

# ***Leitfäden***

---

*für das Verbundvorhaben **Effizienzschub in der Massivumformung durch Entwicklung und Integration digitaler Technologien im Engineering der gesamten Wertschöpfungskette (EMuDig 4.0)** im Förderprogramm: **PAiCE - Digitale Technologie für die Wirtschaft***

---

## 1 Inhaltsverzeichnis

1.	Einleitung.....	5
2	Rückverfolgbarkeit .....	6
2.1	Kleinstmögliche Losgrößenfertigung.....	6
2.1.1	Formalisierung des Prozesses .....	7
2.1.2	PLC-basierte Datenerfassung und -verarbeitung .....	8
2.1.3	Mehrdimensionale Modellierung und datengetriebene Rückverfolgbarkeit.....	8
2.1.4	Praktische Tags und ihr Leistungsvergleich.....	9
2.1.5	Experimentelles Assistenzsystem für die Einzel-/Chargenverfolgung.....	11
2.2	Prozessregelung .....	12
2.2.1	Modellfabrik am Institut für Umformtechnik.....	13
2.2.2	Adaptive Prozesssteuerungsstrategie .....	15
2.2.3	Cloud-Technologien für Datenspeicherung, Berechnung und Assistenzsystem.....	16
2.2.4	Anwendung des maschinellen Lernens als Modellierungsinstrument.....	17
3	Produktionsanlagen.....	19
3.1	Messkonzepte und Auswahl der Sensorik.....	19
3.2	Prognosemodell.....	21
3.2.1	Hydraulische Räderpresse, Otto Fuchs.....	21
3.2.2	Exzenterpresse, Hirschvogel.....	23
3.3	Definition Wartungsmaßnahmen.....	24
3.3.1	Hydraulische Räderpresse, Otto Fuchs.....	24
3.3.2	Exzenterpresse, Hirschvogel.....	27
3.4	Benutzerschnittstelle/ Visualisierung, Datenschnittstelle .....	30
4	Werkzeuge.....	31
4.1	Digitaler Werkzeugpass .....	32
4.2	Objektivierung Standmengenende.....	34
4.2.1	Verschleißmessung.....	35
4.2.3	Standmengenanalyse .....	36
4.2.4	Überwachungskonzept entwickeln .....	37
4.3	Standmengenprognose .....	38
4.3.1	Ermittlung der Dehnungswöhlerlinie .....	40
5	Kurzzeitanalyse.....	42
5.1	Steuerungsbasierte Datenerfassung und -verarbeitung.....	42
5.1.1	Zuordnung .....	42

# EMuDig4.0

5.1.2	Heterogenität .....	43
5.2	Konzept zur Datenmodellierung und Rückverfolgung .....	45
5.2.1	Extract-Transform-Load.....	45
5.2.2	Multidimensionale Datenmodelle.....	45
5.3	Prozessregelung/Mechanismen für die Umsetzung von präventiven Maßnahmen zur Einhaltung der Qualitätsanforderungen .....	48
5.3.1	Einführung .....	48
5.3.2	Ansatz .....	48
5.4	Anomalie-Charakterisierung.....	50
5.4.1	Entwickelte Netzarchitektur.....	50
5.5	Darstellung der Assistenzfunktionen und Integration in ein Assistenzsystem .....	52
5.5.1	Systemarchitektur .....	52
5.5.2	Datenmonitoring .....	53
5.5.3	Rückverfolgung.....	54
5.5.4	Analyse .....	56
5.5.5	Regelung der Qualität.....	58
6	Langzeitanalyse .....	60
6.1	Skalierbare Infrastruktur für die Datenanalyse .....	60
6.1.1	Cloud-Infrastruktur.....	60
6.2	Datenanalyse-Infrastruktur .....	68
6.2.1	Ingestion-Layer .....	68
6.2.2	Batch-Layer.....	69
6.2.3	Serving-Layer .....	69
6.2.4	Umsetzung in EMuDig 4.0 .....	71
6.3	Datenvorverarbeitung.....	72
6.3.1	Ausgangssituation .....	72
6.3.2	IT-Sicherheit.....	73
6.3.3	Datenintegration .....	75
6.3.4	Datenbereitstellung.....	76
6.3.5	Datendokumentation .....	77
6.3.6	Datenvalidierung .....	78
6.3.7	Datenbereinigung.....	78
6.3.8	Datenaggregation.....	78
6.4	Datenanalysemethoden .....	79
6.4.1	Datenvorverarbeitung.....	80
6.4.2	Ausarbeitung von Fragestellungen.....	80

## ***EMuDig4.0***

6.4.3	Statistische Beschreibung der Daten.....	81
6.4.4	Datenselektion und Dimensionsreduktion.....	82
6.4.5	Analysemethoden.....	83
6.4.6	Implementierung.....	88
6.4.7	Aufbereitung und Evaluation der Ergebnisse.....	88

# EMuDig4.0

## 1. Einleitung

Im Rahmen dieses Leitfadens werden die Ergebnisse der Hauptarbeitspakete zwei bis sechs des Projekts „EMuDig4.0“ im Kontext der Ergebnisse ausführlich dargestellt, die in der tatsächlichen Umsetzung bei den Anwendungspartnern erzielt wurden. Ziel dieser Darstellung ist es nicht nur, die Projektergebnisse insgesamt darzustellen, sondern speziell auch die Übertragbarkeit der definierten Modelle und Methoden aus der Modellfabrik in Pilotanwendungen und darüber hinaus zu bewerten.

Dafür folgt der Leitfaden der in Abbildung 1.1 dargestellten Struktur: Projektergebnisse der sechs übergeordneten Themen „Rückverfolgbarkeit“, „Anlagentechnologie“, „Werkzeuge“, Kurzzeit-Analyse“, „Langzeit-Analyse“ und „Factory-Cloud“ werden als eigenständige Module in den Kapiteln zwei bis sechs ausführlich behandelt.

In jedem Kapitel wird zunächst auf die Ausgangssituation in den Pilotanwendungen sowie die sich daraus ergebende Aufgabenstellung eingegangen. Anschließend folgt eine Darstellung der im Projekt verfolgten Methodik im Hinblick auf die eingesetzte Soft- und Hardware sowie die konkret umgesetzten Prozesse. Daraus leiten sich Grenzen der Modelle und Analysen sowie eine Bewertung der Funktionalität der einzelnen Systeme ab. Eine Stellungnahme zum Nutzen, der Verwertung der erarbeiteten Erkenntnisse sowie ihrer wirtschaftlichen Bedeutung rundet jedes Kapitel ab.

Für die ausführliche Darstellung der Ergebnisse aus Sicht der Pilotanwender Hirschvogel und Otto Fuchs sei auf den gemeinsamen Konsortial-Schlussbericht verwiesen, in dem die Projektergebnisse im Kapitel 4, in den Ausführungen zu den Arbeitspaketen 7.1 und 7.2 aus Sicht der Anwender ausführlich dargestellt werden.

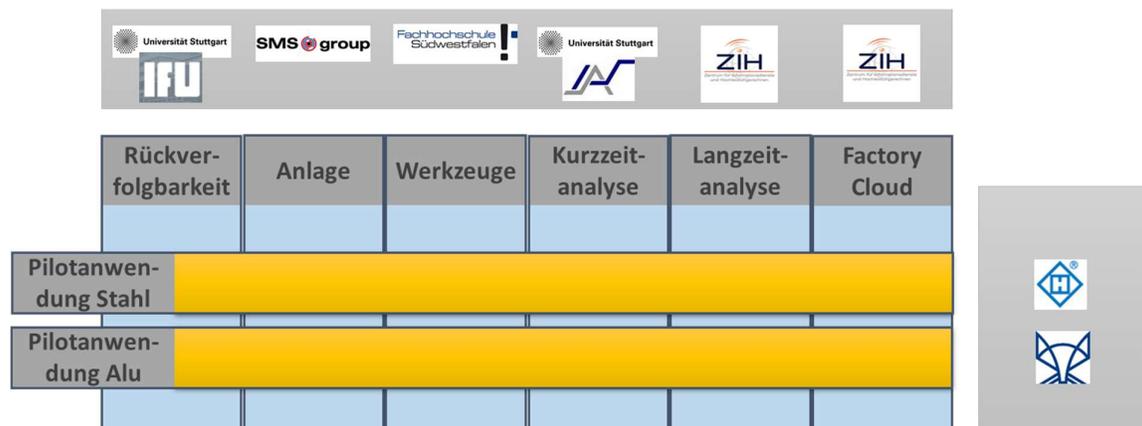


Abbildung 1.1: Übersicht der Leitfadenstruktur

## 2 Rückverfolgbarkeit

Die Verfolgung jedes einzelnen Werkstücks bietet zwei wesentliche Vorteile in der Schmiedetechnik. Erstens erleichtert die Zuordnung des physischen Werkstücks zu den überwachten Prozessinformationen die Ursachenanalyse für die Produktqualität. Zweitens können die nachfolgende Teilprozesse den eingehenden Werkstückeigenschaften angepasst werden, um die Robustheit der Prozesskette des Warmschmiedens zu verbessern. Im ersten Teil dieses Abschnitts wird eine allgemeine Methodik zur Rückverfolgbarkeit und Kennzeichnung von Aluminium- und Stahlschmiedeteilen für die Schmiedeindustrie vorgestellt. Darüber hinaus werden ein Rahmenkonzept für das Streaming und die Verarbeitung großer Mengen von Echtzeitdaten sowie ein mehrdimensionaler Ansatz zur Modellierung und Analyse der Werkstückinformationen für die Einzel- und Chargenverfolgung vorgestellt. Die Werkstückrückverfolgung überträgt die (Teil-)Prozess- und Zustandsmessungen in eine zusammenhängende Big-Data-Struktur, in der die Tools des maschinellen Lernens und der adaptiven Prozesskontrolle aufgebaut werden können. Der zweite Teil des Abschnitts erläutert im Wesentlichen die in der Modellfabrik entwickelte adaptive Steuerungsstrategie, die auf dieser Datenstruktur basiert. Die Echtzeit-Datenerfassung und das High Performance Computing auf einer Echtzeitdatenbank bieten nicht nur eine agile Lösung für die Prozesssteuerung, sondern stellen auch eine flexible webbasierte Plattform für die Visualisierung in jedem Anwendungsbereich dar. Ein vor Ort arbeitendes Assistenzsystem kann die notwendigen Informationen visualisieren und dem Maschinenbediener die Arbeit bei der Entscheidungsfindung erleichtern. Die Prozesshistorie oder -Übersicht kann auf jedem mit dem Internet verbundenen Gerät beliebig dargestellt werden.

### 2.1 Kleinstmögliche Losgrößenfertigung

Die Produktionslinie eines Schmiedewerks beinhaltet je nach Produkt und verfügbarer Infrastruktur sehr unterschiedliche Prozessschritte. Dieser Abschnitt stellt Bestandteile einer verallgemeinerten Methodik vor, die auf möglichst diverse Produktionssysteme und Prozessketten adaptiert werden kann. Werkstücke können entweder auf Chargenebene oder einzeln verfolgt werden. Die Rauheit des Prozess-Schritts, die Haltbarkeit der Tags und die Markierungsgeschwindigkeit bestimmen die Intensität der Verfolgung. Eine komplette Produktionslinie kann Teilprozesse wie Gießen, Extrudieren, Sägen, Drehen, Transfer, Lagern, Erwärmen, Schmieden, Strahlen und Wärmebehandlung enthalten. Die (Teil-)Prozessinformationen können in der Factory-Cloud gespeichert und über eine Master-Identität einem bestimmten Werkstück zugeordnet werden. Es ist nicht notwendig, die komplette Master-Identität auf einem am Werkstück befestigten Tag zu speichern. Lokale oder temporäre Identitäten können auf Maschinen- oder Teilprozessebene zugewiesen und wiederverwendet werden, vorausgesetzt, die entsprechenden Master-Identitäten sind eindeutig. Die Verfolgung eines Werkstücks über eine Master-Identität mit verschiedenen lokalen Identitäten in der gesamten Fertigungslinie erfordert eine Vernetzung der technischen Verfolgungssysteme in Echtzeit. Die manuelle Identifikation, Übertragung und Registrierung von Identitäten kann ein Teil dieses Netzwerkes sein. Mit einer derart gelungenen Vernetzung von Tracking-Systemen reduziert sich das Tracking-Problem auf eine lokalisierte Verfolgung von Werkstücken über einzelne Prozesse hinweg.

## 2.1.1 Formalisierung des Prozesses

Teilprozesse in einer Schmiedeproduktionslinie können in die folgenden Prozesskategorien (PK) eingeteilt werden:

- PK1 - Der Tag bleibt während des gesamten Prozesses bestehen.
- Tag besteht nicht während des gesamten Prozesses und
  - PK2 - Werkstücke werden der Reihe nach bearbeitet
  - PK3 - wird eine Batch-Verarbeitung durchgeführt.

(Teil-)prozesse, die zu PK1 gehören, beschädigen das auf dem Werkstück angebrachte Tag nicht, so dass das Tag nach der Verarbeitung einwandfrei identifiziert werden kann. Prozesse, die zu PK2 gehören, beschädigen das Tag. Die sequenzielle Verarbeitung der Werkstücke ermöglicht jedoch die Verfolgung der Identität innerhalb und nach der Verarbeitung in Bezug auf die vor dem Prozess gescannte Identität. Die Qualität der Teile kann schon während oder mindestens direkt nach dem Teilprozess überprüft werden. Genügen die Teile den Qualitätskriterien nicht, muss der Ausschuss in Echtzeit registriert und aussortiert werden. Dieser Vorgang kann automatisiert oder durch manuelle Auslösung erfolgen. Die zu PK3 gehörenden Teilprozesse beschädigen das Tag ebenfalls. Die individuelle Identität des gefertigten Teils geht hierbei verloren, da viele Werkstücke gemeinsam bearbeitet werden. In diesem Fall ist eine individuelle Neukennzeichnung aufgrund der hohen Produktionsrate oder logistischer Zwänge nicht praktikabel. Solche Prozesse (z.B. Wärmebehandlung, Strahlen) erlauben daher nur eine Chargenverfolgung. Einzelne Informationen von gemeinsam bearbeiteten Werkstücken können in einer "statistischen Darstellung" zusammengeführt, in der Factory-Cloud gespeichert und nur mit Chargen verknüpft werden.

Die Markierungszeit, die bei PK2 länger als die Prozesszeit dauert, behindert die Verfolgung aller Werkstücke einzeln, es sei denn, es werden mehrere Markierungsstationen eingesetzt. Wertvolle zusammenhängende Informationen über Prozesse und Werkstücke können jedoch für einige der Werkstücke mit einer einzigen Markierungsstation erstellt werden.

Die Chargenverfolgung kann durch Online-Messungen angereichert werden, um den Informationsgehalt der Rückverfolgung vor Ort zu verbessern. Fortan kann die Rückverfolgung mit zusätzlichen individuellen Informationen im weiteren Verlauf der Prozesskette durchgeführt werden. Zum Beispiel ist die Werkstückgröße, die in der Regel eine hohe Streuung aufweisen kann, ein wichtiger Faktor für die Prozesssteuerung und für die Analyse der Zusammenhänge. Zu diesem Zweck kann eine Online-Geometriemessung durchgeführt werden. Andere Messsysteme (z.B. ein Wirbelstrom-Messsystem) könnten auch verwendet werden, um andere Eigenschaften der Rohlinge oder des Produkts zu messen.

Die operativen Rückverfolgungsaktionen sind die Markierung mit der Identität und das Scannen von Tags, um die Identität und die entsprechenden Informationen aus der Factory-Cloud abzurufen. Die rechnerischen Rückverfolgungsaktionen sind: Lokalisierung der Identität im Fabriknetzwerk, Bereitstellung der erforderlichen Informationen aus der Factory-Cloud an die Prozessmaschine und Speicherung der erzeugten Produkt- und Produktionsinformationen in der Factory-Cloud. Die Lokalisierung der Identität im Fabriknetzwerk erfolgt entweder durch Scannen der Tags vor Ort oder "virtuell" ohne Hilfe eines physischen Tags am Werkstück. In Form eines „virtuellen Tags“. Wenn ein strenges Übergabeprotokoll mit sequentieller Verarbeitung für eine Reihe von Prozessen implementiert wird, kann eine virtuelle Standortverfolgung mit Hilfe von Näherungssensoren durchgeführt werden. Diese Standortverfolgung kann auch durch digitale Überwachung mit Bildverarbeitung realisiert werden. In Abwesenheit eines strikten Übertragungsprotokolls müssen

# EMuDig4.0

Standortvermittlungen durch Bewegungsverfolgung unterstützt werden, um die virtuelle Verfolgung durchzuführen. Die Bewegungsverfolgung könnte mit einem Netzwerk von Videokameras und Bildverarbeitungstechniken realisiert werden. Kamerastandorte und Bewegungspfade müssen so geplant werden, das eine robuste Konfiguration entsteht.

Die Kategorie eines Teilprozesses kann durch eine Prozessneugestaltung geändert werden, um die Rückverfolgbarkeit zu verbessern. Werkstückförderer oder Werkstückträger mit strikter Einhaltung der Bearbeitungsreihenfolge anstelle einer ungeordneten Übergabe sind triviale Beispiele für solche Verbesserungen. Zumindest können kleine Werkstückbehälter mit reduzierten Losgrößen verwendet werden, um die Qualität der statistischen Darstellung der zugehörigen Informationen zu erhöhen.

## 2.1.2 PLC-basierte Datenerfassung und -verarbeitung

Die individuellen Schnittstellen der Konnektor-Software laufen auf Steuergeräten und die standardisierte Schnittstelle läuft in der Factory-Cloud. Die individuelle Schnittstelle ist grundsätzlich als Zustandsmaschine implementiert, die in den SPS-Code integriert werden kann. Es lassen sich zwei wesentliche Zyklen unterscheiden. Zum einen ein "Lese"-Zyklus, der die Protokollstruktur abtastet und je nach Bussystem Daten extrahiert. Zweitens, ein "Schreib"-Zyklus, der ein REST-Modell verwendet, um die extrahierten Daten zu beschreiben. Die gemessenen Prozessparameter werden konvertiert, auf das Modell abgebildet und an die standardisierte Schnittstelle auf dem Cloud-Server gesendet. Die standardisierte Schnittstelle leitet die Daten zur nachfolgenden Verarbeitung weiter. Sie werden entweder in einer relationalen Datenbank gespeichert, in der Datenanalysetechniken angewendet werden können, oder an eine Echtzeit-Datenbank weitergeleitet, um Ad-hoc-Berechnungen und analytische Online-Verarbeitung (OLAP) durchzuführen. Die Echtzeitdatenbank wird für die Zuordnung von Werkstückidentität und Prozessparametern entlang der gesamten Prozesskette verwendet. Sie ermöglicht eine Online-Verfolgung von Einzelteilen oder Chargen. Das allgemeine Konzept ist in Abbildung 2.1 dargestellt.



