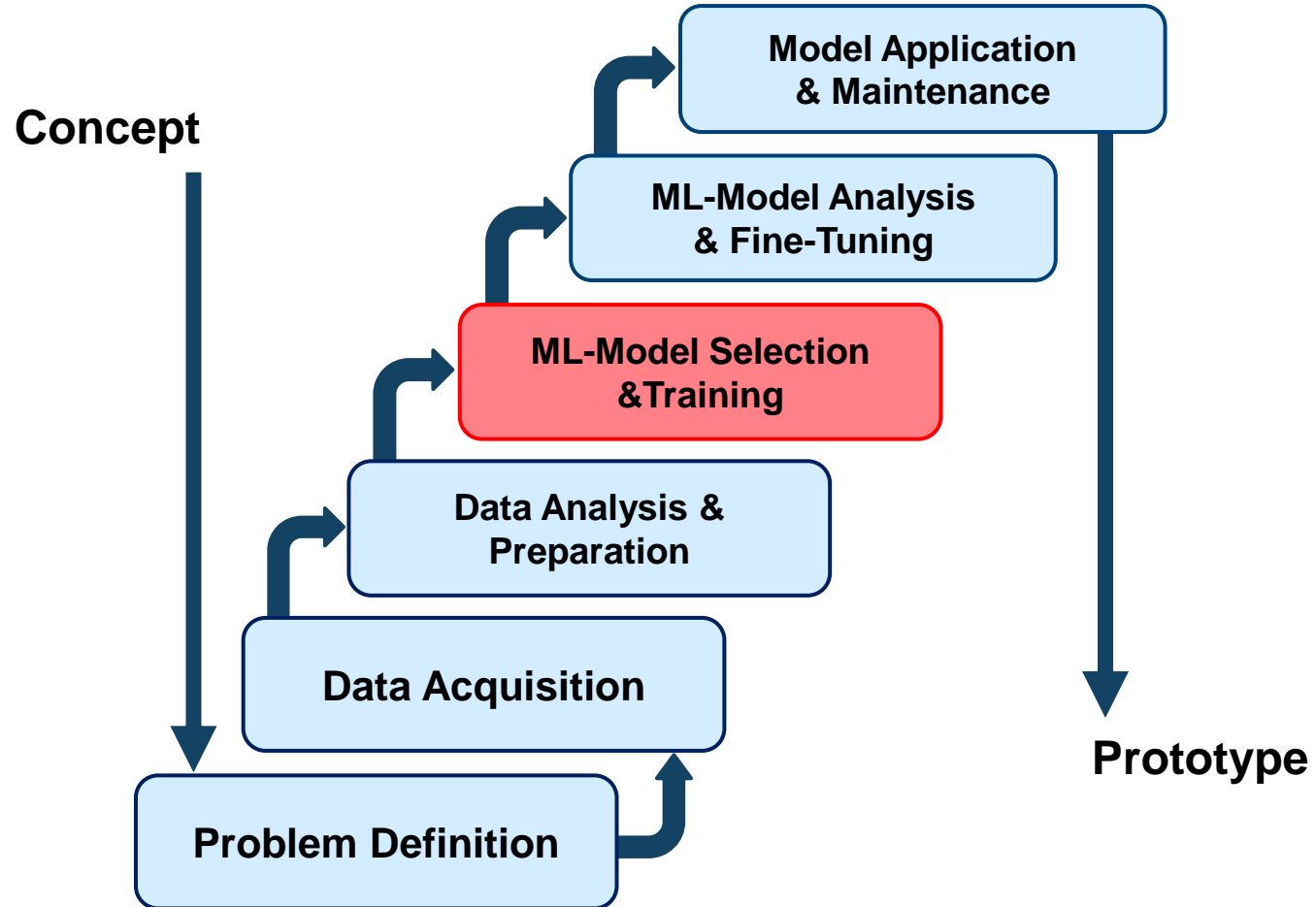
Datenanalyse im Colab Worksheet

- Die PCA-Projektion projiziert mehrdimensionale Merkmale in ein zweidimensionales Diagramm
- **Ziel:** Trennbarkeit zwischen den Merkmalsgruppen
- **Falls nicht vorhanden:** Mehr Daten bereitstellen und / oder Features variieren