Abbildung 2.1 Konzept und Systemarchitektur

## 2.1.3 Mehrdimensionale Modellierung und datengetriebene Rückverfolgbarkeit

Um eine effiziente Überwachung und Verfolgung zu erreichen, muss ein OLAP-Modell unter Berücksichtigung der Datenräume der einzelnen Prozessschritte erstellt werden. Das Modell ist wie in Abbildung 2.2 dargestellt in die Dimensionen Zeit, Raum und Produktinformation unterteilt. Die Dimensionen enthalten eine interne, hierarchische Struktur, in der jede Verzweigung durch ein

# EMuDig4.0

Element beschrieben wird. Grundelemente sind mit Datenvektoren verbunden und markieren die innerste Schicht des Datenmodells (Produkttyp P mit id x). Höhere Schichten (Charge A des Produkttyps P) aggregieren die Informationen der unteren Schichten durch Anwendung von Aggregationsverfahren. So können heterogene Daten in ein homogenes Modell integriert werden, da es möglich ist, verschiedene Schichten mit unterschiedlicher Auflösung zu assoziieren. Dieser Faktor ermöglicht die Verfolgung einer flexiblen Anzahl von Teilen durch die Navigation durch die Schichten. Der Modellierungsansatz ist in Abbildung 2.2 als Würfel dargestellt. Für n-dimensionale Modellstrukturen sind drei Haupteigenschaften erforderlich: 'Gemeinsamkeit', 'Wiederverwendbarkeit' und 'Zusammensetzbarkeit'. Der Aspekt der 'Kombinierbarkeit' wird durch einen Metamodellierungsansatz ergänzt. So können Schwellenwerte und Regeln für Prozessparameter definiert werden, die eine effektive Grenzwertüberwachung ermöglichen. Im Hinblick auf die beschriebene Prozesskette wirken sich unterschiedliche Umweltbedingungen tatsächlich auf die Qualität von Prozess und Produkt aus. Beispielsweise hängt die abnehmende Qualität von Kühlmitteln von zahlreichen Faktoren ab, wie z.B. der Lebensdauer (Zeitabmessung) und der Maschine (Raumabmessung). Folglich können diese Nebenwirkungen formalisiert und auf die Dimensionsstruktur abgebildet und auf den Meta-Würfel projiziert werden. Somit kann ein automatisiertes Monitoring auf der Basis von Expertenwissen realisiert werden.

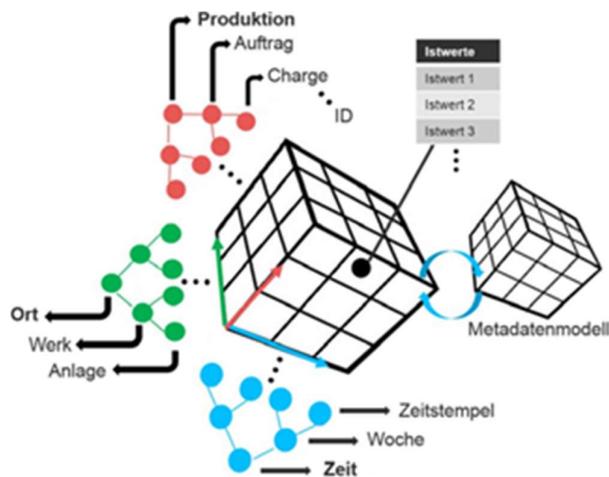


Abbildung .2.2: Mehrdimensionales Modell von Prozess- und Werkstückinformationen.

## 2.1.4 Praktische Tags und ihr Leistungsvergleich

Der Aufbau eines kompletten Trackingsystems wird unter Berücksichtigung der angewandten Teilprozesse, der Teilegeometrie und des Materials konzipiert. Die Diskussion beschränkt sich hier auf den Vergleich der bekannten Tag-Eigenschaften und die Untersuchungen der noch nicht getesteten Eigenschaften. In diesem Zusammenhang werden Lasermarkierungen an heißem Stahl und Aluminium durchgeführt und die Ergebnisse mit der Nadelmarkierung verglichen. Lösungsansätze zur Teile- oder Chargenverfolgung für Teilprozesse wie Zerspanung, Sägen, Transfer oder Lagerung befinden sich bereits in der industriellen Praxis. Die raue Schmiedenumgebung, die möglicherweise Erwärmung, Schmieden, Strahlen und Wärmebehandlung umfasst, beschränkt die einsetzbaren Kennzeichnungstechniken auf RFID, Etikettierung, Nadelmarkierung, Lasergravur und virtuelle Kennzeichnung. Tags dürfen keinen Schaden verursachen, sie müssen bestehen und könnten bei etwa

# EMuDig4.0

500°C für Aluminium und 1200°C für Stahl gekennzeichnet werden. Die Entwicklung der Oxidschicht muss bei der Auswahl der Markierungstechnologie berücksichtigt werden. Ein hitzebeständiger RFID ist bis zu 350 °C am Transponder beständig, wobei eine Größe von 10×7×3 cm auf Behältern zur Chargenverfolgung praktisch ist. Es gibt auch Etiketten, die bis zu 1250 °C haltbar sind und eine flexible Größe und einen flexiblen Inhalt für die Einzel- oder Chargenverfolgung bieten. Die Schwierigkeit entsteht beim Anbringen, Aufbewahren und Entfernen der Etiketten. Obwohl sie nur oberflächlich sind, werden Nadelmarkierung und Lasergravur zu einem Teil des Werkstücks. Sie verändern lokal die Geometrie und sogar die Materialeigenschaften. Tabelle 2.1 vergleicht die Funktionalität der Nadelmarkierung, Ritzmarkierung und Lasergravur auf der Grundlage der Marktforschung und der durchgeführten Experimente.

Tabelle 2.1: Vergleich der Funktionalität von kommerziellen Tags

	Ritzmarkierung	Nadelmarkierung	Faser (Fiber Laser)-Gravur (50W)
<b>Inhalt</b>	Klartext	QR-Code, Klartext	
<b>Markierungstiefe</b>	Begrenzt	Schlagenergie	Strahlungsenergie und wiederholte Schüsse
<b>Markierungszeit</b>	~ 1 Sek.	~ 4 Sek.	~ 8 Sek. for 50 W Faser (erhöht sich mit der Größe und Oberflächenvorbereitung, reduziert sich mit der Leistung)
<b>Klarheit</b>	Fehlender Kontrast	Fehlender Kontrast	besser durch Oberflächenvorbereitung
<b>Größe</b>	Markierungsgröße	Nadelgröße	Flexibel
<b>Auflösung</b>	-	Fest	Flexibel innerhalb eines Kennzeichens
<b>Oberflächenkrümmung</b>	nur leichte Krümmungen erlaubt		Flexibel mit adaptiver Fokussierung
<b>Klartext</b>	Hohe Qualität	Gepunktet	Gepunktet mit flexibler Auflösung
<b>Temperatur</b>	Getestet bei 500 °C für Aluminium, 1200 °C für Stahl		
<b>Kosten</b>	~ 10.000 €		~ 30.000 €
<b>Erkennung</b>	Optische Erkennung, OCR	QR-Code Lesegerät, OCR	QR-Code Lesegerät, OCR (Überlegen mit Markierung auf der Oberflächenvorbereitung.)
<b>Strahlungswiderstand</b>	Im Allgemeinen mangelhaft. Mit der Punktgröße und -tiefe nehmen sie zu.		

QR-Code kann eine erfolgreiche Erkennung auch dann gewährleisten, wenn ein Teil des Etiketts beschädigt ist. Zwei Beispiele finden sich in Abbildung 2.3. Links ein QR-Code, der von einem Faserlaser auf Aluminium bei 500°C erzeugt und anschließend durch Kollisionen mit anderen Teilen während des Schüttelns in einem Behälter beschädigt wurde, wobei die Erkennung jedoch funktionsfähig ist. Rechts ein kaltmarkierter Stahl, der nach drei Stunden Bearbeitung bei 900°C mit einer Oxidschicht bedeckt ist. Sichtbare Markierungen auf der Oxidschicht und auf der sauberen Oberfläche nach Entfernung der Oxidschicht werden beide durch QR-Code-Lesegeräte erkannt. Sollte das QR-Code-Lesegerät ausfallen, ist der beigefügte Klartext gut lesbar.



Abbildung .2.3: QR-Codes, die mit einem Faserlaser erzeugt werden: Aluminium bei 550°C, durch Kollisionen beschädigt (links); Stahl bei RT, bedeckt von einer Oxidschicht, die sich bei 900°C drei Stunden lang entwickelt (rechts). Alle Codes werden erfolgreich erkannt.

Sowohl die Nadelmarkierung als auch die Lasergravur können zur Markierung von QR-Codes und klaren Texten auf der Oberfläche verwendet werden. Da die Auflösung der Nadelmarkierung durch den Nadeldurchmesser begrenzt ist, ist die Beschriftungsgröße nicht flexibel. Die Tiefe der Punkte kann durch die Aufprallenergie beeinflusst werden. Die Überlappung von Punkten durch Vernadelung ist nicht praktisch, aber für einen Lasergravierer ist dies völlig unproblematisch und kann durchgeführt werden, um die Tiefe jedes Punktes und sogar die Auflösung der Gravur zu erhöhen. Ein QR-Code zusammen mit dem Inhalt im Klartext ist durch Nadelmarkierung im Vergleich zur Lasergravur schwer zu erzeugen. Abbildung 2.4 vergleicht eine hochauflösende Lasergravur (bei der ein einzelner Punkt mit vielen winzigen Punkten erzeugt wird) mit dem Ergebnis der Nadelmarkierung, um die Grenzen der Nadelmarkierung und die Möglichkeiten der Lasergravur deutlich zu zeigen. Eine Lasergravur mit Einzelpunktauflösung ist ebenfalls vorgesehen, um den Unterschied zwischen den Punkteigenschaften zu demonstrieren. Die größeren und tieferen Dots bieten eine bessere Beständigkeit gegen Strahlreinigung.

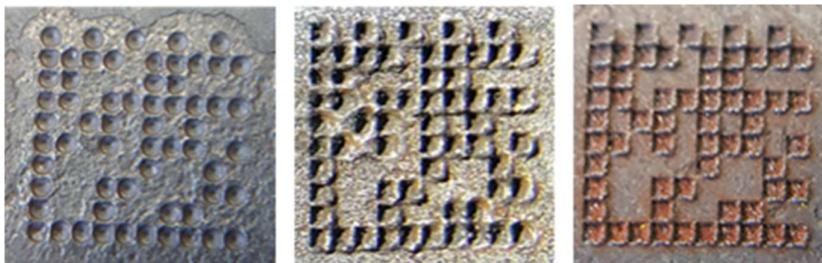


Abbildung 2.4: Durch Nadelmarkierung erzeugte QR-Codes (links), Lasergravur mit niedriger Auflösung (Mitte) und Lasergravur mit hoher Auflösung (rechts), alle mit erfolgreicher optischer Erkennung. QR-Code Größe: etwa 20 mm

## 2.1.5 Experimentelles Assistenzsystem für die Einzel-/Chargenverfolgung

Um die Systemarchitektur zu testen, wird eine Beispiel-Prozesskette verwendet. Die Steuerungen der Umgebung laufen mit einer Taktzeit von 1 ms, was den Verlust relevanter Prozessinformationen aufgrund niedriger Abtastraten verhindert. Der Datenraum umfasst 18 verschiedene Sensorwerte (reale und simulierte) von 5 aufeinander folgenden Prozessschritten. Die standardisierte Schnittstelle

# EMuDig4.0

läuft in der Cloud und leitet die Daten im JSON-Format an die OLAP-Datenbank weiter. Die Übertragung der Daten kann auf verschiedene Weise gesteuert werden. Zum einen kann eine ständig frequentierte Abfrage durch den REST-Server in der Cloud ausgeführt werden. Zum anderen kann über die kundenspezifischen Schnittstellen eine ereignisgesteuerte Datenübertragung realisiert werden.

Bei unterschiedlichen Mengen von Teilen, die verfolgt werden, bildet der ETL-Stapel die neuen Informationen auf den hierarchischen Dimensionsstrukturen ab. Diese Projektion bewirkt die Bildung höherer Schichten für die Chargenverfolgung. Als Ergebnis werden neue Elemente innerhalb der Hierarchie erzeugt. Dadurch können unterschiedliche Prozessdatenstrukturen von unterschiedlichen Produkten und Mengen angepasst werden. Die gleiche Flexibilität gilt für jedes andere Element einer beliebigen anderen Schicht in der hierarchischen Dimensionsstruktur. Somit ist es möglich, unterschiedliche Mengen von Teilen durch die Navigation in der hierarchischen Dimensionsstruktur zu verfolgen. Chargen können mit den entsprechenden Prozessdaten und kumulierten Werten der Messdaten verknüpft werden. Im Gegensatz zu Einzelwerkstücken erfordert die Verfolgung mehrerer Teile die sichere Abschätzung, ob alle Teile die Qualitätsanforderungen erfüllen. Die Navigation zur Chargenebene überwacht automatisch alle relevanten Parameter, die die Charge charakterisieren und gibt dem Anwender einen Überblick über die Durchschnittswerte. Diese Werte werden auf der Basis aller Teile der Charge berechnet. Der gesamte Datenintegrationsprozess wird zyklisch mit paralleler Berechnung ausgeführt, was eine Online-Überwachung und -Verfolgung ermöglicht. Es ist als Web-Schnittstelle für einen flexiblen Zugriff von einer Vielzahl von Geräten aus implementiert. Die erste Ansicht ermöglicht die Grenzwertüberwachung von Prozessparametern, so dass die Überschreitung von Grenzwerten beleuchtet wird. Der Benutzer navigiert in allen drei Hauptdimensionen Zeit, Raum und Produkt. So kann er einzelne Werkstücke oder Chargen in einem bestimmten Prozessschritt zu einem Zeitpunkt seiner Wahl auswählen und die entsprechenden Prozessdaten werden angezeigt. In der zweiten Ansicht sind Trajektorien und Abweichungen zu beobachten, was eine effiziente Kontrolle der Prozessparameter ermöglicht.

## 2.2 Prozessregelung

Die Regelung von Produkteigenschaften in der Metallumformung ist ein aktives Forschungsthema. Aufgrund der kurzen Dauer eines Umformprozesses und der rauen Umgebung ist es schwierig, eine Inline-Prozesskontrolle beim Schmieden mit direkten, korrigierenden Eingriffen während eines Schmiedevorgangs durchzuführen. Stattdessen können Inline-Messungen verwendet werden, um (Teil-)Prozessparameter für zukünftige Werkstücke auf der Grundlage der bereits vorher beobachteten Trends zu schätzen. Dieser Ansatz ist in der Lage, mit den für die Fertigungsanlage und -umgebung spezifischen Schwankungen zurechtzukommen. Material- und werkstückspezifische Abweichungen können bei diesem Ansatz nicht berücksichtigt werden, da Werkstücke mit wahrscheinlich unterschiedlichen Eigenschaften keine eindeutige Identität aufweisen. Ohne Werkstückverfolgung liefert die Prozessdigitalisierung nur lokale Informationen, die spezifisch für jeden Teilprozess und vom gesamten Prozessablauf entkoppelt sind. Daher ist die Zuordnung von Material-, Prozessschritt- und Qualitätsparametern zu den einzelnen Werkstückidentitäten notwendig. Diese Zuordnung bildet den großen Datensatz und das technologische Wissen, das für die Ursachenanalyse notwendig ist. Das gesammelte Wissen in Bezug auf die datenbasierte Modellierung und die auf maschinellem Lernen basierende Modellierung bilden eine flexible adaptive Regelstrategie. Um die erforderliche Menge an realen Prozessdaten zu reduzieren, kann die Finite-Elemente-Analyse (FEA) für das erste Training des maschinellen Lernmodells verwendet werden. Das künstliche neuronale Netz (KNN) wird dann mit

# EMuDig4.0

einer relativ geringen Anzahl von experimentellen Daten weiter verbessert. Die Art der maschinellen Lernarchitektur erlaubt es, weitere Trainingsberechnungen auf der Basis einzelner Datenpunkte oder durch Massenverarbeitung mehrerer Datensätze auf einmal durchzuführen. Sofern die für die Modellierung, das Modelltraining und die Steuerung erforderliche Rechenleistung die Kapazität eines typischen Computers übersteigt, könnte das Cloud-Computing eine alternative Lösung sein.

## 2.2.1 Modellfabrik am Institut für Umformtechnik

Die Modellfabrik im Rahmen vom EMuDig4.0 wurde von den Konsortialpartnern IFU und IAS der Universität Stuttgart aufgebaut. Die Abbildungen 2.5 und 2.6 veranschaulichen die physikalischen Komponenten und das Datenflussnetz der Modellfabrik, die sich aus der folgenden Prozesskette zusammensetzt:

- Bolzenvorbereitung (AL6082-Knüppel mit den Nennmaßen 50 mm × 50 mm Höhe, durch Sägen oder Drehen),
- Kennzeichnung der Knüppel-Identität in einem QR-Code auf jedem Knüppel mit Lasermarkierungstechnologie, die die physische Werkstückverfolgung bis zur Verarbeitung des Knüppels in der induktiven Erwärmungseinheit emuliert,
- induktive Erwärmung, die adaptiv auf der Grundlage der Bolzeigenschaften und der tatsächlichen Werkzeugtemperaturen gesteuert wird,
- Werkzeugschmierung mit kontrolliertem Sprühen auf der Grundlage der Bolzeigenschaften und der aktuellen Werkzeugtemperaturen,
- zweistufiges Warmschmieden von Aluminiumknüppeln, die jeweils adaptiv auf der Grundlage der Bolzeigenschaften, der tatsächlichen Werkzeugtemperaturen und der gemessenen Prozessparameter aus den vorherigen Schritten gesteuert werden,
- Messung der Werkstücktemperatur und der kritischen Abmessungen nach jeder Umformstufe zur Inline-Prozessqualitätskontrolle,
- virtuelle Verfolgung des Werkstücks durch induktive Erwärmung, Schmiede- und Messvorgänge, einschließlich der Echtzeit-Verfolgung von Schrott.

Die Schmierung der Werkzeuge und die Übertragung der Werkstücke von der induktiven Erwärmung bis zur geordneten Ablage auf einem Tisch wird von einem Industrieroboter durchgeführt.

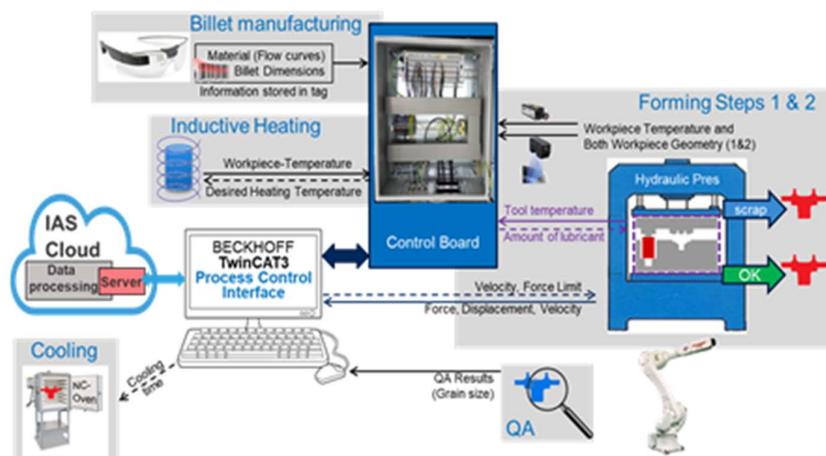


Abbildung 2.5: Modellfabrik – Datenfluss-Netzwerk

## EMuDig4.0

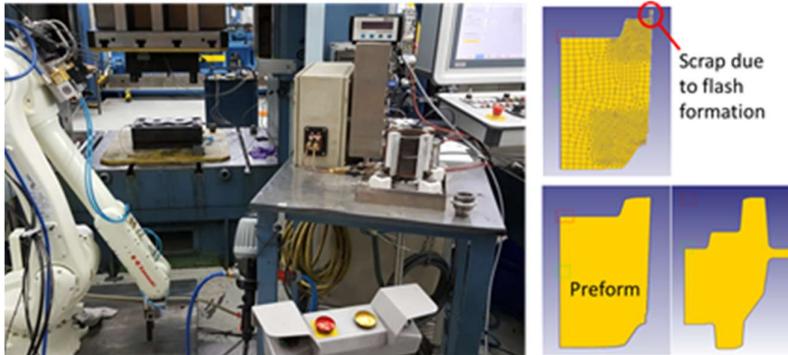


Abbildung 2.6: Modellfabrik und die achsensymmetrischen FEA-Ergebnisse für akzeptable Teile und Ausschuss

Die Steuerung der Presse und der Peripheriegeräte (induktive Erwärmung, Schmieresystem, Geometrie- und Temperaturmessungen) auf der Fertigungsebene wird auf einem Industrie-PC realisiert, der sich im Steuerschrank der hydraulischen Presse befindet. Die Presse kommuniziert direkt mit dem Industrie-PC, während die Peripheriegeräte ihre Schnittstelle zum Industrie-PC über einen zusätzlichen Schaltschrank haben. Der Industrie-PC kommuniziert mit einem Cloud-Server über das Internet: Alle Mess- und Prozessdaten werden an diesen Server gesendet, dort werden adaptive Steuerbefehle berechnet und schließlich ebenfalls über das Internet zurückgeschickt.

Die Vorbereitung des Bolzens durch Sägen oder Drehen wird durch Messen der Bolzenabmessungen und Markieren einer Identitätsnummer (ID) abgeschlossen, die mit den Maß- und Werkstoffinformationen auf dem Bolzen durch einen QR-Code mittels Lasermarkierungstechnologie verbunden ist. In Kombination mit der sequentiellen Bearbeitung und der virtuellen Verfolgung des Werkstücks wird gezeigt, dass die Laserbeschriftung in einer rauen und heißen Schmiedenumgebung, insbesondere für Aluminium, von der induktiven Erwärmung bis zum Schmieden und sogar bis zum anschließenden Sandstrahlen funktioniert. Die Laserbeschriftung mit der bis zu vier Ziffern umfassenden numerischen Identität kann bei Bedarf innerhalb von 3-5 Sekunden aktualisiert werden, je nachdem, ob der QR-Code mit dem entsprechenden Klartext ergänzt wird oder nicht.

Der Herstellungsprozess beginnt danach mit der induktiven Erwärmung des Bolzens. Die induktive Erwärmung des nächsten Bolzens wird während des Schmiedens des vorherigen Werkstücks durchgeführt. Die Modellfabrik demonstriert in diesem Sinne die Praktikabilität der virtuellen Werkstückverfolgungsstrategie bei gleichzeitiger, aber sequentieller Bearbeitung mehrerer Werkstücke. Der physikalische QR-Code wird gelesen und die ID wird an den Cloud-Server gesendet, der die werkstückspezifischen Erwärmungs-, Schmier- und Schmiedeprozessparameter an den Industrie-PC zurückgibt und die Erwärmung unmittelbar anschließend startet. Nachdem das vorherige Werkstück geschmiedet und auf den Stapeltisch zurückgebracht wurde, führt der Roboter die Schmierung der Werkzeuge durch. Nach der induktiven Erwärmung wird die "Anschlaghöhe" zwischen Ober- und Unterwerkzeug durch eine präzise Positionierung des Pressenstößels spezifisch auf die aktuellen Bolzenabmessungen eingestellt. Ziel dieser Stößelpositionierung zur Steuerung der Anschlaghöhe ist die Vermeidung von Gratbildung in der ersten Schmiedeoperation bei geschlossenen Werkzeugen (siehe Abbildung 2.6). Die Vorform wird in geschlossenen Gesenken geschmiedet und sowohl die Weg-Kraft-Messung des Stößels als auch die gemessenen Werkzeugtemperaturen werden an den Cloud-Server gesendet. Bevor das Werkstück in das letzte Schmiedegesenk transportiert wird, bewegt der Roboter das Werkstück kurz unter dem Laserprofilometer und dem Pyrometer zur

# EMuDig4.0

Vermessung der Geometrie und der Temperatur. Diese Daten werden ebenfalls an den Cloud-Server weitergegeben, aber auch dazu verwendet, um zu entscheiden, ob die Teilequalität für die weitere Verarbeitung in der letzten Schmiedestufe noch akzeptabel ist. Wenn die Gratbildung durch die Geometriemessung als „zu groß“ erkannt wird, wird das Werkstück in der zweiten Schmiedestufe nicht weiterbearbeitet und der Roboter bringt das Teil in die Schrottbox. Der Werkstückzähler auf dem Cloud-Server wird automatisch entsprechend dem aktuellen Zustand des Werkstückverfolgungssystems aktualisiert. Währenddessen wird die Anschlaghöhe für die zweite Schmiedestufe wieder werkstückspezifisch angepasst, falls das Teil weiterbearbeitet werden soll. Die zweite Schmiedestufe wird in ähnlicher Weise durchgeführt, mit anschließender Geometrie- und Temperaturmessung, bevor der Roboter das Fertigteil auf den Stapeltisch bringt. Die Messwerte des gesamten Prozesses werden mit der Werkstück-ID verknüpft und definieren später, ob die erreichte Qualität als akzeptabel betrachtet wird oder nicht. Der Roboter legt die bearbeiteten Teile in der gleichen Bearbeitungsreihenfolge auf den Stapeltisch.

## 2.2.2 Adaptive Prozesssteuerungsstrategie

Um die Erfahrung des Operators zu nutzen und unerwartete Systemstörungen zu beseitigen, wird ein "Assistenzsystem" als Hauptgateway des Kontrollsystems programmiert. Das Assistenzsystem ermöglicht es dem Operator, die vom Cloud-Server vorgeschlagenen Prozessparameter zu überprüfen, zu korrigieren und zu genehmigen. Der Cloud-Server übernimmt in diesem Zusammenhang die Aufgaben der Eigenschaftskontroll-Einheit und des Prozessplaners, wie in Abbildung 2.7 dargestellt. Es werden drei Rückkopplungsebenen in Bezug auf unterschiedliche Regelungsstrategien, die auf solche Arten von Fertigungsprozessen anwendbar sind, dargestellt: Die "Echtzeit-Online-Regelstrategie" ist in der Lage, die anstehenden Prozessparameter hinsichtlich der Historie und der tatsächlichen Messungen aus der Bearbeitung des aktuellen Werkstücks anzupassen. Die Prozessparameter werden nur zwischen den Prozessstufen geändert, nicht aber während einer einzelnen Prozessstufe. Die "Langzeit-Online-Regelstrategie" berücksichtigt Prozessparameter, die nicht schnell verändert werden können. Werkzeugtemperaturen zum Beispiel gehören zum Grundanliegen dieses Regelkreises. Die "Langzeit Offline-Regelstrategie" erfordert die mikrostrukturelle oder detaillierte Qualitätsanalyse als Grundlage für Vorschläge zu den Grenzwerten der Prozessparameter. Werkzeugtemperatur, Bolzenerwärmungstemperatur und Umformgeschwindigkeit haben allesamt Einfluss auf die mikrostrukturellen Eigenschaften und sind das Grundanliegen dieser Art von Regelkreis. Alle Regelkreise sind durch das Assistenzsystem wirksam und bedürfen in jedem Fall der Zustimmung des Bedieners. Die Online-Messung von Prozessparametern und Werkstückeigenschaften (wie Geometrie und Temperatur) tragen zur Datenbank des vom Regler verwendeten Metamodells bei. Zu diesem Zweck ist Ausschuss ebenso nützlich wie akzeptable Teile bei der Verbesserung des Metamodells. Theoretisch ist es möglich, das Metamodell nach jedem Werkstück oder nach einer Produktionscharge zu aktualisieren. Künstliche neuronale Netze (KNN) stellen einen geeigneten Rahmen dar, der die Möglichkeit der Modellverbesserung mit jedem einzelnen Datenpunkt oder mit einer Gruppe von Datenpunkten bietet.

# EMuDig4.0

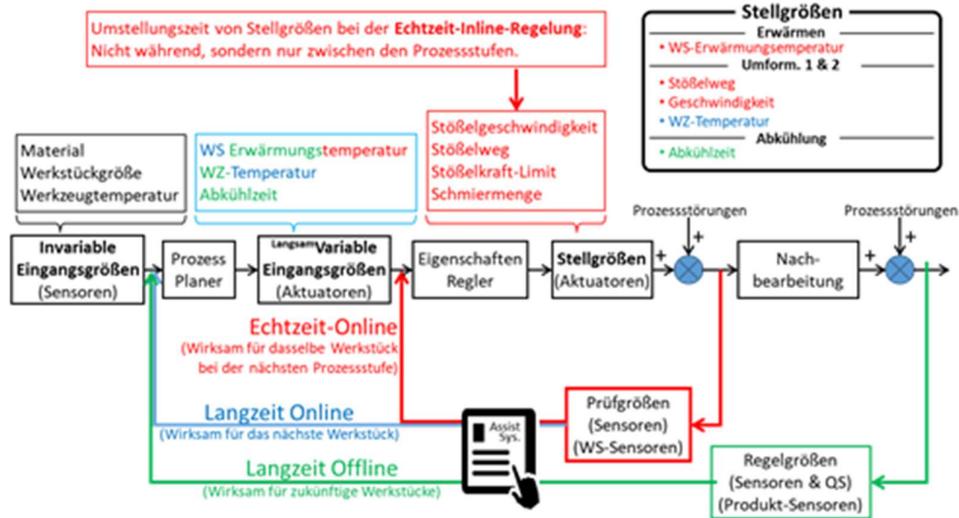


Abbildung 2.7: Kontrollstrategien der Modellfabrik

## 2.2.3 Cloud-Technologien für Datenspeicherung, Berechnung und Assistenzsystem

Cloud-Technologien bieten eine flexible Plattform für das Internet der Dinge (IoT). Die Speicherung des Zustands des Herstellungsprozesses und der Prozessdaten an einem gemeinsamen Ort ermöglicht die Analyse großer Datenmengen, um starke oder schwache Korrelationen zu identifizieren. Diese Korrelationen eröffnen den Weg zu einem tieferen Verständnis der Prozesse, die für den Menschen normalerweise nicht sichtbar sind. Die internetbasierte Handhabung von Daten ist auch der flexibelste Ansatz, um sie an jedem Ort und in jedem Format nach Bedarf verfügbar zu machen. Zustandswerte und aktuelle Werte können dem Bediener der Druckmaschine auf jedem Bedienfeld oder sogar auf einem verschleißfreien Visualisierungsgerät zur Verfügung gestellt werden. Die Gesamtleistung der Produktionskette kann dem Produktionsingenieur auf einem webbasierten Tool am Computer (oder auf dem Tablett) demonstriert werden. Cloud-Technologien für Extract-Transform-Load (ETL), Datenspeicherung, Modellierung, Berechnung, Überwachung und Verfolgung werden vom IAS an der Universität Stuttgart entwickelt.

Das entwickelte Konzept, das aus solchen Elementen besteht, ist in Abbildung 2.1 dargestellt. Die gemessenen Prozessparameter werden konvertiert und auf die standardisierte Speicherschnittstelle, die auf dem Cloud-Server läuft, übertragen. Die standardisierte Schnittstelle leitet die Daten zur weiteren Verarbeitung weiter. Sie werden entweder in einer relationalen Datenbank gespeichert, in der Datenanalysetechniken angewendet werden können, oder an eine Echtzeit-Datenbank weitergeleitet, um Ad-hoc-Berechnungen und analytische Online-Verarbeitung (OLAP) durchzuführen. Die Echtzeitdatenbank wird für die Zuordnung von Werkstückidentität und Prozessparametern entlang der gesamten Prozesskette verwendet. Sie ermöglicht eine Online-Verfolgung von Einzelteilen oder Chargen.

Um eine effiziente Überwachung und Verfolgung von Prozessdaten zu erreichen, wurde ein OLAP-Modell unter Berücksichtigung der Datenräume und der entsprechenden Prozessschritte erstellt. Das Modell ist in die Dimensionen Zeit, Raum und Produktinformationen unterteilt, Abbildung 2.2. Die Dimensionen enthalten eine interne, hierarchische Struktur, in der jede Verzweigung durch ein Element beschrieben wird. Grundelemente sind mit Datenvektoren verbunden und markieren die

# EMuDig4.0

innerste Schicht des Datenmodells (Produkttyp P mit id x). Höhere Schichten (Charge A des Produkttyps P) kumulieren die Informationen der unteren Schichten durch Anwendung von Aggregationsverfahren. So können heterogene Daten in ein homogenes Modell integriert werden, da es möglich sein kann, verschiedene Schichten mit unterschiedlicher Auflösung zu assoziieren. Dieser Faktor ermöglicht die Verfolgung einer flexiblen Menge von Teilen durch die Navigation durch solche Schichten. Der Modellierungsansatz ist in Abbildung 2.2 als Würfel dargestellt.

Das grundlegende und entscheidende Assistenzsystem, das der Maschinenbediener verwendet, ist auf dem Bedienungspult der Presse installiert. Die Abbildung 2.8 zeigt den Parameterüberwachungs- und Dateneingabebereich des im Modellwerk verwendeten Assistenzsystems. Das Eingabedeck auf der rechten Seite in orangefarbener Farbe empfängt die vom Cloud-Server vorgeschlagenen Prozessparameter und ermöglicht es dem Bediener, diese Werte zu kontrollieren, zu akzeptieren oder zu ändern, bevor sie an die Geräte des Fertigungssystems gesendet werden. Die manuelle Anpassung der Parameter durch den Bediener ist einerseits eine fehleranfällige Aufgabe. Andererseits ist der Bediener verpflichtet, für jedes Peripheriegerät und für jede Schmiedestufe innerhalb der richtigen Zeitperiode die richtigen Parameter anzugeben. Daher erleichtert der automatisierte Datenfluss zum richtigen Ziel (zur Presse oder zu den Peripheriegeräten: Schmiersystem, induktives Erwärmungssystem usw.) hinsichtlich der Verfolgung der Werkstücke in der Prozesskette die Aufgabe des Bedieners drastisch.

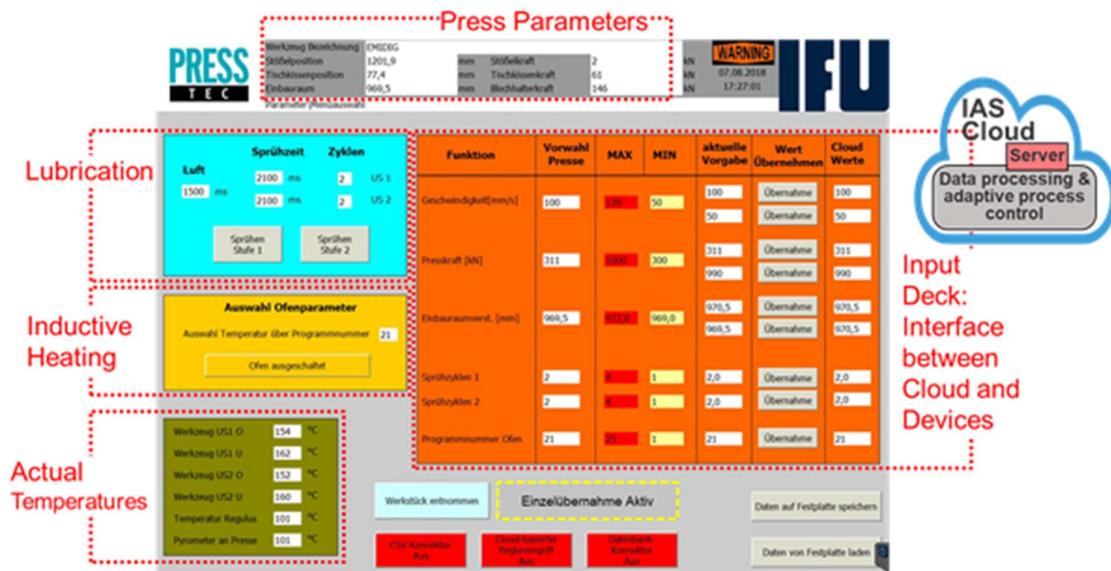


Abbildung 2.8: Schnittstelle des Assistenzsystems auf dem Bedienpult der Presse.

## 2.2.4 Anwendung des maschinellen Lernens als Modellierungsinstrument

Künstliche neuronale Netzwerke (KNN) mit einer verborgenen Schicht mit 10 Neuronen sollen als Regelalgorithmus im Cloud-Server dienen. Für jede Stellgröße (z.B. Anschlaghöhe, induktive Erwärmungstemperatur) wird ein separates KNN bevorzugt. Um die Anzahl der für das Training notwendigen Experimente zu reduzieren, werden FEA zum Vor-Training der ANN verwendet. Abbildung 2.9 illustriert die Responsfläche für den durch 1000 FEA trainierten Prozessparameter

## EMuDig4.0

„Anschlaghöhe“ in roter Farbe. Kugeln werden in dieser Abbildung als Datenpunkte verwendet, um die Grathöhe (in Form von Kugeldurchmessern) nach der ersten Schmiedestufe darzustellen. Die Grathöhe soll im Wesentlichen durch die richtige Kombination von Werkstückhöhe, Anschlaghöhe, Schmierstoffmenge und Werkstückerwärmungstemperatur minimiert werden. Die Verteilung der Datenpunkte in der Abbildung 2.9 ist so, dass die Gruppe mit der minimal möglichen Grathöhe durch eine hypothetische Oberfläche von den Datenpunkten mit unerwünschten Grathöhen getrennt ist. Auf dieser hypothetischen Fläche kann der optimale Wert der Anschlaghöhe gefunden werden. Die trainierte KNN ist in der Lage, diese hypothetische Fläche als Antwortfläche zu erkennen (oder zu lernen), die zur Berechnung von Anschlaghöhe verwendet werden kann. Tausende von Datenpunkten, die durch die FEA des Schmiedeprozesses gewonnen werden, ermöglichen ein erstes Training. Danach wurde eine Reihe von Schmiedeeexperimenten durchgeführt, die eine akzeptable Qualität oder eine inakzeptable Ausschussproduktion ergaben. Das KNN wird mit nur 46 experimentellen Datenpunkten weiter trainiert und die Responsefläche entwickelt sich, um die experimentellen Erkenntnisse zu erfassen. Dieses neue Modell ermöglicht es nun, eine geringere Menge an Schrott zu generieren. Die Weiterentwicklung des KNN kann auch im Anschluss an abgeschlossene Produktionsprozesse durchgeführt werden. Dabei wird die Regelungssoftware im Cloud-Server entsprechend überarbeitet, um eine Online-Anpassung des Regelungsmodells zu realisieren.