Workshop – Daten und KI in der Produktion

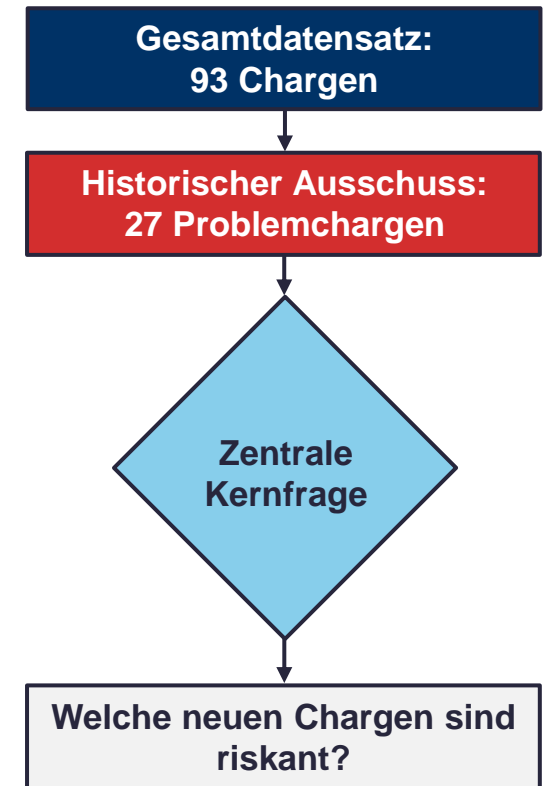
ML-Model Selection & Training



ML-Model Selection & Training

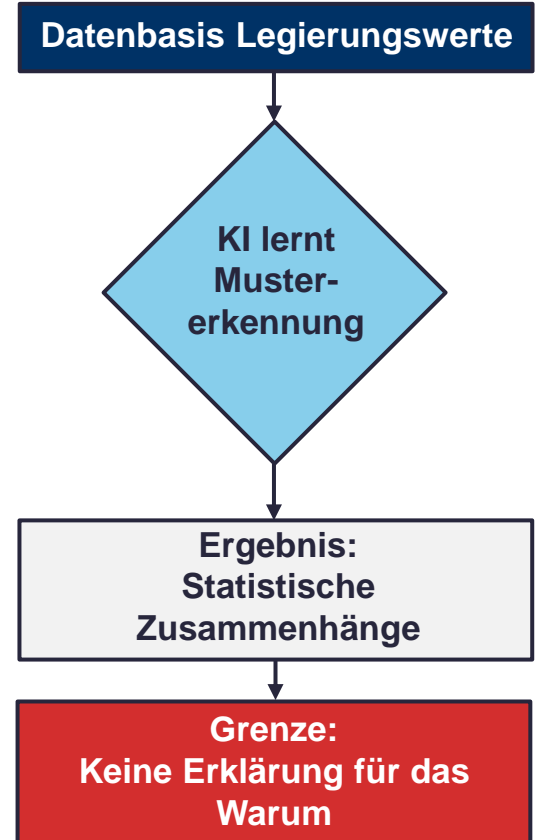
Quantifizierung des Problems

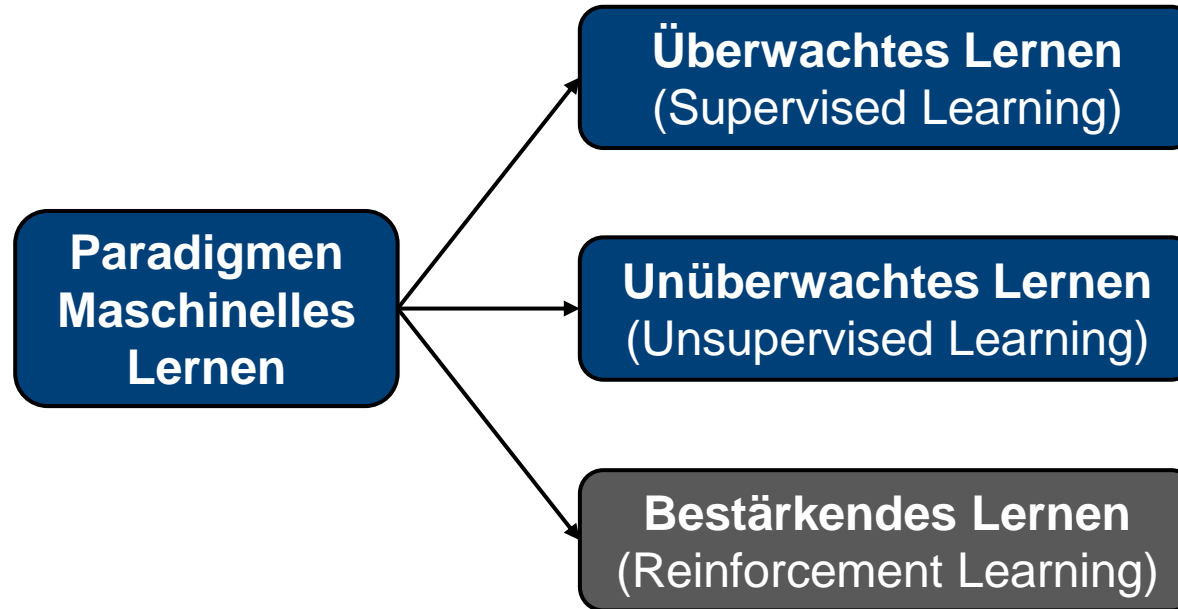
- **Historische Daten:**
Die Analyse stützt sich auf einen bestehenden Datensatz von insgesamt **93 Materialchargen**
- **Fehlerquote:**
Darunter befinden sich **27 identifizierte Problemchargen**, die in der Vergangenheit zu Ausschuss geführt haben
- **Die Kernfrage (Prädiktion):**
Wie lassen sich aus dieser Datenbasis zuverlässige Muster ableiten, um vorherzusagen, **welche neuen Chargen riskant sind?**



Funktionsweise – Was genau lernt die KI?

- **Die Kernkompetenz:**
Die KI lernt ausschließlich, komplexe Muster und Auffälligkeiten in den vorliegenden Materialdaten zu erkennen
- **Fehlendes Kausalverständnis:**
Das Modell besitzt kein echtes "Verständnis" für die physikalischen oder metallurgischen Ursachen hinter den Daten.
- **Wichtige Erkenntnis (Grenzen der KI):**
Die KI liefert keine Erklärung für das *Warum* (keine Ursachenforschung). Die Ergebnisse basieren rein auf statistischen Zusammenhängen (Korrelation, nicht Kausalität).





ML-Model Selection & Training

Systemvergleich der Algorithmen

Modell	Braucht Labels	Typ	Vorhersagequalität	Besonderheit
Isolation Forest	Nein	Unsupervised	Mittel	Erkennt auch unbekannte Fehlertypen
Local Outlier Factor	Nein	Unsupervised	Mittel	Gut bei lokal geclusterten Anomalien
Elliptic Envelope	Nein	Unsupervised	Schwach	Nur bei normalverteilten Daten sinnvoll
Random Forest	Ja	Supervised	Stark	Kombination von Entscheidungsbäumen

Der **Random Forest** übertrifft unüberwachte Ansätze durch die gezielte Nutzung historischer Fehlermuster und bietet somit die höchste operative Vorhersagesicherheit für die Qualitätssicherung.

ML-Model Selection & Training

Random Forest Classifier

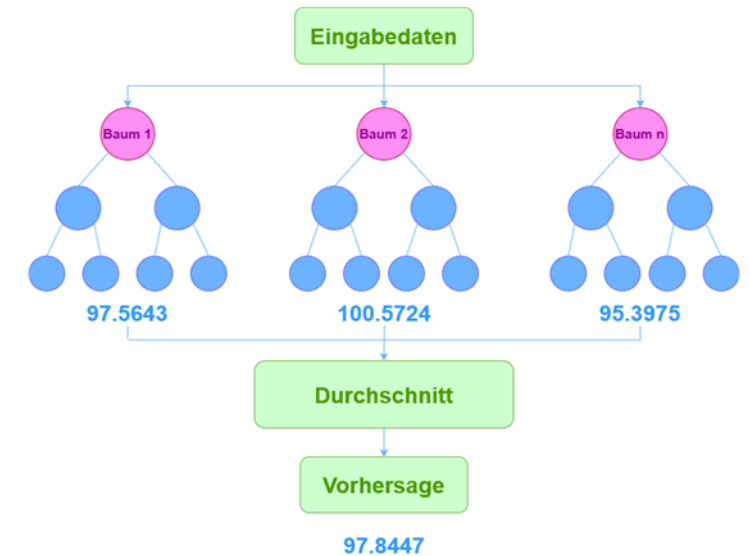
- **Entscheidungsbäume**
- **Jeder Baum bekommt parallel einen Teil der Daten und trifft daraufhin Entscheidung**
- **Gesamtergebnis über Mittelwert**

Robust:

- Durch Menge an Entscheidungsbäumen wird Varianz verringert
- Führt zu stabilen, generalisierbaren Vorhersagen
- Wenig anfällig für Ausreißer

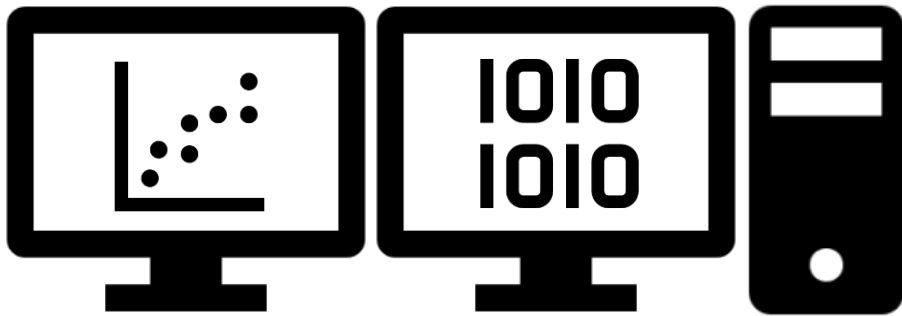
Keine Skalierung notwendig:

- Verarbeitung „roher“ Daten
- Ersparnis an Code und Rechenzeit



Der Random Forest Classifier bietet durch seine Kombination von Einfachheit, Vielseitigkeit und Effektivität eine leistungsfähige Lösung für z.B. binäre Klassifikationsaufgaben.

Code Exkurs



Kategorien von Klassifikationsergebnissen

Status	Beschreibung	Zustand	Vorhersage
True Positive (TP)	Modell sagt ein positives Ergebnis korrekt voraus.	Positiv	Wahr
True Negative (TN)	Modell sagt ein negatives Ergebnis korrekt voraus.	Negativ	Wahr
False Positive (FP)	Modell sagt ein positives Ergebnis falsch voraus.	Positiv	Falsch
False Negative (FN)	Modell sagt ein negatives Ergebnis falsch voraus.	Negativ	Falsch

ML-Model Selection & Training

Wesentliche Metriken zur Gütebewertung

Metrik	Aussagekraft	Berechnung
Accuracy	Wie oft liegen wir mit unserer Vorhersage zu den einzelnen Klassen richtig? Ziel: Hohe Genauigkeit bei der Bestimmung der Schuldigen und Nicht-Schuldigen.	$\frac{TP + TN}{TP + FP + FN + TN}$
Precision	Wie oft haben wir die korrekte Vorhersage von allen je Klasse positiv gemachten Vorhersagen gemacht? Ziel: Die Inhaftierten sollen auch wirklich schuldig sein.	$\frac{TP}{TP + FP}$
Recall	Wie oft haben wir die richtige Vorhersage von allen eigentlich richtigen Vorhersagen gemacht? Ziel: Wir wollen alle Schuldigen inhaftieren.	$\frac{TP}{TP + FN}$
F1-score	Mischung aus Recall und Precision	$(1+\beta^2)(Precision*Recall)/(\beta^2 *Precision+Recall)$ mit $\beta = 1$

Modellevaluation – Die Gefahr der „Accuracy-Falle“

- **Asymmetrische Datenverteilung:**
Die Datenbasis ist mit **66 fehlerfreien Chargen** (71 %) sowie **27 Problemchargen** (29 %) stark unbalanciert.
- **Limitation der Accuracy:**
Die isolierte Betrachtung der Vorhersagegenauigkeit ist **irreführend**.
Die bloße Vorhersage der Majoritätsklasse erzeugt **ohne echten Lerneffekt** sofort eine scheinbare Genauigkeit von 71 %.
- **Praktisches Risiko:**
Diese Metrik maskiert das systematische **Ignorieren** kritischer Problemchargen und **verfehlt** den industriellen Nutzen.
- **Methodische Lösung:**
Zur validen Modellevaluation wird der **F1-Score** angewendet. Dieser garantiert die zuverlässige Identifikation der relevanten Minderheitsklasse.

Konfusionsmatrix und Kategorisierung der Vorhersagen

- **Struktur der Matrix:**

Die Matrix unterteilt alle Vorhersagen in vier **spezifische Felder** (TP/FP/TN/FN).

- **Ideale Verteilung:**

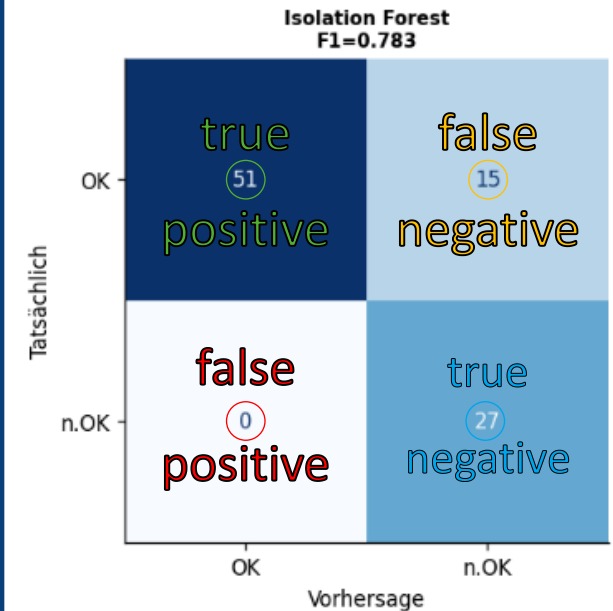
Korrekte Vorhersagen finden sich auf der **Hauptdiagonalen**. Das analytische Ziel erfordert die Maximierung exakt dieser Treffer.

- **Fehlklassifikationen:**

False Positive und False Negative bezeichnen inkorrekte Zuweisungen. Diese Fehler bedürfen einer **systematischen Minimierung**.

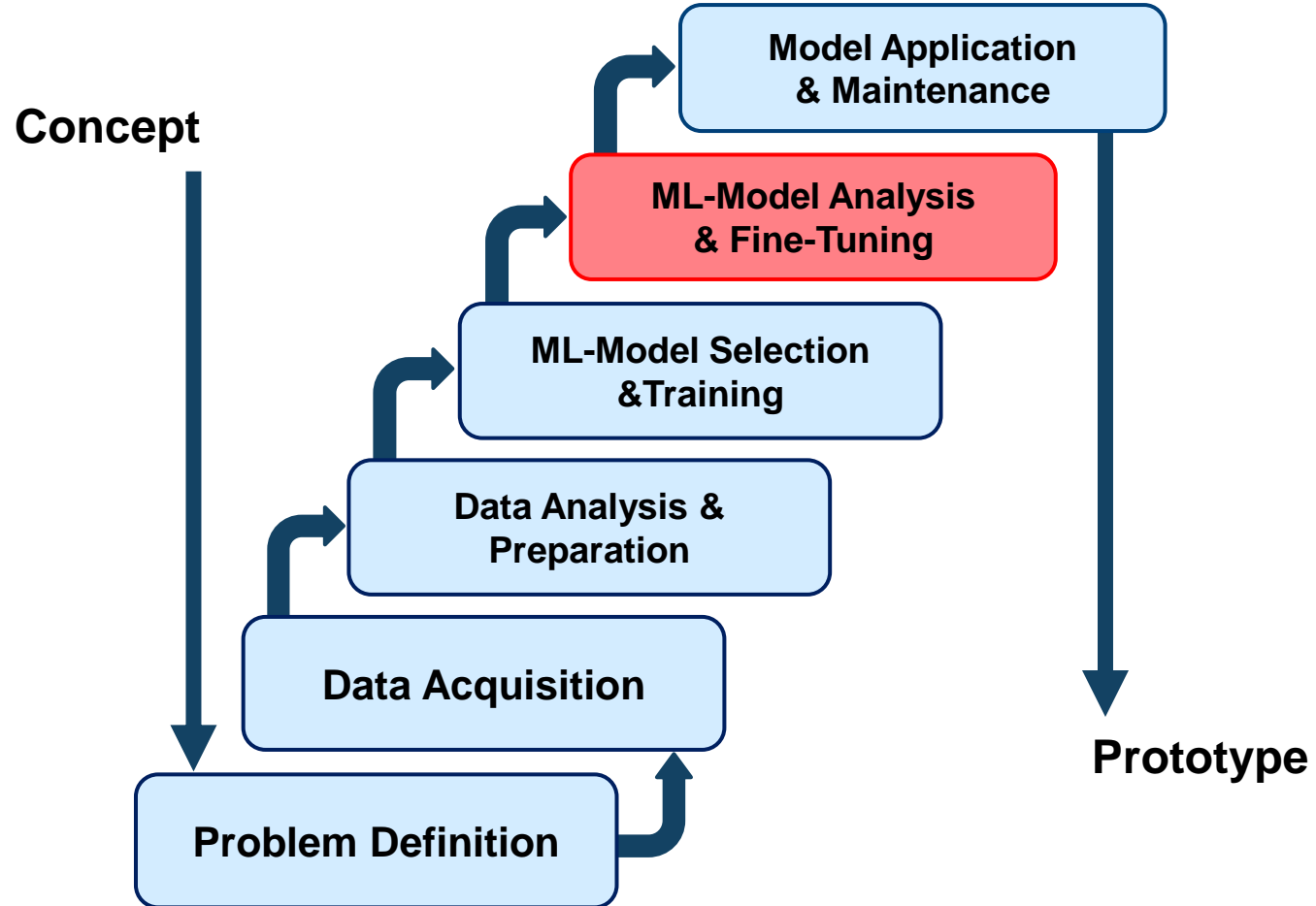
- **Wirtschaftliches Risiko:**

Die irrtümliche Freigabe einer fehlerhaften Charge als einwandfrei verursacht die **höchsten Produktionskosten**. Dieses spezifische Szenario erfordert höchste präventive Aufmerksamkeit.