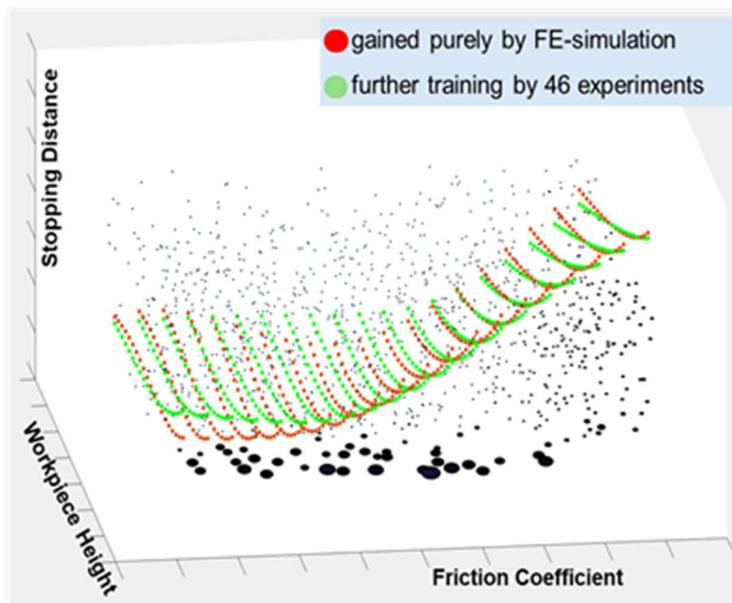


Abbildung 2.9: Responseflächen, die vom trainierten KNN für "Anschlaghöhe" erzeugt werden. Die Datenpunktdurchmesser repräsentieren die Grathöhe des Werkstücks.

## 3 Produktionsanlagen

Angestrebtes Ergebnis im Schwerpunkt „Produktionsanlagen“ war die Realisierung eines ganzheitlichen Messkonzepts für Frühindikatoren von Bauteil- und Baugruppenschäden sowie die Ableitung von Prognosemodellen zur Vorhersage ungeplanter Stillstände in der Massivumformung. Anhand zweier repräsentativer Produktionslinien der Anwendungspartner sollten die Modelle erprobt werden. Dies betraf die Planung gezielter Wartungsmaßnahmen zur Erhaltung der Anlagenverfügbarkeit beispielsweise auf Basis eines verbliebenen Abnutzungsvorrats sowie ihre Durchführung. Dafür wurde zum einen eine hydraulische Räderpresse (Hersteller: Müller Weingarten, Betreiber Otto Fuchs) und zum anderen eine Exzenterpresse (Hersteller: SMS, Betreiber: Hirschvogel) für Komponenten des Automobilbereichs ausgewählt.

### 3.1 Messkonzepte und Auswahl der Sensorik

Zu Projektbeginn wurden in Workshops mit den Anlagenbetreibern Hirschvogel und Otto Fuchs mögliche Schadensfälle identifiziert, die sich signifikant auf die GAE auswirken können. Nachfolgend wurde geprüft, welche der identifizierten Schadensfälle potentiell mit vorhandenen Signalen der Automatisierung oder zusätzlicher Sensorik direkt oder indirekt beobachtbar sind. Unter Einbeziehung einer Wirtschaftlichkeitsbetrachtung wurden folgende Bauteile bzw. Teilsysteme zur Überwachung ausgewählt.

Tabelle 3.1: Hydraulische Räderpresse, Otto Fuchs

Überwachungsmodul	Bewertungsgrundlage	Zusätzliche Sensorik / Hardware
Hauptpumpen	Vibration, Druckverlauf	Beschleunigungssensoren, Drucksensoren an Hauptpumpeneinzelleitungen
Hauptpumpenanstellung	Schwenkwinkelregelungsverhalten	keine
Hauptpumpen- und Antriebslager, Kupplung	Vibration	Beschleunigungssensoren
Ventile Stoßelanstellung	Druckverlauf	Drucksensoren an Hauptpumpeneinzelleitungen
Steuerdruckpumpen	Vibration, Ladeschaltzyklus	Vibrationssensoren
Steuerdruckpumpen- und Antriebslager, Kupplung	Vibration	Beschleunigungssensoren
Steuerdruckspeicher	Druckverlauf, Vorspanndruck	Gasdrucksensoren
Stoßelanstellung interne Leckage	Stoßelbewegung in Ruhelage	keine
Hydraulik Medium	Ölstand, Kontamination	Optische Feststoffkontaminationsmessung, Wassersättigungssensor

# EMuDig4.0

Tabelle 3.2: Exzenterpresse, Hirschvogel

Überwachungsmodul	Bewertungsgrundlage	Zusätzliche Sensorik / Hardware
Hydraulisches Kupplungs-Bremssystem	Druck-, Positionsverläufe und Ventilschaltverhalten, Konformität mit Ablaufsteuerung	Drucksensoren für Kupplungs-, Brems- und Speicherdrücke, Eingangsklemmen für Ventiltrückmeldung, Drehzahlgeber mit erhöhter Auflösung
Hauptantrieb	Motor Drehmomentbildender Strom im Leerlauf	keine
Pressenständer, Stößelverstellung	Pressenständerdehnung, Vibration, Position Stößelverstellung	Beschleunigungssensoren
Prozessführung hinsichtlich Umformkräften	Umformenergie, Schmierstoffsprühung, Pressenständerdehnung, Werkstücktemperatursteuerung	Dehnungsmessstreifen Pressenständerdehnung, Drehzahlgeber mit erhöhter Auflösung
Exzenter-, Druckstangen und Exzenterlager	Temperaturverläufe	keine
Tischausstoßer	Druckverläufe, Ausfuhr-, Einfuhrweg	Drucksensoren für Tischausstoßer
GEA	Motortemperaturverläufe, Verlauf drehmomentbildende Ströme, Regelungsverhalten Positionsregelungen	Temperatursensoren an Antrieben
Greifer	Greiferendschaltersignalverlauf, Konformität GEA und Tischausstoßer	Eingangsklemmen für hochauflösende Rückmeldung Greiferendschalter
Hydraulik Medium	Ölstand, Kontamination, diverse Öleigenschaften, Öltemperatur	Optische Feststoffkontaminationsmessung, Wassersättigungssensor, Multifunktionaler Ölzustandssensor (HydacLab)
Pumpen, Pumpen- und Antriebslager, Kupplung	Vibration	Beschleunigungssensoren
Induktivofen	Materialtemperatur, Spannung, Leistung, Frequenz, Taktzeit	keine

Bei beiden Demonstratoranlagen wird das Messdatenerfassungssystem ibaPDA der Firma iba AG verwendet, das jeweils auf einem Windows-Betriebssystem auf einem Hutschienen-PC im Feldschrank nahe der Maschine betrieben wird. In Abbildung 3.1 ist der Signalfluss schematisch am Beispiel der Demonstratoranlage von Hirschvogel dargestellt. Über einen Busmonitor werden Mess- und Steuerungsdaten vom Automationssystem erfasst. Die Messsignale von zusätzlichen Sensoren werden über das dezentrale Ein-/Ausgabesystem ibaNet750-BM-D mit daran angeschlossenen Eingangsmodulen sowie über Zentraleinheiten ibaPADU-S-CM mit angeschlossenen ibaMS8xICP und ibaMS16xAI erfasst. Auf demselben PC werden die Messdaten mit der proprietären Genius CM Software verarbeitet, sodass Schäden mittels Kennwerten diagnostiziert und entsprechende Alarmer generiert werden.

# EMuDig4.0

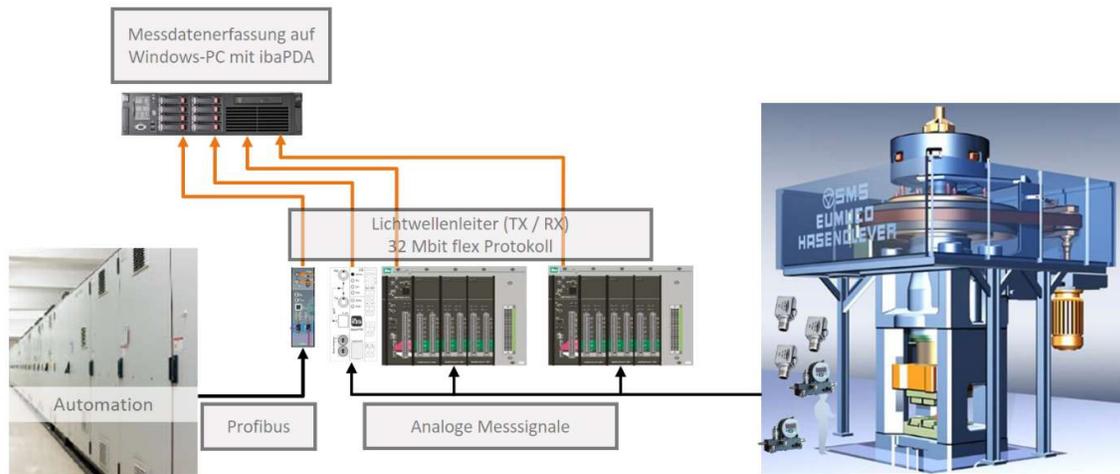


Abbildung 3.1: exemplarischer Signalfluss an der Hirschvogel-Press

## 3.2 Prognosemodell

### 3.2.1 Hydraulische Räderpresse, Otto Fuchs

Von den unter Modul 3.1 genannten Überwachungen konnten folgende Module durch auftretende Verschleißerscheinungen und schadensbedingte Ausfälle validiert werden, sodass sowohl die Analysemethoden selbst als auch die Schwellwerte zur Alarmgenerierung optimiert und verifiziert werden konnten. Dabei werden in der folgenden Auflistung nur Überwachungsmodule genannt, die nachblickend als wirtschaftlich sinnvoll erachtet werden.

Tabelle 3.3: wirtschaftliche Überwachungsmodule – hydraulische Räderpresse, Otto Fuchs

Überwachungsmodul	Bewertungsgrundlage	Zusätzliche Sensorik / Hardware
Hauptpumpen	Druckverlauf	Drucksensoren an Hauptpumpeneinzelleitungen
Hauptpumpenanstellung	Schwenkwinkelregelungsverhalten	keine
Ventile Stößelanstellung	Druckverlauf	Drucksensoren an Hauptpumpeneinzelleitungen
Steuerdruckpumpen	Ladeschaltzyklus	keine
Steuerdruckspeicher	Druckverlauf	keine
Stößelanstellung interne Leckage	Stößelbewegung in Ruhelage	keine
Hydraulik Medium	Ölstand	keine

## EMuDig4.0

Darauf basierend wurden Prognosemodelle für den Tankölstand und den Steuerdruckpumpenzustand erstellt. In Abbildung 3.2 ist exemplarisch eine Prognose für den Tankstand dargestellt. Mittels Regression werden Tankfüllstand (radiometrisch erfasst), Öltemperatur und Pressenposition miteinander verrechnet, um den entsprechenden Füllstand in Ruhe und bei Referenztemperatur zu erhalten. Aus dem Verlauf dieses Wertes (türkiser Graph) lässt sich eine Leckölrate (Steigung der blauen Linie) ableiten, sowie eine Prognose, wann spätestens aufgefüllt werden muss (Schnittpunkt der blauen Linie mit der rot gezeichneten Ausfallgrenze).

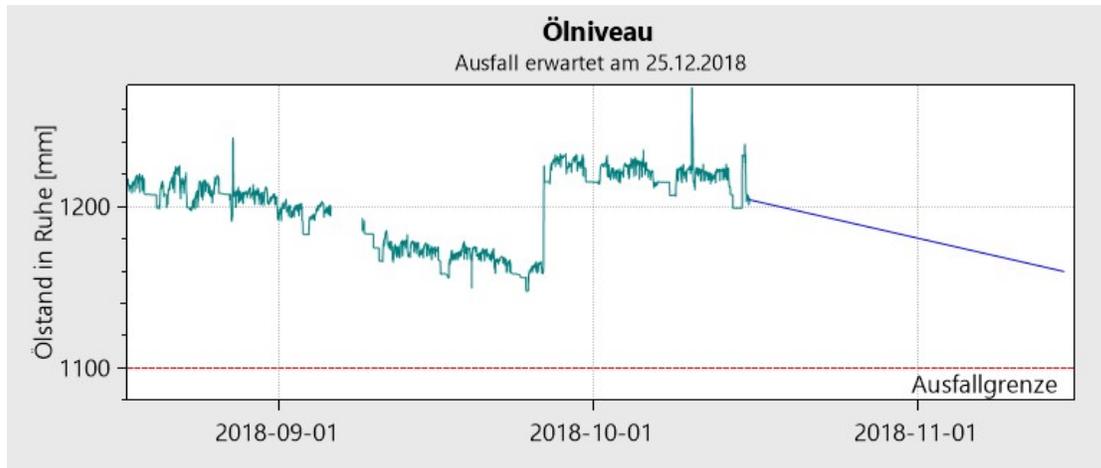


Abbildung 3.2: Prognosemodell Tankstand – Ölniveau. Durch Lineare Regression können Leckölrate und Nachfüll-Intervall bestimmt werden.

Nach selbiger Vorgehensweise wurde eine Prognose für die Steuerhochdruckpumpe umgesetzt, allerdings ist die Aussagekraft des vorhergesagten Ausfallszeitpunkts fraglich. Für die Entwicklung anspruchsvollerer, qualitativ hochwertigerer Prognosen von Ausfällen durch fortschreitende Schädigungen ist eine genaue Klassifizierung der aufgetretenen Schadensarten an den jeweiligen Bauteilen und eine hohe Schadensfallzahl Grundvoraussetzung. Beides ist hier nicht der Fall (2 Ausfälle der Pumpe während der Projektlaufzeit).

# EMuDig4.0

## 3.2.2 Exzenterpresse, Hirschvogel

Von den unter Modul 3.1 genannten Überwachungen werden folgende Module als potentiell wirtschaftlich sinnvoll erachtet und weiterverfolgt.

Tabelle 3.4: wirtschaftliche Überwachungsmodule, Exzenterpresse, Hirschvogel

Überwachungsmodul	Bewertungsgrundlage	Zusätzliche Sensorik / Hardware
Hydraulisches Kupplungs-Bremssystem	Druck-, Positionsverläufe und Ventilschaltverhalten, Konformität mit Ablaufsteuerung	Drucksensoren für Kupplungs-, Bremsdruck, Eingangsklemmen für Ventilrückmeldung, Drehzahlgeber mit erhöhter Auflösung
Hauptantrieb	Motor drehmomentbildender Strom im Leerlauf	keine
Pressenständer, Stößelverstellung*	Pressenständerdehnung, Vibration, Position, Stößelverstellung	keine
Prozessführung hinsichtlich Umformkräften*	Umformenergie, Schmierstoffsprühung, Pressenständerdehnung, Werkstücktemperatursteuerung	Drehzahlgeber mit erhöhter Auflösung
Tischausstoßer	Druckverläufe, Ausfuhr-, Einfuhrweg	Drucksensoren für Tischausstoßer
GEA	Verlauf drehmomentbildende Ströme, Regelungsverhalten, Positionsregelungen	keine
Greifer	Greiferendschaltersignalverlauf, Konformität GEA und Tischausstoßer	Eingangsklemmen für hochauflösende Rückmeldung Greiferendschalter
Induktivofen	Materialtemperatur, Spannung, Leistung, Frequenz, Taktzeit	keine

\*: Wurde im Projektzeitlauf nicht untersucht/ entwickelt, sodass der potentielle Nutzen nicht bewertet werden kann.

Für die Entwicklung von Prognosen hat sich im Fall der Exzenterpresse lag die Herausforderung speziell bei der Klassifizierung von Störungen sowie in der Sicherstellung der nötigen Messdatenqualität. Die verfügbaren Störmeldungen basieren auf subjektiven Einschätzungen der Anlagenbediener und Instandhalter. Messsignale von zentraler Bedeutung waren bis zuletzt nicht immer zuverlässig. Zurzeit wird noch geprüft, ob mit der vorhandenen Datenbasis die Entwicklung eines Störungsprognosemodells ausreichend genau möglich ist.

## 3.3 Definition Wartungsmaßnahmen

### 3.3.1 Hydraulische Räderpresse, Otto Fuchs

Im Fall der hydraulischen Räderpresse haben sich einige, wiederholende Schadens- und Mängelarten ergeben, die sich mit der zuvor ausgewählten Sensorik und Automationssignalen einfach erkennen und klar voneinander unterscheiden lassen. Ein Ausschlag der Kennwerte, die aus den Mess- und Automationssignalen berechnet werden, ergeben eine klare Indikation für spezifische Instandhaltungsmaßnahmen. Zur Verdeutlichung werden im Folgenden einige Beispiele genannt.

**Bsp. 1:** Der Kennwert für die Ventile zwischen Hauptpumpe P4 und den Anstellzylindern schlägt wie in Abbildung 3.3 dargestellt bei den markierten Zeitpunkten aus und überschreitet den definierten Grenzwert, sodass ein Alarm ausgelöst wird. Dies deutet auf ein Festsitzen eines Ventils hin, was sich häufig durch manuelles Betätigen der Ventile einfach beheben lässt. Ist dies nicht möglich oder es verklemmt sich kurz darauf wieder, ist das Ventil zu wechseln.