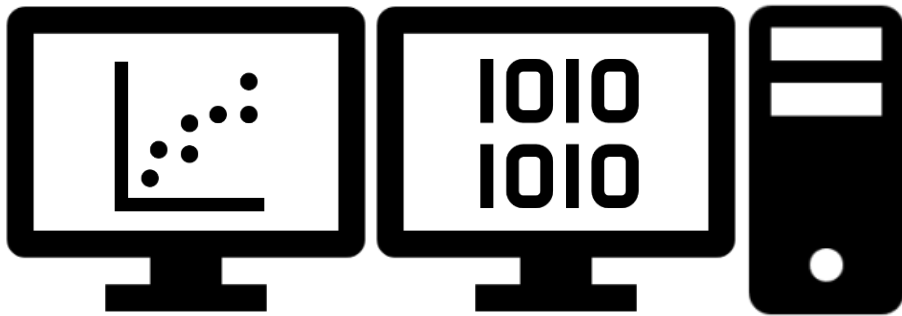


Workshop – Daten und KI in der Produktion

ML-Model Analysis & Fine-Tuning



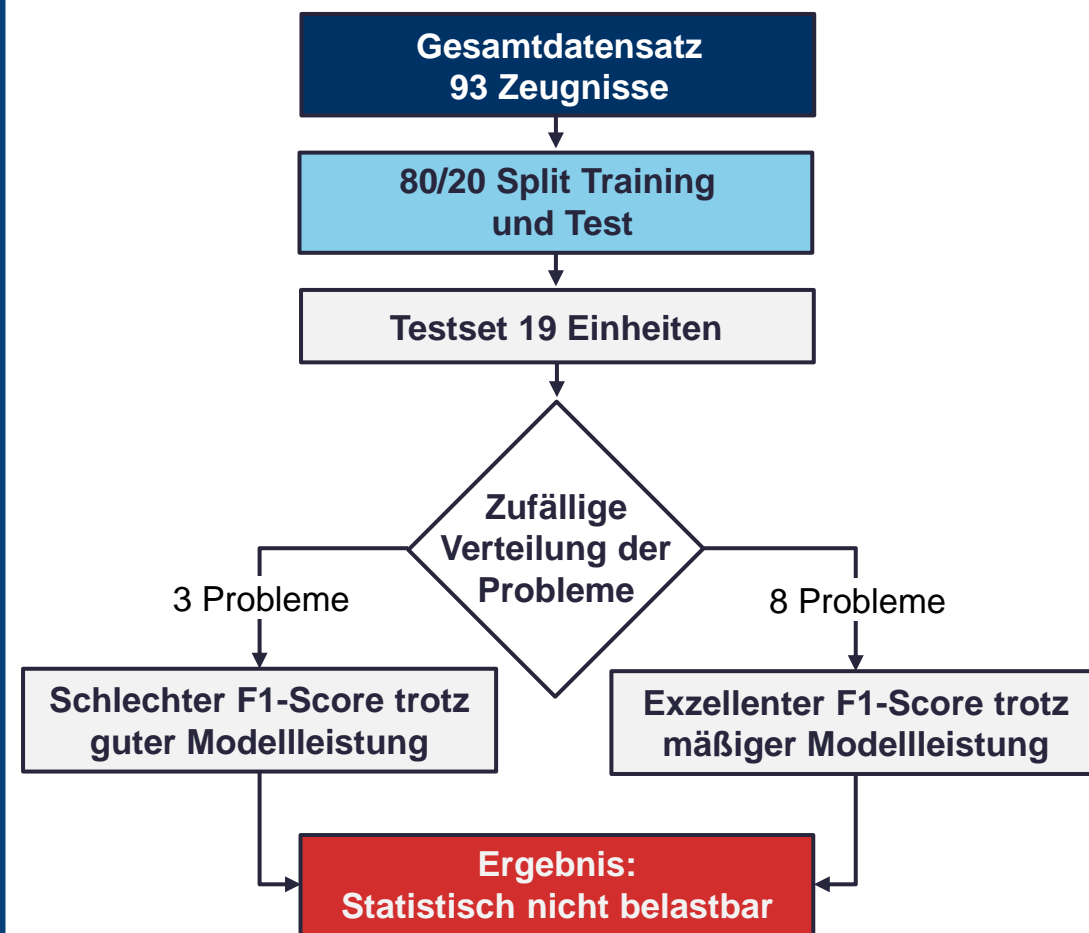
Code Exkurs



ML-Model Analysis & Fine-Tuning

Validierungsstabilität bei limitierten Datenmengen

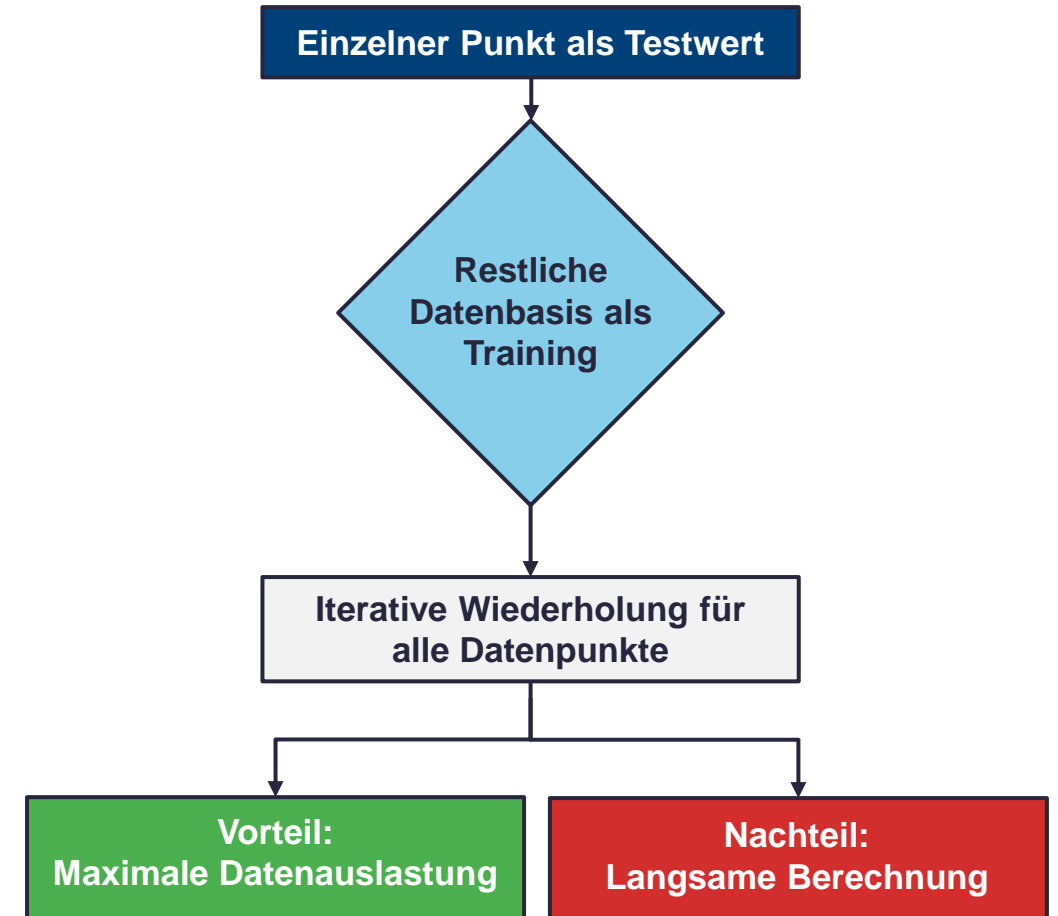
- **Statistische Varianz:**
Ein klassischer **80/20 Split** verursacht bei kleinen Datensätzen erhebliche **Schwankungen** der Ergebnisse.
- **Geringe Stichprobengröße:**
Das Testset von lediglich **19** Einheiten bietet **keine belastbare Basis** zur Evaluierung seltener Problemchargen.
- **Zufallseffekte:**
Die **zufällige Verteilung** kritischer Chargen im Testset **verzerrt** den F1 Score sowie die wahrgenommene Modellgüte massiv.
- **Mangelnde Validität:**
Das Ergebnis spiegelt primär den **Zufall der Datenaufteilung** statt der tatsächlichen Leistungsfähigkeit des Modells wider.



ML-Model Analysis & Fine-Tuning

Methodik der Leave One Out Cross Validation

- **Grundprinzip:**
Ein einzelner Datenpunkt dient isoliert als **Testwert**. Die restliche Datenbasis wird vollständig für das **Training** des Modells verwendet.
- **Iterativer Prozess:**
Dieser Vorgang wird für **jeden einzelnen Punkt** des
- **Methodischer Vorteil:**
Das Verfahren garantiert die **maximale Auslastung** der verfügbaren Informationen.
Es stellt die **exakteste Methode** zur Validierung kleiner Datensätze dar.
- **Systembedingter Nachteil:**
Iterativer Ansatz erfordert einen **hohen Rechenaufwand**.
Bei umfangreichen Datenmengen führt dies zu einer erheblichen **Verlangsamung der Prozesse**.



Das Phänomen Overfitting in der Modellentwicklung

- **Definition des Effekts:**
Das Modell erfasst die Trainingsdaten **zu detailliert** und integriert dabei auch **irrelevante zufällige** Schwankungen.
- **Praktische Auswirkungen:**
Die Auswertung liefert **exzellente** Ergebnisse bei vertrauten Trainingsdaten jedoch **mangelhafte** Prognosen bei neuen und unbekanntem Werten.
- **Fehlerhafter Lernprozess:**
Die Künstliche Intelligenz prägt sich **spezifische Eigenschaften** einzelner Chargen ein, anstatt allgemeingültige und übertragbare Muster zu extrahieren.
- **Zusammenfassendes Fazit:**
Das System hat die vorliegenden Informationen lediglich **auswendig gelernt**, anstatt die fundamentalen Zusammenhänge wirklich zu verstehen.



ML-Model Analysis & Fine-Tuning

Ergebnisse des Random Forest

- **Perfekte Klassifikation:**

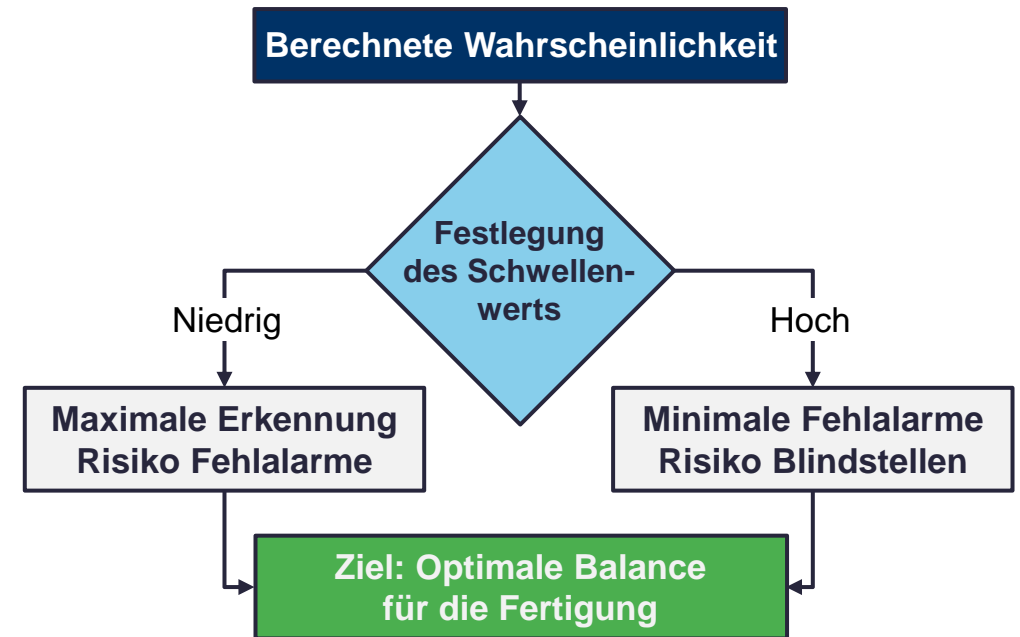
Die Evaluation weist durchgehend einen **F1-Score von exakt 1.0** auf. Dieser Idealwert resultiert aus einer **klaren Grenze** im Raum der Merkmale und einer mathematischen Trennbarkeit der vorliegenden Daten.

- **Kritische Reflexion:**

Eine absolut fehlerfreie Vorhersage tritt in der industriellen Realität äußerst **selten auf**. Ein solches Resultat erfordert zwingend **eine methodisch kritische Überprüfung** der Datenbasis auf mögliche Fehlerquellen.