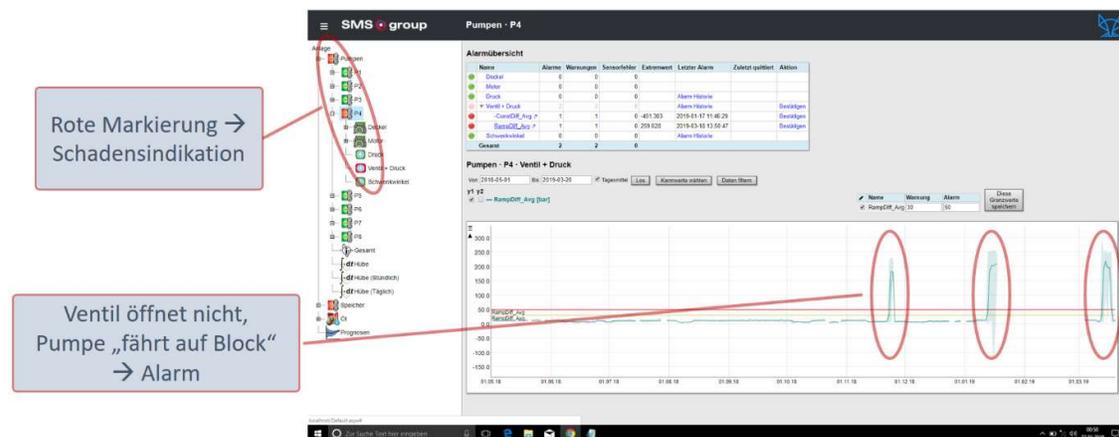


Abbildung 3.3: Ventilkennwerte Hauptpumpe P4/Anstellzylinder – Kreise: Festsitzende Ventile; zu Ersetzen bei wiederholtem Auftritt

**Bsp. 2:** Der Kennwert für die Leistungsfähigkeit von Hauptpumpe P3 schlägt wie in Abbildung 3.4 dargestellt bei den markierten Zeitpunkten aus und unterschreitet den definierten Grenzwert, sodass ein Alarm ausgelöst wird. Beim nächsten Stillstand werden die Leitungsanschlüsse geprüft. Sofern diese ohne Beanstandung sind, wird die Pumpe durch eine neue ersetzt.

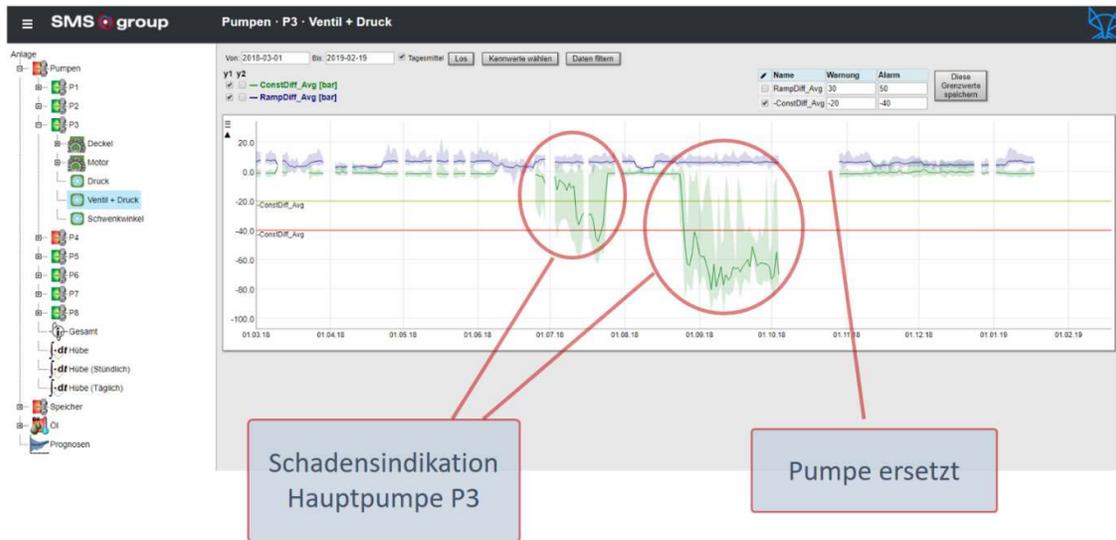


Abbildung 3.4: Kennwert Leistung Hauptpumpe P3 – Kreise: Schadenssituationen im Vorfeld des Pumpentauschs

Abbildung 3.5 zeigt Schäden an Bauteilen der ausgebauten Hauptpumpe P3. Von links nach rechts sind die Steuerplatte, exemplarisch die Kopfseite eines Kolbens und ein Ausschnitt der Zylindertrommel, sowie die Nahaufnahme in einen Kolbenraum der Zylindertrommel dargestellt.



Abbildung 3.5: Schäden an Komponenten der fehlerhaften Hauptpumpe A3 nach dem Ausbau.

# EMuDig4.0

**Bsp. 3:** Der in Abbildung 3.6 dargestellte Kennwert für die Leistungsfähigkeit der Steuerhochdruckpumpe überschreitet im markierten Zeitbereich den definierten Grenzwert, sodass ein Alarm ausgelöst wird. Nach Ausschluss des Speicherzustands bzw. des Vorspanndrucks des Speichers als Ursache, werden die Leitungsanschlüsse der Pumpe geprüft. Sofern diese ohne Beanstandung sind, wird die Pumpe durch eine neue ersetzt.

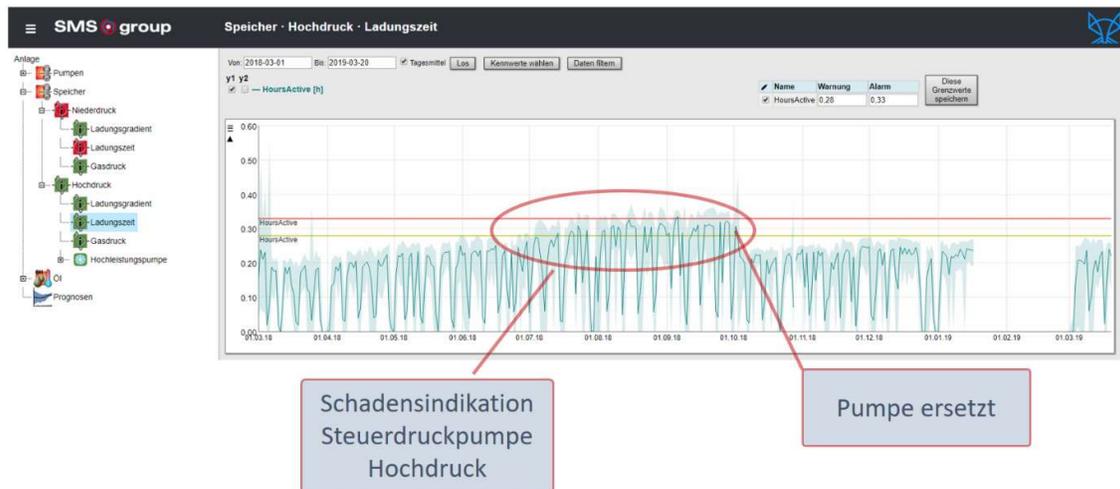


Abbildung 3.6: Kennwerte der Leistungsfähigkeit der Steuerhochdruckpumpe vor Austausch.

**Bsp. 4:** Der Kennwert für das Absacken des Stößels in Ruhestellung, bei der die Position durch Schließen der Schaltventile gehalten werden soll, ist in den markierten Zeitbereichen erhöht. Im ersten Fall links wird der Grenzwert wie in Abbildung 3.7 dargestellt überschritten, sodass ein Alarm ausgelöst wird. Dies deutet auf eine Leckage hin, was aus der Erfahrung der Hydrauliker der SMS in den meisten Fällen auf eine interne Leckage im Zylinder zurückzuführen ist. Dies hat sich in diesem Fall bestätigt, der Kennwert ist nach dem Wechsel der Dichtringe drastisch reduziert.



Abbildung 3.7: Kennwert für Stößel in Ruhestellung. Nach Übersteigen eines Grenzwerts vermuten Hydraulik-Techniker eine interne Leckage im Zylinder. Nach Austausch der Dichtringe sinkt der Kennwert wieder unter den Schwellwert.

## 3.3.2 Exzenterpresse, Hirschvogel

Im Fall der Exzenterpresse von Hirschvogel hat sich die Entwicklung der Kennwerte aus organisatorischen Gründen, regulatorischen Hindernissen und mangelnder Datenkonsistenz stark verzögert, sodass die Kennwertberechnungen zurzeit noch in der Validierungs- und Optimierungsphase sind. Exemplarisch werden im Folgenden vier aufgetretene Mängel in der Auswirkung auf die bereits vorhandenen Kennwerte bzw. Messdaten dargestellt, für welche die daraus abzuleitenden Wartungsmaßnahmen benannt werden können.

**Bsp. 1:** In Abbildung 3.8 sind Kennwerte des Hydrauliksystems für die Kupplung dargestellt. Beim markierten Datum wurde das Pressensicherheitsventil aufgrund der zuvor erhöhten Zeit vom Schaltbefehl des Ventils bis zum Druckaufbau ausgetauscht. Nach dem Wechsel ist der entsprechende Kennwert reduziert.

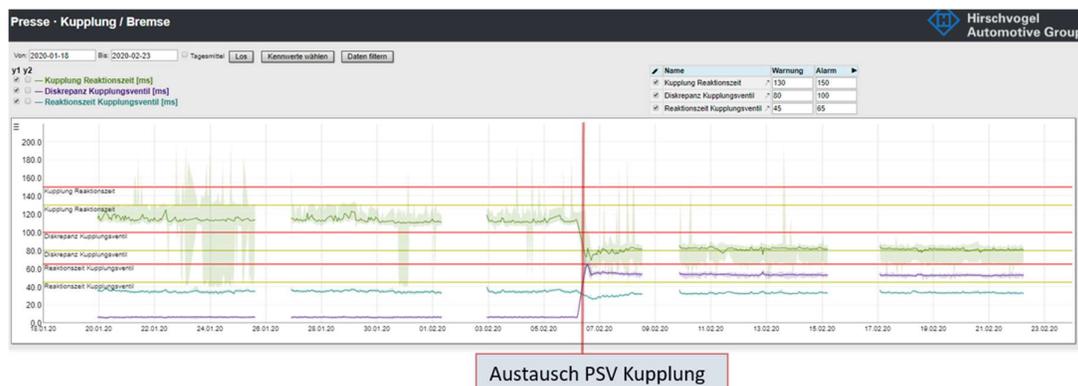


Abbildung 3.8: Status des Hydrauliksystems. Reaktionszeit zwischen Schaltbefehl des Ventils bis zum Druckaufbau erhöht. Nach dem Austausch der PSV-Kupplung sinkt Zeit messbar ab.

**Bsp. 2:** Abbildung 3.9 zeigt Trendgraphen von Kennwerten zu den Drehmomentverläufen der GEA Vor-/Rückwärtsbewegung für vorne links (oben) und vorne rechts (unten) dargestellt. Jeweils darunter sind über einen Tag gemittelte Drehmomentverläufe entsprechend dem Datum auf dem der Cursor im Trendgraph steht und Referenzkurven, die anhand eines Zeitraumes ermittelt wurden, in dem die Motorsteuerung einwandfrei funktioniert hat. Der Zeitraum, in dem die Motorsteuerung einwandfrei funktioniert hat, entspricht im Bild dem Zeitraum zwischen dem 2. und 3. Marker von links (ca. 3,5 Monate). Die Referenzkurven wurden basierend auf Daten eines Monats innerhalb dieser Zeitspanne ermittelt. Die Kennwerte in den Trendgraphen sind im Zeitraum, in dem die Motorsteuerung einwandfrei war, konstant deutlich niedriger als davor und danach. Anhand der Drehmomentverläufe ist zu erkennen, dass die Kennwerterhöhungen außerhalb dieses Zeitraums einem zwischen links und rechts gegensinnigem Offset der Drehmomentverläufe entspricht. Die Motoren arbeiten dann gegeneinander. In diesem Fall ist die Parametrierung der Antriebssteuerung zu prüfen.

# EMuDig4.0

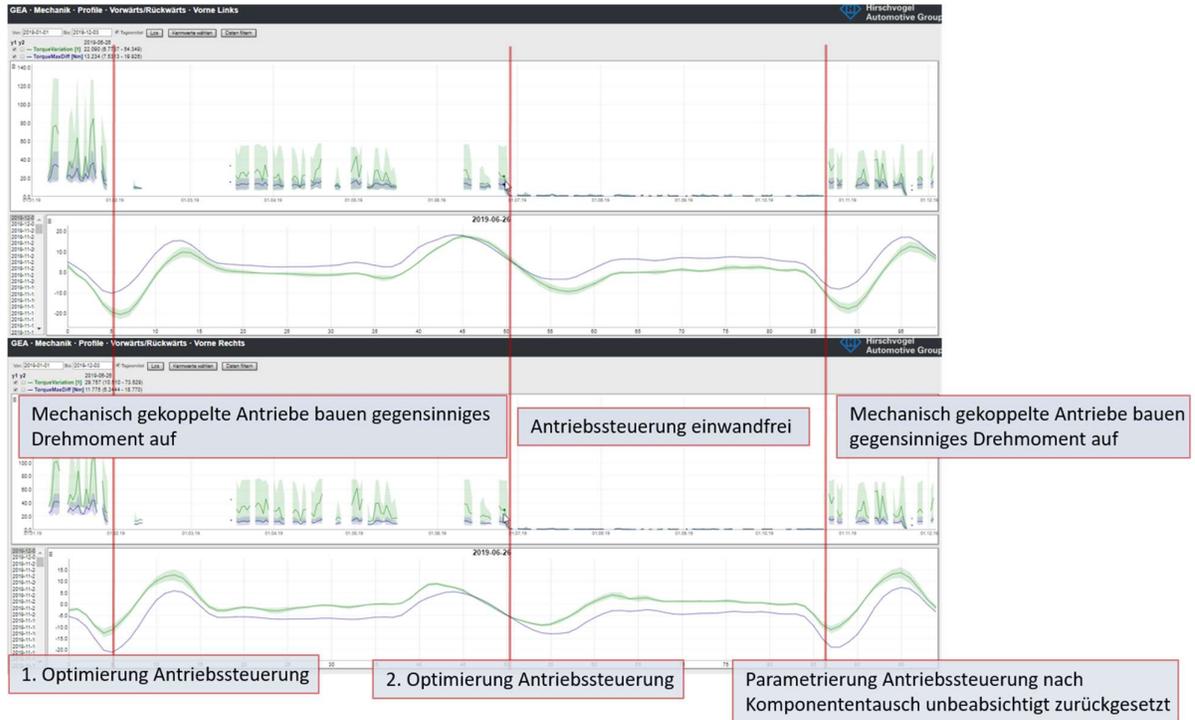


Abbildung 3.9: Trendgraphen von Kennwerten zu den Drehmomentverläufen der GEA Vor-/Rückwärtsbewegung für vorne links (oben) und vorne rechts (unten)

**Bsp. 3:** In Abbildung 3.10 ist der Trendgraph der Kennwerte für mechanisches Spiel aller Achsen dargestellt. Im markierten Zeitraum ist Kennwert für die Heben-/Senken-Achse vorne links stark erhöht. Nach Einstellung des Zahnspiels ist der Kennwert deutlich reduziert.

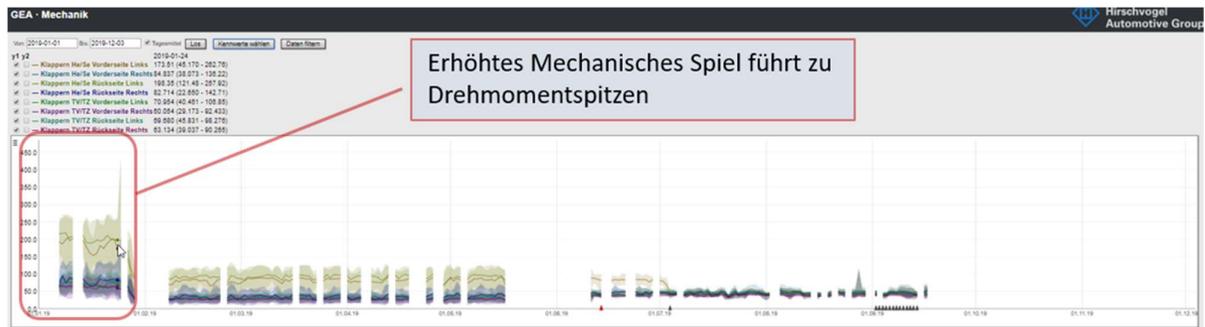


Abbildung 3.10: Trendgraph „mechanisches Spiel der Achsen“ – zu hohes Spiel führt zu Drehmomentspitzen. Neueinstellung reduziert dies deutlich

# EMuDig4.0

**Bps. 4:** Im Laufe des Projektes wurde ein Schaden am Dämpfungszyinders des Kupplungssystems festgestellt, der zu hohen Spitzen im Kupplungsdruck beim Einkuppeln und ungewollt schneller Beschleunigung der Kurbelwelle auf Nenndrehzahl geführt hat. In Abbildung 3.11 ist der Trendgraph zu zwei Kennwerten basierend auf dem Kupplungsdruckverlauf dargestellt. Der Kennwert im Trendgraph grün gibt die Höhe der Druckspitze beim Einkuppelungsbeginn wieder. Bei beschädigten Dämpfungszyindern ist dieser Kennwert stark erhöht. Nach Austausch der defekten Dämpfungszyinder ist dieser Kennwert deutlich reduziert. Danach wurden temporär weitere Maßnahmen getroffen, um die Druckspitzen beim Einkuppeln weiter zu vermindern und ein weiches Einkuppeln zu erreichen. Dabei wurde erreicht, dass die Kupplung langsamer Druck aufbaut. Da aber die Steuerung nicht zeitgleich angepasst wurde, hat dies zu einer schlechteren Abstimmung mit der Bremse geführt und zu einem härteren Einkuppeln, was sich in höheren Druckspitzen bei Einkuppelungsende widerspiegelt (lila Kurve).