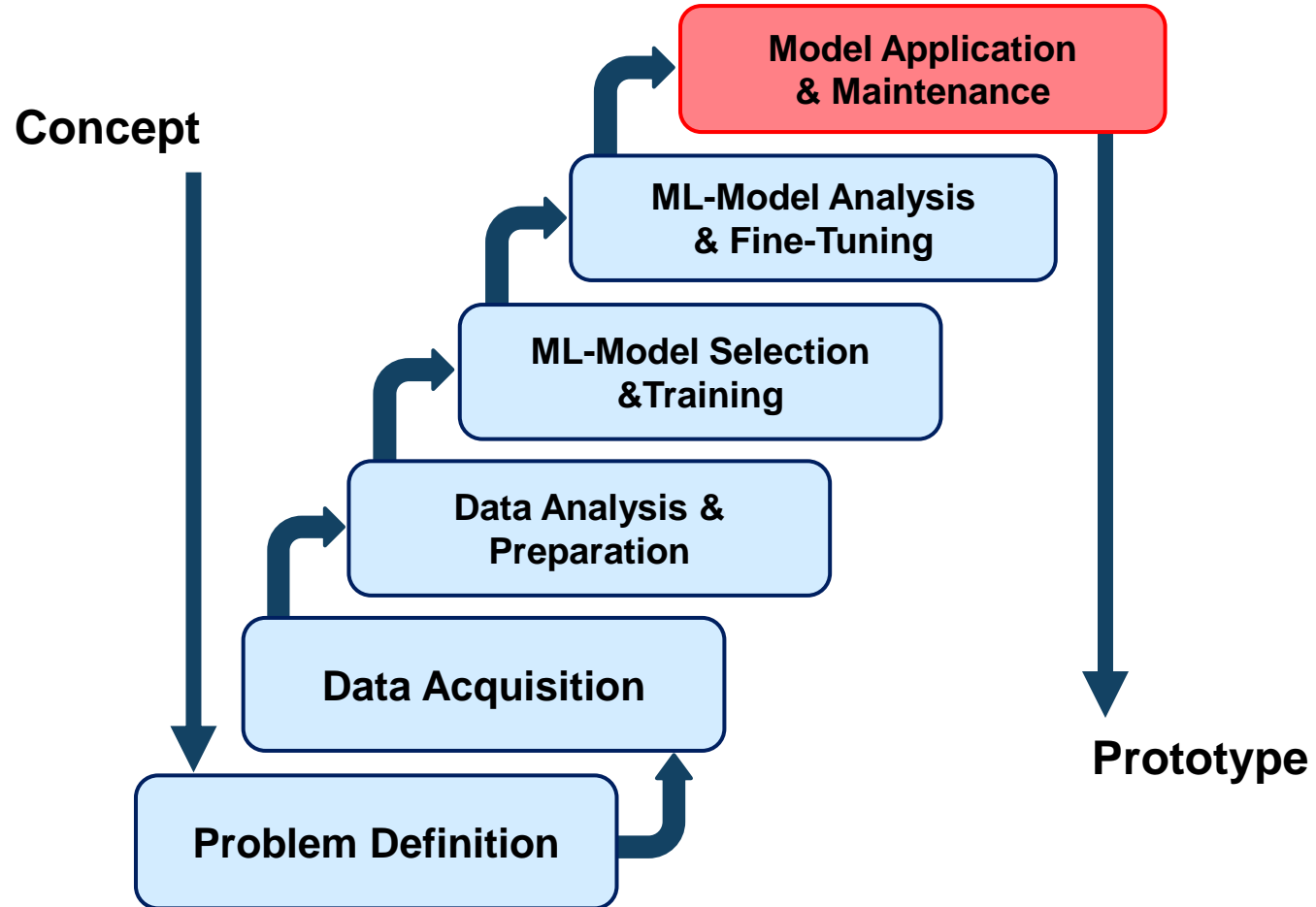
Festlegung des Schwellenwerts beim Isolation Forest

- **Definition des Schwellenwerts:**
Der Schwellenwert bestimmt, ab welcher berechneten Wahrscheinlichkeit ein Bauteil als **kritisch** eingestuft wird.
- **Niedriger Schwellenwert:**
Diese Einstellung maximiert die **Sensitivität**. Es werden nahezu alle potenziellen Problemchargen erfasst, was jedoch die Anzahl der **Fehlalarme** erhöht.
- **Hoher Schwellenwert:**
Diese Konfiguration fokussiert **die Spezifität**. Die Präzision der Warnungen steigt, wodurch unnötige Produktionsunterbrechungen minimiert werden.
- **Strategische Abwägung:**
Die Wahl des optimalen Schwellenwerts erfordert **eine Balance** zwischen maximaler Sicherheit und wirtschaftlicher Effizienz.



Workshop – Daten und KI in der Produktion

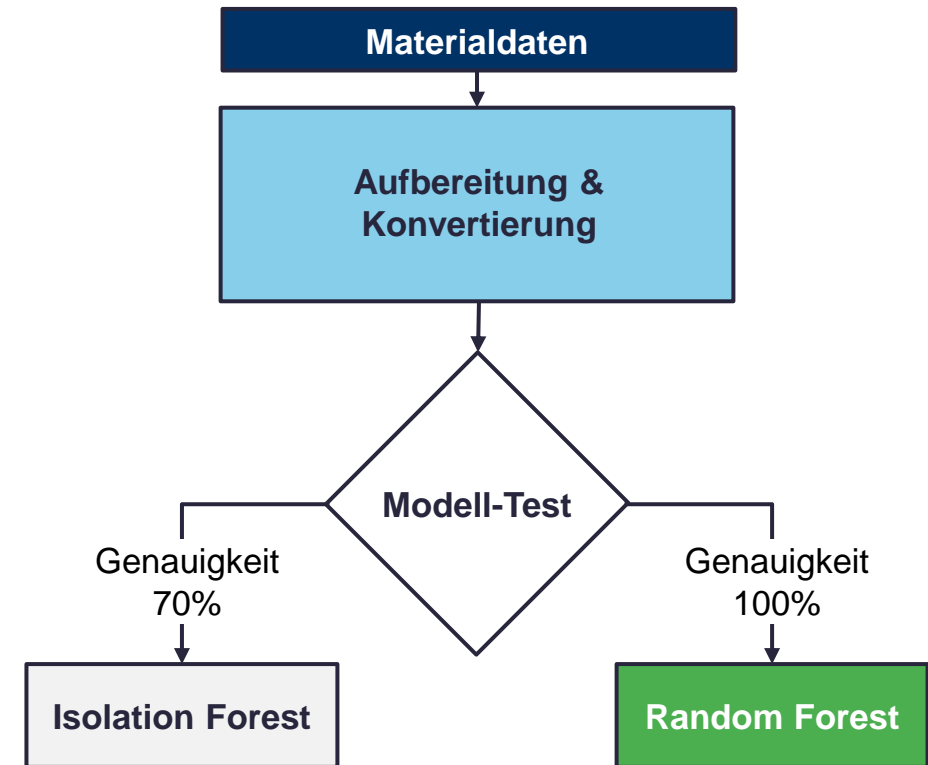
Model Application & Maintenance



Model Application & Maintenance

Zusammenfassung der Lösung

- **Datenbasis:**
ca. 100 Materialzeugnisse im PDF-Format mit mechanischen und chemischen Werten wie Streckgrenze und Kohlenstoffgehalt.
- **Vorbereitung:**
Einsatz einer Pipeline zur Datenkonvertierung der PDFs in ein ML geeignetes Format.
- **Modellvergleich:**
Evaluierung zwischen Random Forest (lernt überwacht) und Isolation Forest (lernt unüberwacht).
- **Entscheidung:**
Wahl des **Random Forest**, da die Klassen im Merkmalsraum gut voneinander trennbar sind.

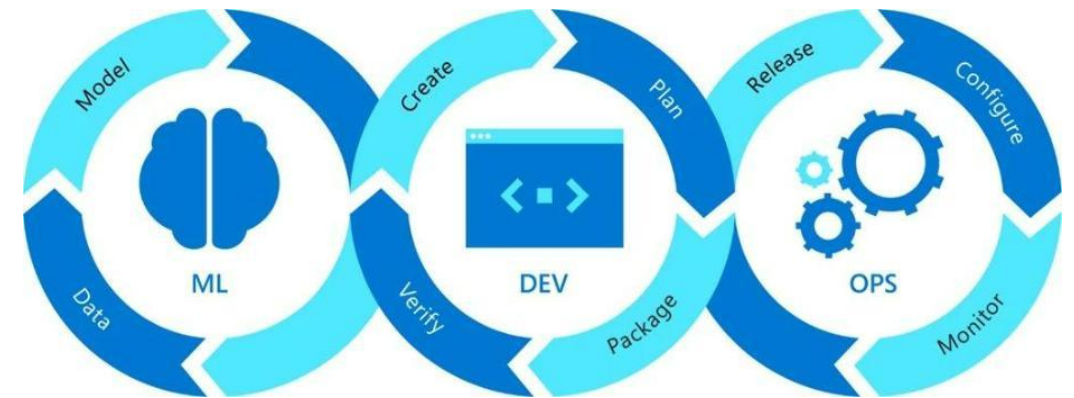


Der Prozess: Von den Daten zum KI-System

Modelleinsatz und Wartung

Fragestellungen, die es zu beantworten gilt:

- Wo muss das Modell integriert werden?
- Welche Schnittstellen existieren zu anderen Systemen?
- Echtzeit-Fähigkeit notwendig?
- kontinuierliches Monitoring der KI
- Feedbackmöglichkeiten umsetzen
- regelmäßige Aktualisierung des Modells durchführen
- weiteres Sammeln von Daten



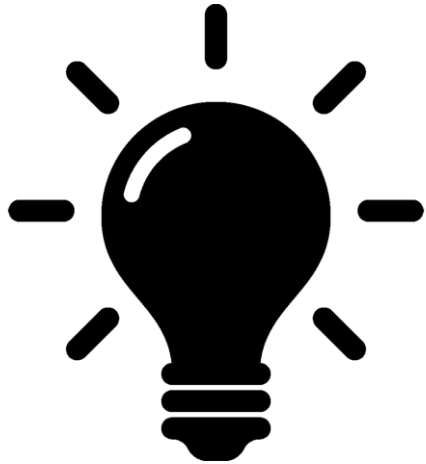
<https://azure.microsoft.com/de-de/blog/mlops-blog-series-part-1-the-art-of-testing-machine-learning-systems-using-mlops/>

PAUSE



Quelle: geralt auf pixabay.com

Reflektion



Quelle: geralt auf pixabay.com

Lessons Learned der Vollmann Group



Weitere Potenzialentfaltung bei Vollmann

- **Strategische Wissenserweiterung**
Gezielte Erweiterung des Wissens nach abgeschlossenem Screening möglich.
- **Erschließen weiterer Datenquellen**
Gezielte Suche und Bewertung von Dokumentendaten in weiteren Fertigungs- und Firmenprozessen
- **Skalierung der Datenextraktion**
Erkenntnisse aus entwickeltem Demonstrator verallgemeinern & interdisziplinären Extraktor entwickeln
- **Systemintegration und Chargen-Labels**
Integration des Datenextraktors in Enterprise Resource Planning Systeme sowie automatisierte Kennzeichnung von Material-Chargen nach Güte



Das Ziel ist es, durch Wissensausbau, optimierte Prozesse und automatisierte Kommunikation die Effizienz und Innovation in der Zusammenarbeit weiter zu steigern.

KI-Implementierung: Akademische vs. Industrielle Perspektive

KI-Implementierung

- **Akademisch:** Entwicklung von Algorithmen für präzise Voraussagen
- **Industriell:** Prognosequalitäten zur frühen Effizienzsteigerung nutzen

Theorie vs. Praxis

- **Akademisch:** Fokus auf Forschung, Validierung und Optimierung
- **Industriell:** Schnelle Implementierung und Fokus auf den ROI

Verbesserungsvorschläge

1. Wissenstransfer zwischen Forschung und Praxis stärken
2. Optimierung der Implementierung in definierter Zeit
3. Langfristiger Mehrwert durch nachhaltige Nutzung von KI-Ansätzen
4. Bildungsinitiativen für Mitarbeiter in KI und Datenwissenschaft



<https://www.proxi.id/images/blog/blog-academic-programs.png> &
https://img.rolandberger.com/content_assets/content_images/captions/Roland_Berger_INS_967_Industry_4.0_IT_image_caption_none.jpg

Eine erfolgreiche KI-Einführung in Unternehmen erfordert einen ausgewogenen Mix aus akademischer Präzision und industrieller Flexibilität, unterstützt durch fortlaufende Bildung und Kooperationen.

Strukturierte Entwicklung und Implementierung von KI-Lösungen

Vorgehensweise im Leitfaden:

Der im Workshop vorgestellte Prozess wird im Leitfaden detailliert beschrieben; von der Problemstellung bis zur fertigen Lösung.

Praxisbezug:

Ein realer Anwendungsfall (z. B. Bauteilklassifikation) sowie ausgewählte Code-Snippets (Python, TensorFlow) veranschaulichen jeden Entwicklungsschritt.

Datenqualität & Ethik:

Im Fokus stehen die Integration hochwertiger Daten sowie der Umgang mit Datenschutz, Fairness und Compliance in industriellen KI-Projekten.

Nachhaltige Weiterentwicklung:

Durch Modellüberwachung, Re-Training und Verfahren wie Random Forest wird der kontinuierliche Einsatz sichergestellt.



Der Leitfaden beschreibt einen strukturierten KI-Prozess – von der Problemdefinition über Modellierung bis zur praktischen Umsetzung und Wartung im Unternehmen.

Good Practices

Die Firma Hans Berg

Szenario

- Optimierung der Gütebewertung im Produktionsprozess zur Reduktion von Ausschuss

Herausforderung

- Falschklassifizierung durch maschinelle Güteprüfung trotz einwandfreier Bauteile
- Fehlende Integration metallographischer Prüfungen in automatisierte Bewertung

→ **Resultat:** Unnötig hoher Ausschuss trotz gut produzierter Teile

Aufgabenstellung

Optimierung des Ausschusses durch verbesserte Bewertung der Teilegüte

Lösung

- **KI-Modell:** Automatisierte Verknüpfung von Produktionsdaten mit Prüfergebnissen zur intelligenten Gütebewertung
- **Maschinelles Lernen:** Valide Zusammenhänge und Reduktion von Fehlklassifikationen
- **Ziel:** Minimierung von Ausschuss und Entlastung des Prüfprozesses



Ergebnis:

Signifikant gesteigerte Klassifizierungsgenauigkeit der Bauteile, wodurch der Ausschuss erheblich reduziert wurde.

Good Practices

Die Firma STRIKO

Szenario

- Optimierung der Entwicklung von Berstscheiben

Herausforderung

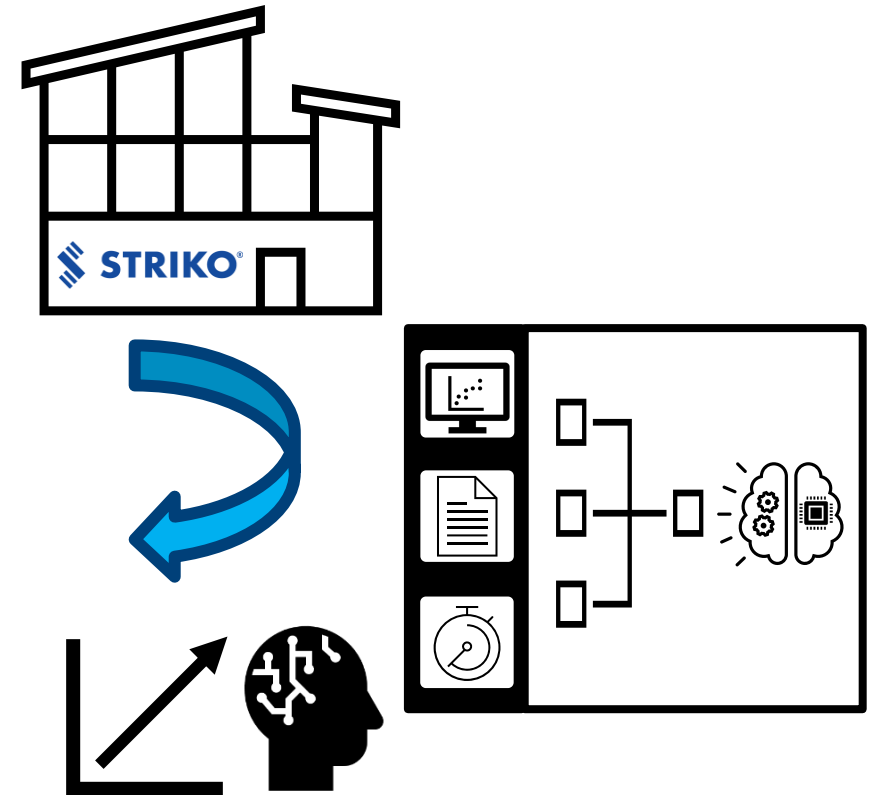
- Entwicklung überwiegend durch Expertenwissen und Trial-and-Error
 - Mangelnde systematische Datennutzung und Wissensarchivierung
- **Resultat:** Hoher zeitlicher Aufwand und inkonsistente Produktqualität