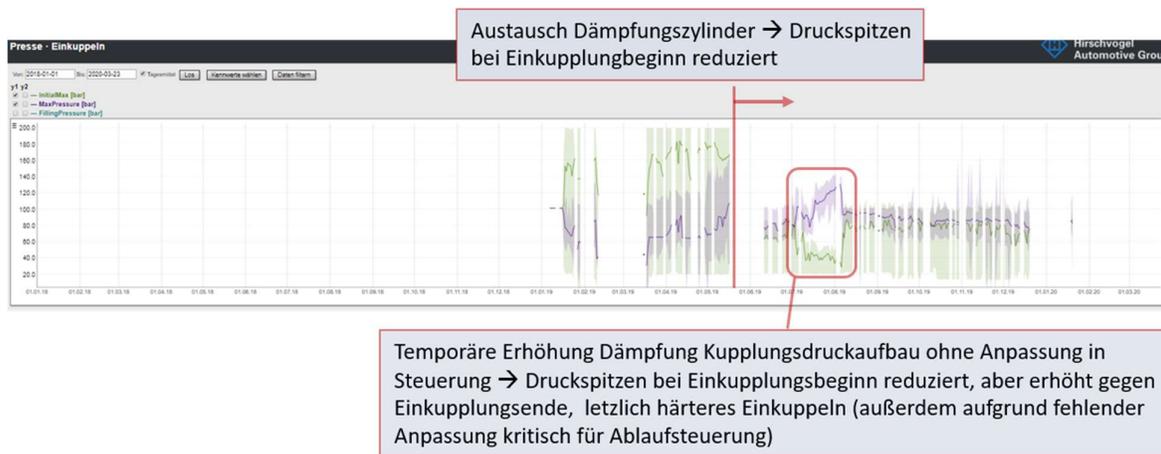


Abbildung 3.11: Kupplungsdruckverlauf (Kennwerte) -

## 3.4 Benutzerschnittstelle/ Visualisierung, Datenschnittstelle

Die Benutzerschnittstelle des Zustandsüberwachungssystems Genius CM ist ein in Abbildung 3.12 dargestelltes Webinterface und kann von jedem beliebigen PC im angeschlossenen Firmennetzwerk aufgerufen werden. Auf der linken Seite der Webseite ist der Anlagenbaum dargestellt, in der die Analysemodule aufgeteilt nach den überwachten Baugruppen aufgeführt sind. Im Anlagenbaum wird durch die Farbgebung signalisiert für welche Baugruppen keine Warnungen (grün), nicht quitierte Warnungen (gelb) oder nicht quitierte Alarme (rot) vorliegen. Durch Klicken auf die Baugruppen im Anlagenbaum können die jeweilige Alarmübersicht und –Historie aufgerufen und dort Meldungen quittiert werden, sodass der Alarmstatus wieder zurückgesetzt wird. Auf der untersten Ebene sind die den Baugruppen zugeordneten Analysemodule dargestellt. Durch Klicken auf die Analysemodule wird die Darstellung des Kennwerttrendgraphen aufgerufen. Im Beispiel in Abbildung 3.12 ist das Analysemodul „Ventil + Druck“ von Pumpe P3 aufgerufen. Der Trendgraph zeigt den zeitlichen Verlauf des diagnostizierten Schadens / Mangels und die dazugehörigen Grenzwerte, deren Überschreitung zur Auslösung von Warnungen bzw. Alarmen führt. Rechts über dem Trendgraphen sind Eingabemöglichkeiten zur Änderung der Grenzwerte.

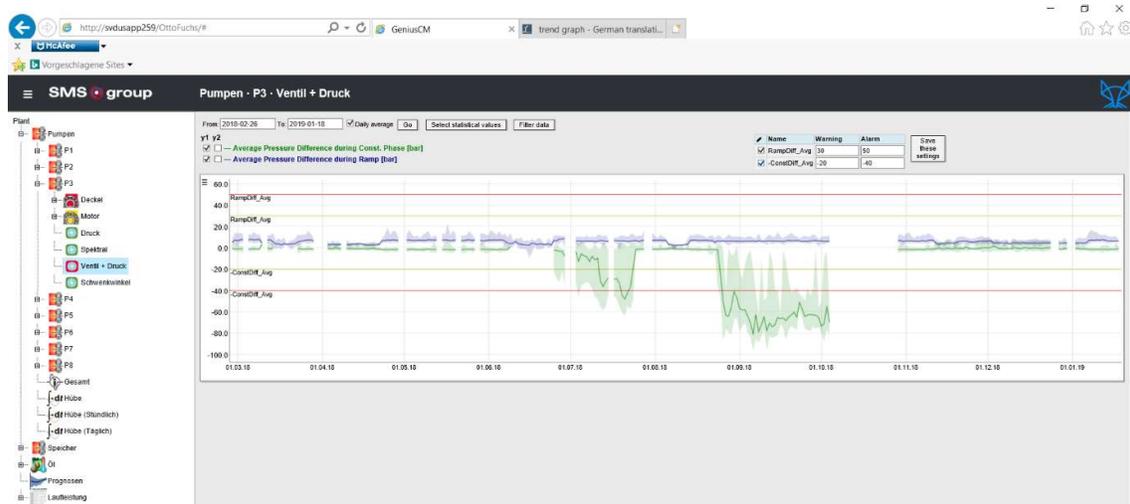


Abbildung 3.12: Darstellung des Web-Interface zur Zustandsüberwachung

# EMuDig4.0

Für die Anbindung an die Standort-IT / das ERP-System des Betreibers werden die Alarmstati per OPC-DA-Server bereitgestellt. In Abbildung 3.13 ist die Struktur der Analysesoftware inklusive der Datenschnittstelle für die Standort-IT / das ERP-System dargestellt, wobei im Unterschied zur Umsetzung bei Hirschvogel eine andere Möglichkeit der Anbindung an die Automation dargestellt ist. Die Analyseergebnisse des Zustandsüberwachungssystems Genius CM werden in einer SQLite-Datenbank gespeichert. Bei Bedarf können alle Datenbankdaten frei zugreifbar gemacht werden, z.B. mithilfe eines Wrappers, sodass die Daten dann mit open source Produkten wie PostgreSQL abgerufen werden können.

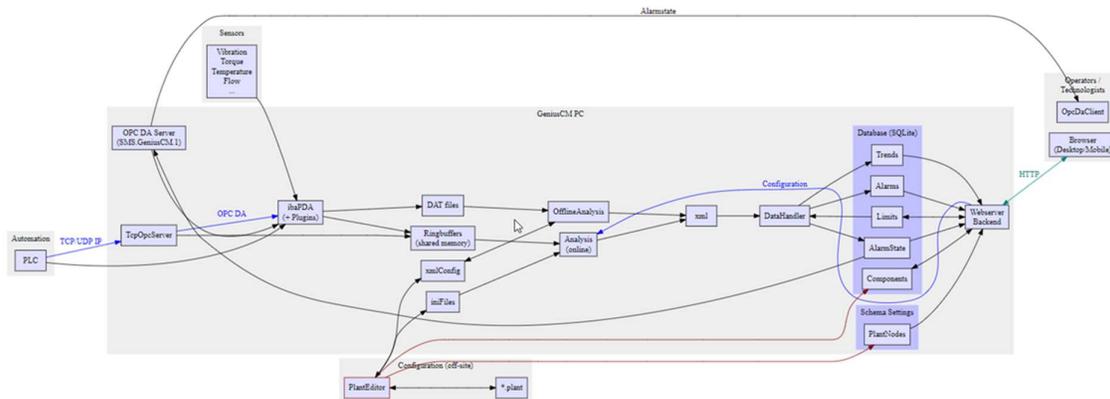


Abbildung 3.13: Struktur der Analysesoftware incl. Datenschnittstellen zu Standort-IT/ERP-System

## 4 Werkzeuge

Produktionswerkzeuge der Massivumformung unterliegen einer extremen Belastung, die sich aus einem Belastungskollektiv bestehend aus mechanischen, thermischen, tribologischen und chemischen Beanspruchungen zusammensetzt. Die Belastbarkeit der Werkzeuge hängt im Wesentlichen ab von den verwendeten Werkzeugwerkstoffen, sowie der Wärme- und Oberflächenbehandlung. Das Verhältnis Belastung und Belastbarkeit bestimmt schlussendlich die erreichbare Standmenge. Die betriebliche Praxis zeigt aber immer wieder, dass die Standmengen auch bei gleichen Produktionswerkzeugen teilweise erhebliche Schwankungen unterliegen.

Wachsender Kostendruck und steigende Rohstoffpreise drängen Industrieunternehmen zur nachhaltigen Produktion. Zur Senkung der Werkzeug- und Rüstkosten und damit der Herstellkosten von Umformteilen ist die Maximierung der Werkzeugstandmenge erforderlich. Bisherige Untersuchungen haben immer wieder gezeigt, dass das komplizierte tribologische System Prozess/Produktionswerkzeug eine realitätskonforme Ursachenforschung unter Laborbedingungen mit vertretbarem Aufwand und akzeptabler Aussagewahrscheinlichkeit nicht zulässt. Grund dafür ist insbesondere, dass unter den Bedingungen der betrieblichen Praxis (diskrete/zufällige) Ereignisse/Anomalien auftreten, die unter Laborbedingungen nicht nachzustellen sind. Daher wurden die folgenden Untersuchungen mit Industriepartnern mit realen Produktions-/Werkzeugdaten durchgeführt.

# EMuDig4.0

In einem ersten Schritt sollen alle für die Werkzeugstandmenge und deren Steuerung relevanten Daten und Informationen in einem **digitalen Werkzeugpass (4.1)** zusammengefasst werden. Dies bildet die Grundlage um strukturiert Abhängigkeiten und/oder Muster erkennen zu können.

Eine Steigerung der Werkzeugstandmenge soll durch Reduzierung der heute sehr großen Standmengensteuerung erreicht werden. Ein wesentlicher Baustein hierfür ist die Implementierung einer Methode zur **objektiven Beurteilung des Standmengenendes (4.2)**.

Ein **adaptives Prognosemodell (4.3)**, mit dem unter gegebenen Produktionsbedingungen die Standmenge / Reststandmenge des Werkzeugs vorherbestimmt werden kann, dient der präzisieren Kalkulation und einer besseren Planbarkeit der Produktionsabläufe.

## 4.1 Digitaler Werkzeugpass

Für die Werkzeugstandmenge und deren Steuerung sind eine Vielzahl von Parametern verantwortlich. Diese umfassen die Bereiche Herstellung des Werkzeugmaterials, Werkzeugherstellung sowie die Produktion der Schmiedeteile. In Abbildung 4.1 sind ein beispielhafter Prozess und ausgewählte Parameter, die für die Werkzeugstandmenge von Bedeutung sein können, dargestellt.

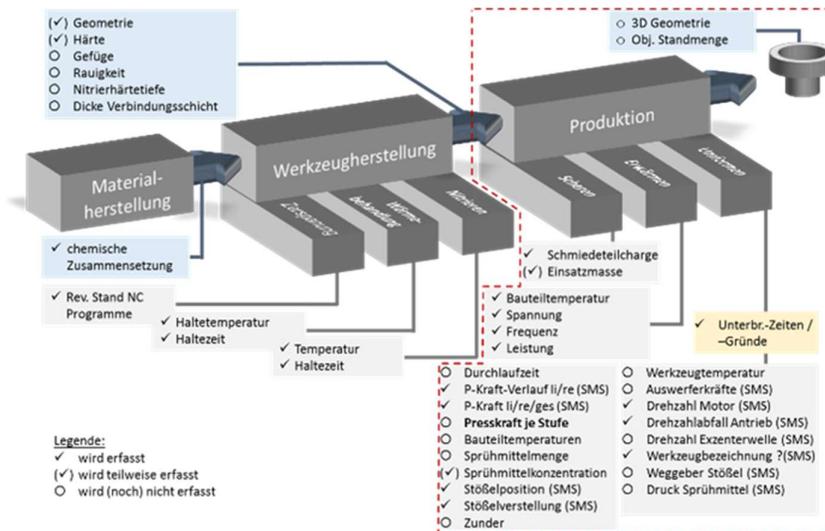


Abbildung 4.1: Beispielhafter Parameterraum

In Abbildung 4.2 ist der prinzipielle Workflow für den digitalen Werkzeugpass dargestellt. Bei der Einführung eines digitalen Werkzeugpasses ist als erstes zu klären welche Parameter berücksichtigt werden sollen. Anschließend muss geprüft werden ob diese Parameter schon erfasst werden. Für Daten, die noch keiner Erfassung unterliegen, sind neue Sensoren (z.B. bei physikalischen Größen) oder neue Verfahren zu implementieren (z.B. bei organisatorischen Informationen).

# EMuDig4.0

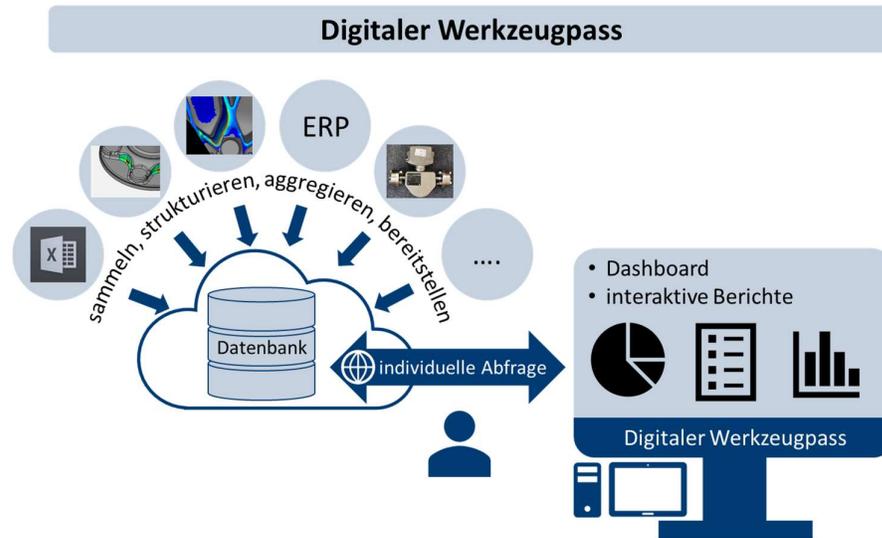


Abbildung 4.2: Workflow „Digitaler Werkzeugpass“

Die erforderlichen Daten stehen in der Regel in unterschiedlichen Formaten / Systemen zur Verfügung, die in einem einheitlichen System aggregiert werden müssen (z.B. PostgreSQL). Spätestens hierbei sollten die Daten auf die notwendige Information reduziert werden.

Eine erste Datenauswertung kann dann über eine Visualisierung mit einem Dashboard erfolgen. Im Rahmen des Teilprojektes wurde hierfür PowerBI von Microsoft verwendet. Hiermit können auf einfache Weise individuelle Berichte erzeugt werden. Typische Fragestellungen auf dieser Auswerteebene sind:

- Mittlere Standmenge von Werkzeug x unter Berücksichtigung des Ausbaugrundes
- Standmengenstreuung von Werkzeug x
- Anzahl Werkzeuge pro Zeitintervall
- Härtewerte der ausgewählten Werkzeuge

In Abbildung 4.3 ist beispielhaft ein im Projekt entstandenes Dashboard dargestellt.

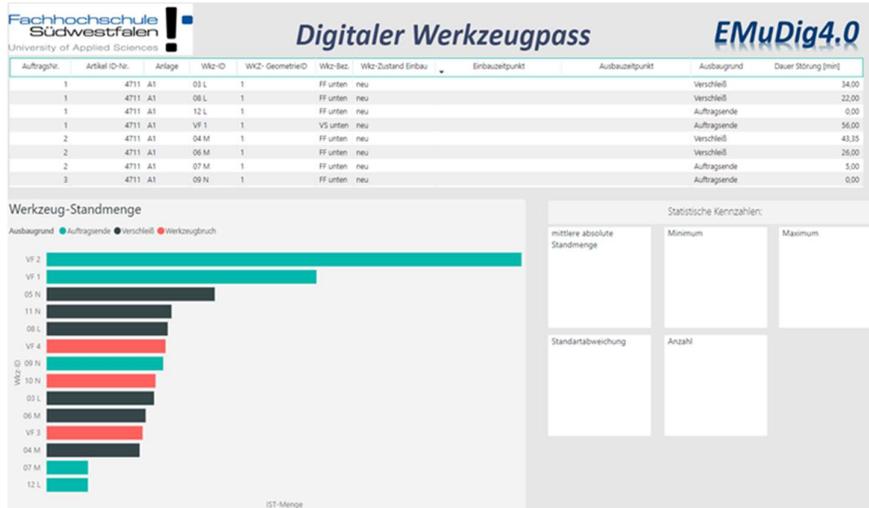


Abbildung 4.3: Dashboard zum digitalen Werkzeugpass (Beispiel, keine realen Daten)

Derartige Auswertungen können dann auf unterschiedlichen Plattformen (PC, mobile Endgeräte) erstellt und/oder abgerufen werden. Die strukturiert abgelegten Daten stehen anschließend auch für tiefer greifende Analysen mit Werkzeugen der Data Analytic zur Verfügung.

## 4.2 Objektivierung Standmengenende

Die Beurteilung ob ein Umformwerkzeug sein Standmengenende erreicht hat wird häufig vom Bedienpersonal durch visuelle Begutachtung vorgenommen. Diese subjektive Vorgehensweise ist ein wesentlicher Faktor für die hohe Standmengenstreuung von Umformwerkzeugen. Dies wurde auch im Rahmen einer Studie<sup>1</sup> von mehreren Unternehmen aus der Massivumformung bestätigt. Aus diesem Grund sollte die Entscheidung bezüglich des Standmengenendes objektiviert werden. Das folgende Bild stellt hierzu die Vorgehensweise dar.

<sup>1</sup> Herbertz, Rainer; Labs, Rainer (2013): Entwicklung eines Konzeptes zur Korrelation Lebensgeschichte / Standmenge für warmProduktionswerkzeuge, Studie im Auftrag des Industrieverband Massivumformung e.V., Hagen

# EMuDig4.0

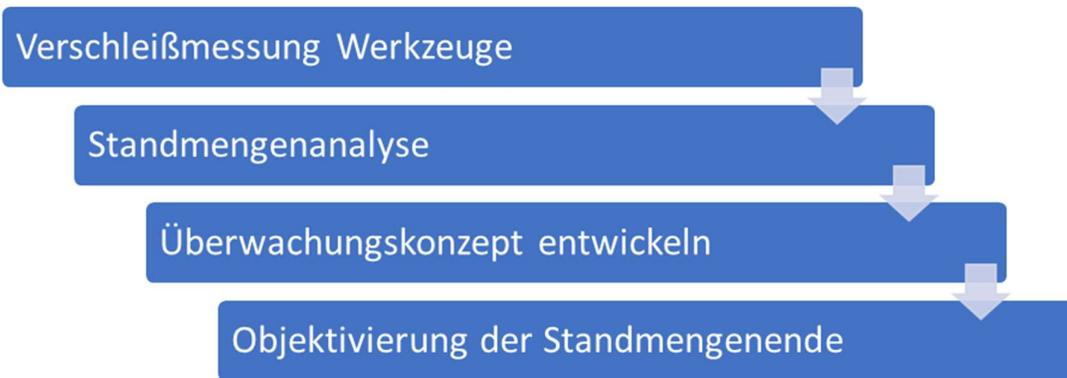


Abbildung 4.4: Vorgehensweise bei der Objektivierung des Standmengenendes

## 4.2.1 Verschleißmessung

Die Messung des Werkzeugverschleißes ist eine wichtige Voraussetzung für die objektive Beurteilung des Standmengenendes. Die Ermittlung des Werkzeugverschleißes erfolgt indirekt über die Vermessung von Schmiedeteilen.

In einem ersten Schritt sind die Verschleißkriterien zu spezifizieren. Typische Verschleißbilder sind Riefen oder Auswaschungen von Radien. Für diese Kriterien werden dann Grenzwerte benötigt. Sind die initialen Punkte festgelegt muss für die entsprechenden Merkmale eine Messstrategie entwickelt werden.

Im Projekt wurden dem Prozess in regelmäßigen Abständen Bauteile entnommen, die nach der Abkühlung und einem Reinigungsstrahlen mit einem 3D Scanner (s. Abbildung 4.5) vermessen wurden.