Aufgabenstellung

Systematische und datengetriebene Optimierung der Berstscheibenentwicklung

Lösung

- **KI-Modell:** Zusammenhang zwischen Konfiguration und Berstverhalten erkennen
- **Maschinelles Lernen:** Einsatz trainierbarer Algorithmen zur kontinuierlichen Modellverbesserung auf Basis vorhandener Testdaten
- **Ziel:** Generalisierbare Modelle zur Vereinfachung und Lösbarkeit komplexer Aufgaben



Ergebnis:

Deutliche Reduktion von Versuchs- und Testaufwand in der Entwicklung von Bertscheiben.

Beispiel-Firma

Die Firma Beulco

Szenario

- Automatisierung der Bauteilerkennung und -zählung im Produktionsprozess

Herausforderung bisheriger Lösung

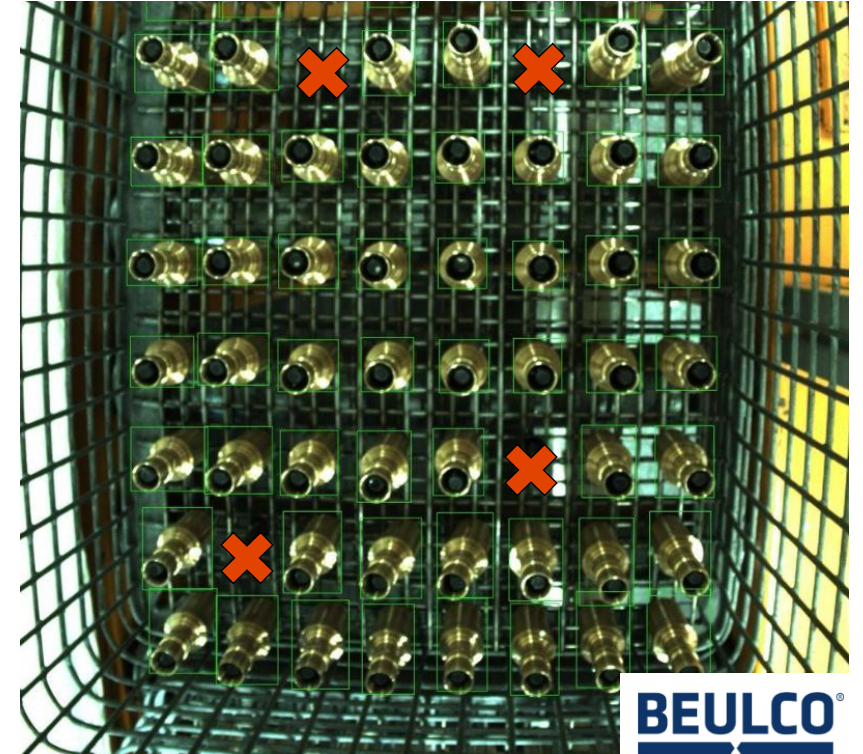
- Hoher manueller Aufwand
- Fehleranfälligkeit bei der Bauteilerfassung
- **Fazit:** präzisen und automatisierten Lösung

Aufgabenstellung

- **KI-gestützte Lösung** zur Erkennung und Zählung auf Basis von Bilddaten

Lösung

- **KI-Modell:** Einsatz einer Computer-Vision-Architektur zur automatisierten Erkennung
- **Maschinelles Lernen:** Kontinuierliche Verbesserung der Erkennungsgenauigkeit
- **Software:** Entwicklung einer intuitiven Anwendung für die Mitarbeiter
- **Ziel:** Prozesssicherheit verbessern, Erfassungsfehler reduzieren, Mitarbeiter entlasten



Ergebnis:

Effizienzsteigerung und Reduktion von Erfassungsfehlern durch KI-basierte Bauteilerkennung.

Unser Weg zur KI-Integration

- **Herausforderungen identifizieren:** Finden und definieren Sie KI-Anwendungen
- **Datenbasis aufbauen:** Nutzen Sie vorhandene Daten und erweitern Sie gezielt
- **Daten vorbereiten:** Fokus auf Pre-Processing → max. KI-Effizienz
- **Effizientes Projektmanagement:** Code-Aufwand definieren und Projekte zeitlich effizient umzusetzen
- **Verbesserungsschritte:** Meilensteine zur kontinuierlichen Evolution
- **Modulare Anwendung:** stufenweiser Systemaufbau für Flexibilität und Souveränität
- **Erfolg durch Transfer:** Adaption des Ansatzes für Ihre KI-Projekte
- **Neue innovative Pfade:** zusätzliche Chancen durch Kollaborationen
- **Übertragbarkeit:** prozessorientierter Ansatz für breite Anwendbarkeit in verschiedenen Unternehmensbereichen



https://sickconnect.com/wp-content/uploads/2023/06/shutterstock_2114019005-scaled.jpg

Profitieren Sie von den heutigen Erfahrungen für Ihren eigenen Erfolg: Integrieren Sie KI in Ihre Unternehmensprozesse.

Feedback für die zukünftige Ausrichtung des Arbeitskreises

- *Welche neuen Erfahrungen und Erkenntnisse haben Sie gewonnen?*
- *Hat das Konzept die Inhalte gut vermittelt?*
- *War die Menge an Informationen gut aufnehmbar?*
- *Gibt es Verbesserungsvorschläge?*
- *Welche neuen Themenideen haben Sie durch den heutigen Termin?*
- *Gibt es noch offene Fragen oder Gedanken zum Thema?*
- *Möchten Sie diese Thematik noch weiter vertiefen?*
- *Haben Sie ergänzende Themenvorschläge, die noch behandelt werden sollen?*
- *... ?*



Um die Inhalte möglichst gut zu vermitteln, sind wir auf Ihr Feedback angewiesen. Helfen Sie uns dabei den Lernprozess maximal ideal zu gestalten und lassen Sie ihre Ideen miteinfließen.

Vielen Dank für Ihre Teilnahme!

Wir freuen uns auf ein baldiges Wiedersehen.

Ihre Ansprechpartner



Martin Stock, B. Eng.

CAD CAM Center Cologne (4C)

Technische Hochschule Köln
Betzdorfer Straße 2
Telefon: +49 221 8275 4134
Mail: Martin.Stock@th-koeln.de



Martin de Fries, M. Sc.

CAD CAM Center Cologne (4C)

Technische Hochschule Köln
Betzdorfer Straße 2
Mobil: +49 1590 1956049
Mail: Martin.deFries@th-koeln.de



Prof. Dr. rer. nat. Dr. h.c. Margot Ruschitzka

CAD CAM Center Cologne (4C)

Technische Hochschule Köln
Betzdorfer Straße 2

Mail: Margot.Ruschitzka@th-koeln.de