Abbildung 4.5: 3D Digitalisierungssystem

Alternativ ist auch die Vermessung der Bauteile mit Linienscannern möglich. Neueste Untersuchungen zeigen, dass dies mittlerweile auch schon im warmen Zustand möglich ist.

## 4.2.3 Standmengenanalyse

Das Verschleißkriterium im Projekt war ein Radius, der bei fortschreitendem Verschleißverlauf ausgewaschen wird.

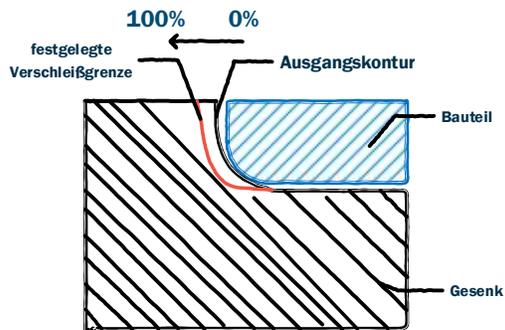


Abbildung 4.6: Verschleißkriterium „Auswaschung Radius“

Im Projektverlauf wurden insgesamt 16 Werkzeuge (Fertigform unten), die auf Grund von Verschleiß ausgefallen sind, genauer analysiert.

Für diese Werkzeuge wurden zur Identifikation des Standmengenendes die üblichen Begutachtungsmethoden angewendet. Somit ist die Standmenge dieser Werkzeuge subjektiv beeinflusst. Um diesen Effekt zu quantifizieren wurde jeweils das letzte gefertigte Bauteil digitalisiert.

Die Ergebnisse sind in Abbildung 4.7 dargestellt. Es zeigt sich, dass etwa 16 % der Standmengenschwankung durch die subjektive Beurteilung des Standmengenendes begründet ist.

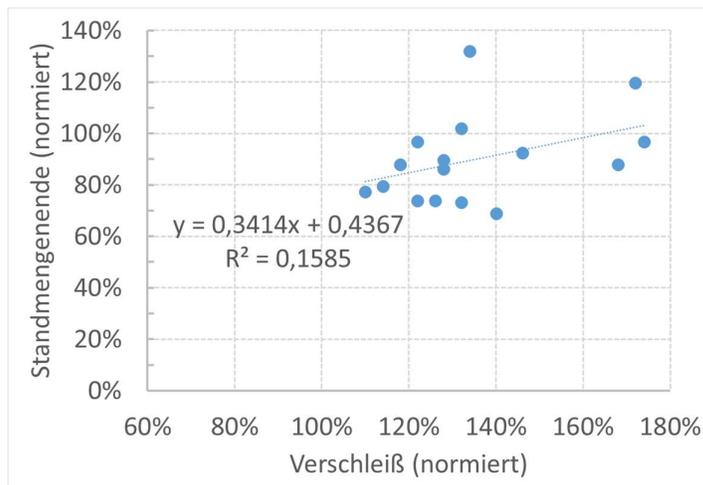


Abbildung 4.7: Maximaler relativer Werkzeugverschleiß von 16 Werkzeugen (Ausbaugrund: Verschleiß)

Wird die Streuung bei der Beurteilung des Standmengenendes (PP) reduziert (z.B. durch Einführung einer Methodik zur Objektivierung des Standmengenendes), führt dies zu einer Verschiebung der Mittellage

# EMuDig4.0

der Standmengenverteilung (SM) (s. Abbildung 4.8). D.h. durch eine Reduzierung der Standmengenstreuung wird die mittlere absolute Standmenge gesteigert!

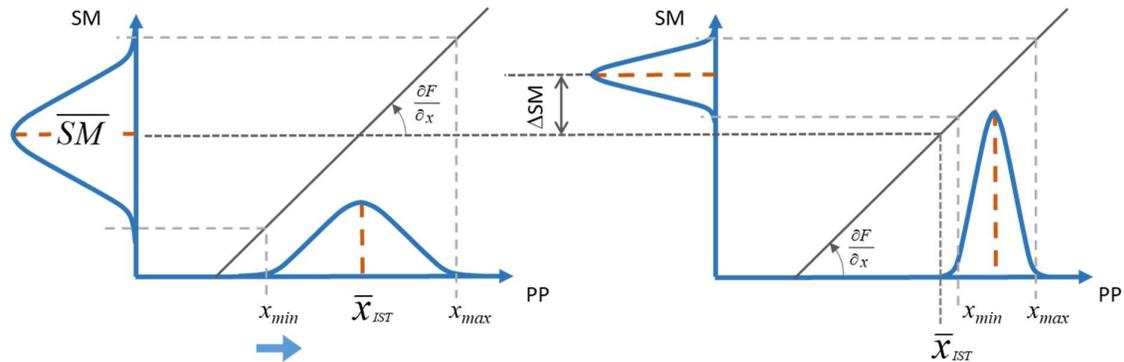


Abbildung 4.8: Standmengensteigerung durch Reduzierung der Standmengenstreuung

In dem untersuchten Fall betrug die mittlere Standmenge 90% bei einem mittleren Verschleiß von 135% (s. Abbildung 4.9). Würde der maximal festgestellte Verschleiß (174%) als neuer Grenzwert definiert und die Streuung der Standmengenbeurteilung eliminiert (d.h. alle Werkzeuge würden bei gemessenen 174% Verschleiß ausgebaut) würde sich die mittlere Standmenge auf 103% erhöhen (Delta 13%).

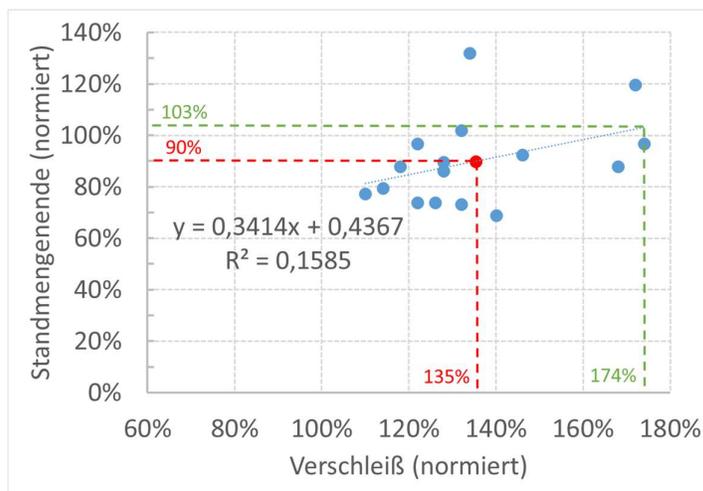


Abbildung 4.9: Potential der Objektivierung des Standmengenendes

## 4.2.4 Überwachungskonzept entwickeln

Zur Identifikation des richtigen Zeitpunktes des Werkzeugausbaus wurde eine Methodik zur Werkzeugüberwachung entwickelt. Diese Methodik sieht vor, dass nach etwa 50% der erwarteten Werkzeugstandmenge ein Bauteil dem Prozess entnommen und vermessen wird.

# EMuDig4.0

Anhand der ermittelten absoluten Verschleißmaße und unter Nutzung bisheriger Verschleißverläufe wird eine Reststandmenge prognostiziert. Hierzu kann eine Variationsbreite (Minimale / maximale Reststandmenge) angegeben werden (s. Abbildung 4.10).

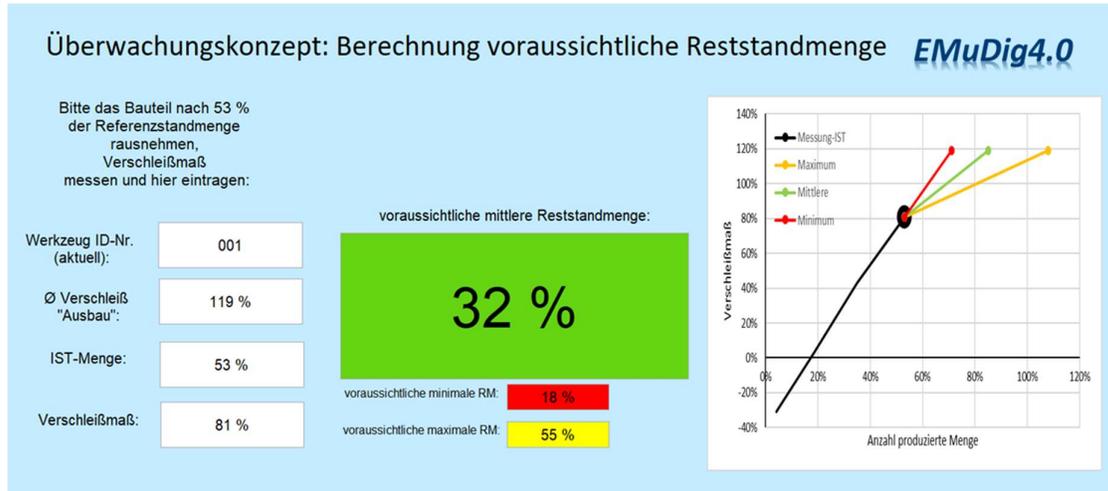


Abbildung 4.10: Werkzeugüberwachung und Prognose der Reststandmenge

Zum Zeitpunkt der minimal prognostizierten Reststandmenge ist ein weiteres Bauteil dem Prozess zu entnehmen und zu vermessen. In Abhängigkeit der gewünschten Genauigkeit und des notwendigen Aufwands kann dieses Vorgehen weiter wiederholt werden. Das letzte gefertigte Bauteil ist auf jeden Fall zu vermessen, um hierdurch die Prognosen der Reststandmenge zu verbessern.

Diese Methodik kann Bestandteil des digitalen Werkzeugpasses sein.

## 4.3 Standmengenprognose

Ermüdungsrisse / Brüche sind nach abrasivem Verschleiß der zweithäufigste Ausfallgrund von Werkzeugen der Massivumformung. Ungeplante Werkzeugausfälle kosten Produktionszeit und können bei der Kostenkalkulation nur unzureichend berücksichtigt werden. Hierdurch entsteht die Gefahr, dass man entweder zu teuer anbietet (und den Auftrag wahrscheinlich nicht erhält) oder aber zu preiswert anbietet und so wirtschaftliche Verluste entstehen können. Aus diesem Grund soll durch eine adaptive Standmengenprognose eine zuverlässigere Kalkulation möglich werden.

Durch das entwickelte adaptive Prognosemodell wird die zu erwartende Gesamthubzahl ermittelt (siehe Abbildung 4.11), aus der dann die Werkzeugstandmenge abgeleitet werden kann (Zur Fertigung eines Bauteils werden je nach Prozesssituation 2-3 Hübe benötigt). Das Prognosemodell kann bereits in der Kalkulationsphase eingesetzt werden, da

- die Dehnungswöhlerlinie werkstoffabhängig ist. Die Werkzeugwerkstoffe beschränken sich in einem Unternehmen aber auf eine überschaubare Anzahl von Standardwerkstoffen. Die Dehnungswöhlerlinien müssen einmal erstellt werden und stehen dann auch für zukünftige Kalkulationen zur Verfügung.

# EMuDig4.0

- die maximale Dehnungsamplitude bauteilspezifisch ist. Sie kann daher über FEM Simulationen vor Angebotsabgabe ermittelt werden.

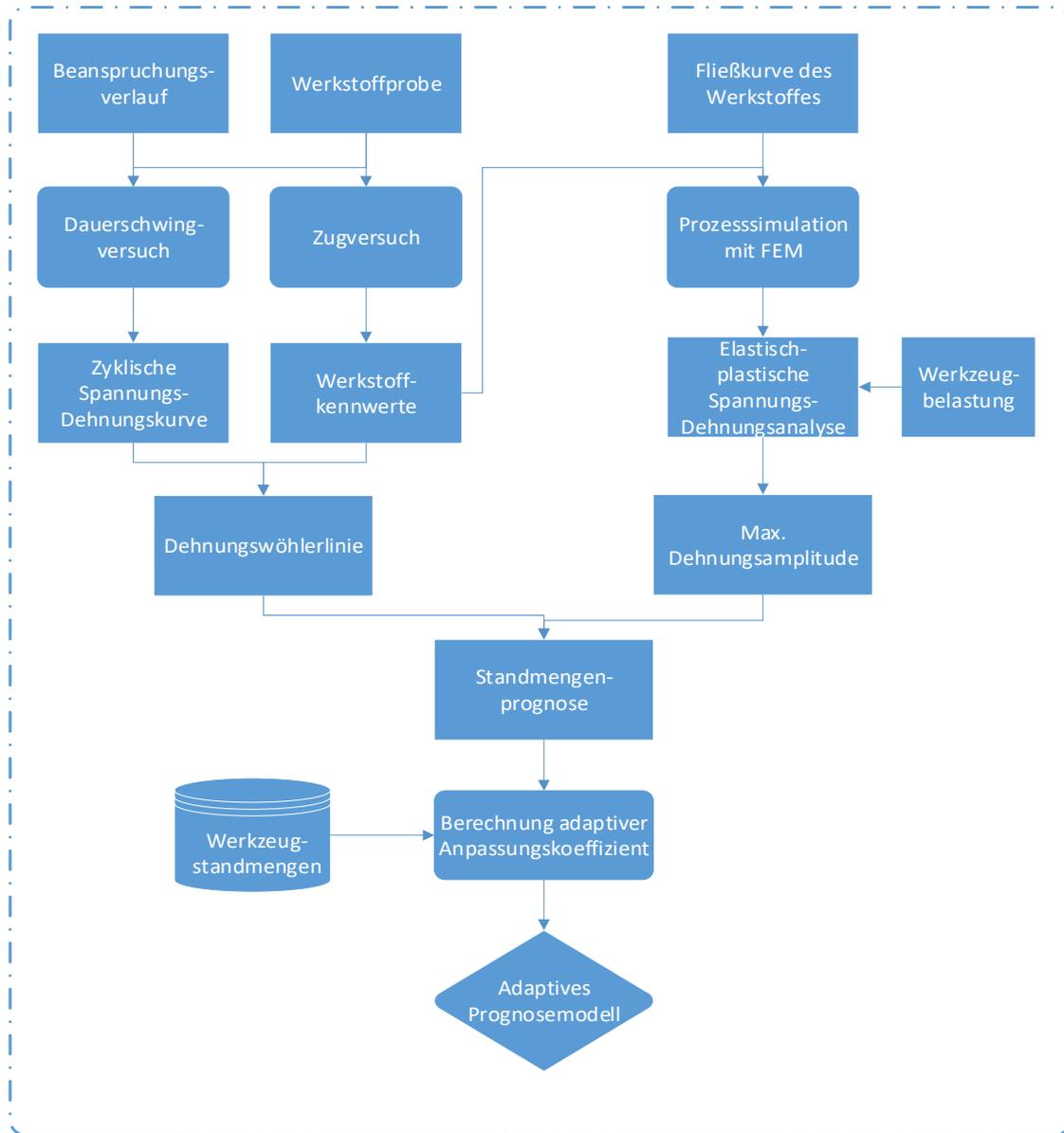


Abbildung 4.11: Ablaufdiagramm zur Erstellung eines adaptiven Prognosemodells

Folgende Voraussetzungen müssen für die Anwendung des hier beschriebenen Prognosemodells gegeben sein:

1. Die Umformkraft ist konstant
2. Die lokale Temperatur an der rissgefährdeten Stelle ist bekannt
3. Es liegt eine Werkstoff-/Probewöhlerlinie vor, bestenfalls für die lokale Werkzeugtemperatur
4. Die lokalen Spannungen und Spannungsgradienten können mittels FEM richtig ermittelt werden

## 4.3.1 Ermittlung der Dehnungswöhlerlinie

Die zyklische Spannungs-Dehnungskurve (ZSD-Kurve) bildet eine der wichtigsten Darstellungen zur Beschreibung des Verhaltens eines Metalls oder einer Legierung bei Wechselverformung<sup>2</sup>.

Umformwerkzeuge versagen im Allgemeinen bereits im Kurzzeit- / Zeitfestigkeitsbereich (LCF – Low Cycle Fatigue)

In der Vergangenheit wurden verschiedene Algorithmen entwickelt, um ZSD Kurven unter Nutzung statischer Werkstoffkennwerte für den LCF Bereich zu berechnen<sup>3</sup>. Ein Vergleich von Algorithmen die für den hier vorliegenden Anwendungsfall geeignet erschienen, hat eine Spannweite der prognostizierten Belastungszyklen von 10 bis 5000 ergeben. Dieses Ergebnis macht deutlich, dass derartige Verfahren nicht für eine hinreichend genaue Prognose von Werkzeugstandmengen geeignet sind.

Im experimentellen Bereich existieren verschiedene Methoden für die Ermittlung der ZSD-Kurve. Der Versuchsaufwand dieser Methoden ist allerdings sehr unterschiedlich. Für das durchgeführte Vorhaben wurde der Incremental Step Test ausgewählt, da

- der Versuchsaufwand und somit die entstehenden Kosten für eine Dehnungswöhlerlinie relativ gering sind.
- bereits gute Ergebnisse im Bereich von Umformwerkzeuge erzielt wurden

In Abbildung 4.12 ist der prinzipielle Versuchsablauf bei der Durchführung des Incremental Step Test dargestellt. Da die Umformwerkzeuge im Betrieb warm werden, müssen die Versuche bei unterschiedlichen Temperaturen durchgeführt werden.

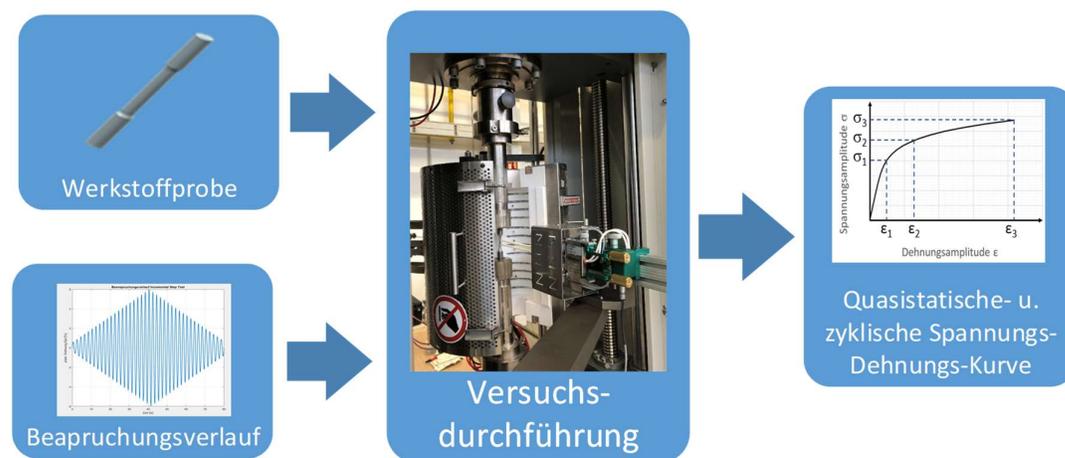


Abbildung 4.12: Incremental Step Test

<sup>2</sup> Christ, Hans Jürgen (1991): Wechselverformung von Metallen. WFT Werkstoff-Forschung und -Technik, vol 9. Springer, Berlin, Heidelberg

<sup>3</sup> Wächter, Michael; Zur Ermittlung von zyklischen Werkstoffkennwerten und Schädigungsparameterwöhlerlinien; Dissertation 2016; TU Clausthal

# EMuDig4.0

Aus Zugversuch und zyklischer Spannungs-Dehnungskurve wird dann die Dehnungswöhlerlinie ermittelt (s. Abbildung 4.13)

$$\varepsilon = \varepsilon_{el} + \varepsilon_{pl} = \frac{\sigma'_f}{E} (2N)^b + \varepsilon'_f * (2N)^c$$

$$\sigma'_f = R_m(1 + A_g) \quad \text{zyklischer Schwingfestigkeitskoeffizient}$$

$$b = \frac{-n'}{1 + 5n'} \quad \text{zyklischer Schwingfestigkeitsexponent (-0,05 bis -0,12 nach Radaj u. Vormwald)}$$

$$\varepsilon'_f = \ln(1+A) \quad \text{zyklischer Duktilitätskoeffizient}$$

$$c = \frac{-1}{1 + 5n'} \quad \text{zyklischer Duktilitätsexponent (-0,5 bis -0,7 nach Radaj u. Vormwald)}$$

Kennwert	Bedeutung	Quelle
R <sub>m</sub>	Zugfestigkeit	Zugversuch
A <sub>g</sub>	Gleichmaßdehnung	Zugversuch
N	Anrisspielzahl	
A	Bruchdehnung	Zugversuch
n'	zyklischer Verfestigungskoeffizient	zyklische σ - ε Kurve

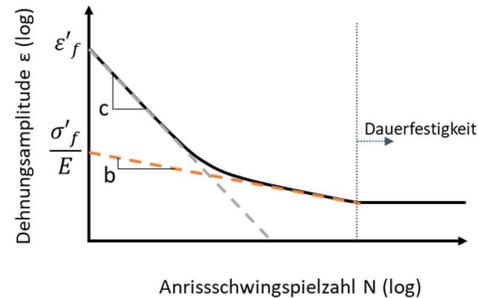


Abbildung 4.13: Ermittlung der Dehnungswöhlerlinie

Prinzipiell kann so mit nur einer Probe einer Dehnungswöhlerlinie ermittelt werden. Bei den durchgeführten Untersuchungen wurden Dehnungswöhlerlinien bei 2 Temperaturen und jeweils 6-facher Belegung erstellt.

Alle 10 Dehnungswöhlerlinien werden als Mittellinie zusammengefasst, da ein signifikanter Einfluss der Temperatur nicht nachweisbar ist.

In einem erstellten Tool können die relevanten Werkzeugdaten eingetragen werden. Als Ergebnis erhält man einen prognostizierten Standmengenbereich (min/max) in dem die tatsächliche Werkzeugstandmenge mit einer statistischen Wahrscheinlichkeit von 99% liegen wird.

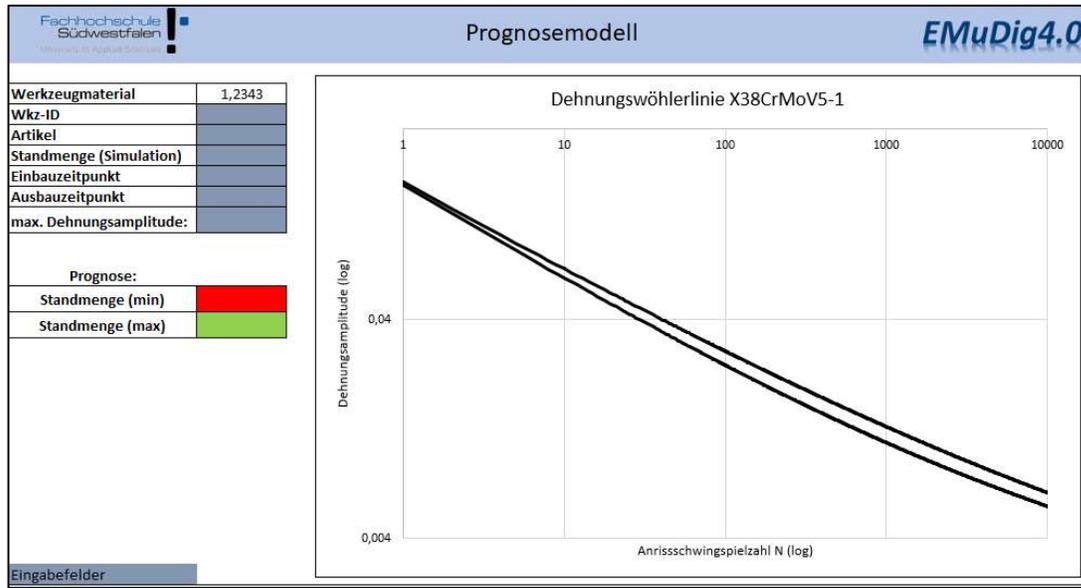


Abbildung 4.14: Prognosemodell zur Standmengenvorhersage

Durch das Erfassen der tatsächlichen Werkzeugstandmengen kann das Prognosemodell durch Veränderung der zyklischen Verfestigungskoeffizienten angepasst werden.

## 5 Kurzzeitanalyse

### 5.1 Steuerungsbasierte Datenerfassung und -verarbeitung

#### 5.1.1 Zuordnung

Für eine datengetriebene Optimierung der Produkt- und Prozessqualität ist die sichere Zuordnung von Prozess- und Bauteileigenschaften essentiell. Bei der Rückverfolgung einzelner Werkstücke oder Chargen entlang von Prozessketten stellt die Erfassung und Zuordnung von Prozessdaten zu den entsprechenden Losen eine große Herausforderung in der industriellen Umsetzung dar. Die Grundlage für eine durchgängige Rückverfolgung eines möglichst kleinen Fertigungsloses ist daher die Integration einer geeigneten Infrastruktur, zum Beispiel einer Smart-Tag-Lösung, die für den Einsatz unter den zuvor skizzierten Umgebungsbedingungen geeignet ist. Anforderungen ergeben sich in besonderem Maß durch die hohen Prozesstemperaturen, die durch die induzierten elektromagnetischen Störfelder eine dezentrale Informationshaltung auf den Smart Tags erschweren. Der zu erfassende Datenraum beinhaltet sowohl die Sensordaten als auch die durch die ausgewählte robuste Identifikation des jeweiligen Bauteils verfügbaren Produktdaten. Um ein durchgängiges, bauteilbezogenes Datenmodell umzusetzen, registriert sich jedes Bauteil vor der Bearbeitung an der jeweiligen Prozessstation. Dies wird durch einen Lesevorgang realisiert. Die im Zuge der Bearbeitung an der Station anfallenden Prozessdaten werden dem Bauteil bzw. dessen im Datenmodell hinterlegter Identifikation direkt zugeordnet. Über den durch die Registrierung erfassten Zeitstempel können Bearbeitungs- und Verwendungszeit von Bauteil und Anlage ermittelt werden. Die Sensordaten werden mit einem Zeitstempel versehen und in das Datenmodell integriert. Auf diese Weise wird sukzessiv und in Echtzeit

## EMuDig4.0

ein für jedes Bauteil individuelles Datenmodell generiert und im Produktionsablauf stetig aktualisiert. Entlang der gesamten Wertschöpfungskette wird das digitale Abbild der einzelnen Werkstücke in prozessabhängigen Zyklen geladen und erweitert. Am Datenmodell ist damit auch der aktuelle Bearbeitungsstand ablesbar.

Zeichnet sich das zu Grunde liegende logistische System weniger durch eine sequentielle Abfolge als eher durch eine komplexe Produktionssystematik aus, ist die jeweilige Ablauflogik entsprechend in Software zu realisieren. Um dies umzusetzen, hilft es, die Ablauflogik zunächst als gerichteten Graphen zu modellieren, der dann in Programmcode überführt werden kann. Das eingangs beschriebene Problem einer fehlenden Zuordnung von Prozessdaten zu Bauteileigenschaften wird dadurch behoben. Für die Konzeption der softwareseitigen Zuordnungslogik ist es wichtig und sinnvoll Ausnahmesituationen wie Ausschussproduktion oder Bedienungsfehler seitens des Werkers zu berücksichtigen. Dies kann bspw. über die Möglichkeit einer binären Eingabe umgesetzt werden, bei der der Werker lediglich einen Knopf zu betätigen hat. In der Software wird dann die entsprechende Ausnahmeroutine eingeleitet, die die Zuordnung der entstehenden Prozessdaten anpasst. Der digitale Zwilling der Werkstücke ist die Grundlage weiterführender datenanalytischer Betrachtungen. Auf der Basis einer robusten Identifikationstechnologie und einer sicheren Zuordnungslogik gemäß der logistischen Abläufe kann eine Rückverfolgung und darauf aufbauende datengetriebene Qualitätsüberwachung konzipiert und umgesetzt werden.

### 5.1.2 Heterogenität

Die Heterogenität der Datenquellen stellt in Bezug auf Datenformate, Protokolle und Abstraten eine weitere Problematik dar. Jedes System entlang der Prozesskette verfügt über eine eigene Steuerung und ein eigenes, autarkes Bussystem, über das Sensorwerte gelesen und Aktoren eingestellt werden können. Um die Heterogenität der Informationsflüsse abstrahieren zu können, wurde eine Art Konnektoren-Ansatz (vergl. Abbildung 5.1) entwickelt, der für jedes dieser Systeme eine individuelle Kommunikationsschnittstelle vorsieht und die Informationen in einheitlicher Form einer höheren Steuereinheit zur Verfügung stellt. Die erfassten Daten können dadurch bspw. einheitlich in die Cloud übertragen werden. Der Konnektor ist somit in einer verteilten Architektur realisiert. Das entwickelte Modul besteht aus einer individuellen Schnittstelle für jede Datenquelle in der Feldebene sowie auf Steuerungsebene und einer standardisierten Schnittstelle zur Weiterleitung der aufgezeichneten Daten an eine übergeordnete Einheit. Die einzelnen, individuell auf die jeweilige Datenquelle angepassten Schnittstellenmodule können auf einzelnen im Feld verteilte Controller oder einer zentralen Speicherprogrammierbaren Steuerung (SPS) laufen, die jeweils eine einheitliche bspw. in der Cloud realisierte Schnittstelle ansprechen. Die für die Steuergeräte entwickelte Datenerfassungskomponente ist als Zustandsautomat realisiert, der durch seinen generischen Aufbau flexibel in bestehenden SPS-Code eingebunden werden kann. Die Implementierung erfolgt als Mealy-Machine, die zyklisch von der SPS aufgerufen wird. Dieser "Read"-Zustand ist je nach Bussystem individuell implementiert. Er tastet die Protokollstruktur ab und trennt Daten und Metadaten. Um die Daten für die weitere Verarbeitung zu erfassen und zu speichern, werden extrahierte Eingangs- und Ausgangssignale auf globale Variablen abgebildet. Die auf diese Weise zyklisch am Bussystem und der entsprechend anliegenden Peripherie abgefragten Sensordaten werden dann von den globalen Variablen auf ein einheitliches REST-Modell abgebildet. Der "Schreib"-Zustand sendet dann zyklisch alle Variablen in die Cloud unter Verwendung des REST-Modells. Der größte Teil der von den Fertigungssystemen aufgezeichneten Prozessdaten sind Zeitreihendaten. Daher werden die

# EMuDig4.0

Sensorwerte konvertiert und auf das REST-Modell abgebildet, das abwechselnd Zeitstempel, Metadaten und die entsprechenden Daten enthält. Aufgrund der Tatsache, dass der größte Teil der Prozessdaten, die an Fertigungssystemen aufgezeichnet werden, Zeitreihendaten sind, werden die gemessenen Prozessparameter in Abhängigkeit der Protokollstruktur konvertiert, auf das Modell abgebildet und an die standardisierte Schnittstelle weitergegeben. Diese Software-Komponente ist generisch und kann wiederverwendet werden. Die standardisierte Schnittstelle, die in der Cloud läuft, leitet die Daten entsprechend der Zykluszeit der SPS (bspw. alle 7ms im Falle der Stuttgarter Modellfabrik am IFU) weiter. Die Zeitspanne zwischen der Durchführung von zwei aufeinander folgenden Prozessschritten schwächt die harten Echtzeitanforderungen ab. In der Cloud findet die weitere Verarbeitung der Daten in einer Echtzeitdatenbank statt, die es ermöglicht Ad-hoc-Berechnungen durchzuführen und Analyseverfahren gemäß dem Online Analytical Processing (OLAP) zu implementieren. Diese Systemkomponente wird für die eindeutige Zuordnung von Werkstücken und Prozessparametern genutzt und ermöglicht je nach produktionstechnischen Umwelt- und Rahmenbedingungen eine Online-Verfolgung von Einzelteilen bzw. Chargen. Eine detaillierte Beschreibung erfolgt in Modul 5.2.

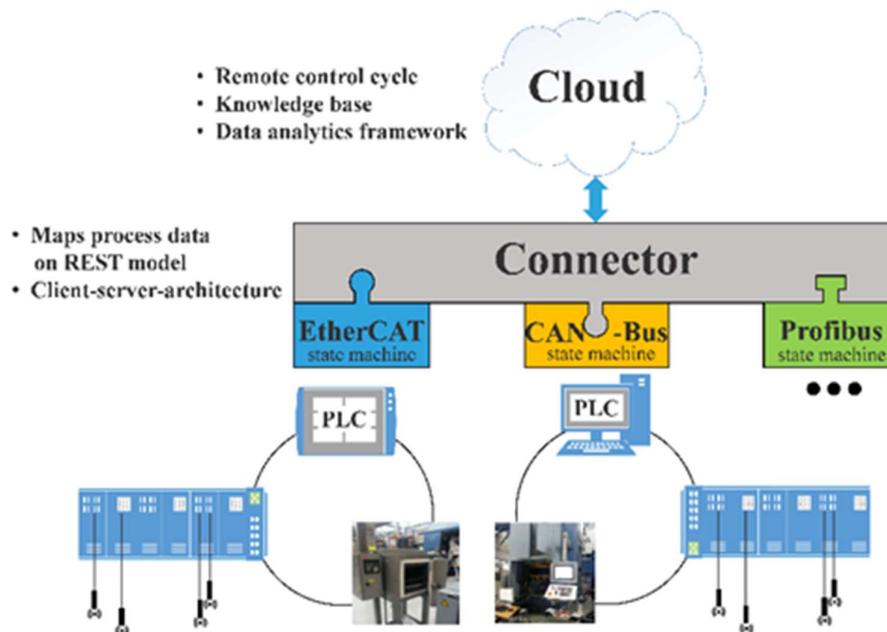


Abbildung 5.1: Darstellung der Datenerfassung und -verarbeitung

Die Übertragung der Daten kann auf verschiedene Arten gesteuert werden. Serverseitig kann über Webservices eine konstant frequentierte Abfrage des zentralen Rechners erfolgen. Alternativ kann clientseitig ein ereignisgesteuerter Datenübertragungsvorgang eingerichtet werden. Die Entscheidung für eine Lösung ist abhängig von den bestehenden Echtzeitanforderungen. Bei der eingangs beschriebenen Prozesskette ist eine zyklische, serverseitige Datenabfrage ausreichend für eine effektive Qualitätsüberwachung. Zusätzlich ermöglicht der Zustandsautomat die Rückführung von Steuerdaten. Nach der Durchführung einer Datenanalyse in der Cloud werden die Ergebnisse in Form von Steuerdaten zurückgeleitet. So ist es möglich, abhängig von den in den vorangegangenen Prozessschritten erfassten, bauteilbezogenen Sensordaten die passenden Stellgrößen für den nächsten Prozessschritt bzw. das nächste Werkstück zu wählen. Eine Anpassung der einzustellenden

# EMuDig4.0

Parameter über mehrere Werkstücke hinweg ist ebenfalls umsetzbar. Die Analyse dient der Adaption des aktuellen Prozessgeschehens und wird für eine situationsbasierte Qualitätsüberwachung genutzt.

## 5.2 Konzept zur Datenmodellierung und Rückverfolgung

### 5.2.1 Extract-Transform-Load

Um ein effiziente Qualitätssicherung realisieren zu können, werden OLAP-Modelle auf Grundlage der Datenräume der einzelnen Prozessschritte erstellt. Dazu wurde ein ETL-Stack (Extract-Transform-Load) implementiert, der die Daten der verschiedenen Datenquellen extrahiert und durch Transformationen in ein strukturiertes Modell überführt, das In-Memory gehalten und konsolidiert wird. So können neben den in Abschnitt 5.1 beschriebenen Prozess- und Produktdaten auch weitere Datenquellen, wie Auftrags- und Unternehmensdaten aus Datenbanken oder ERP-Systemen auf dieser höheren Ebene der Datenverarbeitung integriert werden. Die automatisierte Datenaufbereitung und –verarbeitung ist in Abbildung 5.2 dargestellt.

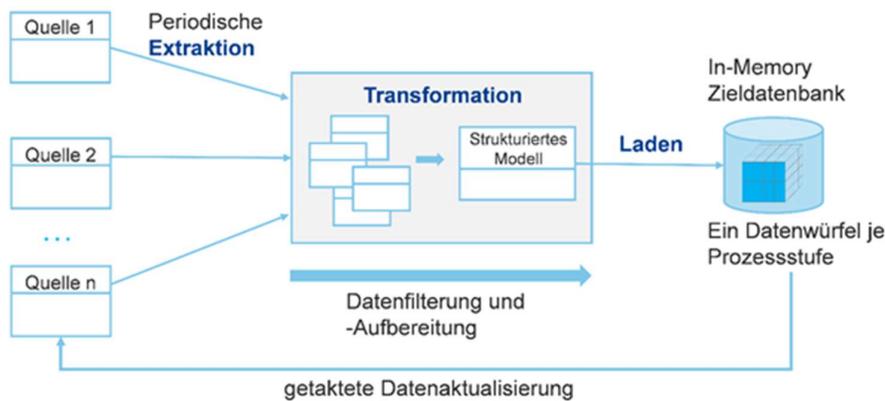


Abbildung 5.2 ETL-basiertes Konzept zur Homogenisierung des Datenraumes

### 5.2.2 Multidimensionale Datenmodelle

Zur multidimensionalen Modellierung der Daten wird im vorliegenden Fall eine Echtzeitdatenbank genutzt, die die Daten In-Memory verarbeitet und Online Analytical Processing (OLAP) ermöglicht. OLAP bietet die Möglichkeit, Daten in einer multidimensionalen Struktur zu modellieren und zu analysieren. Dabei ist die multidimensionale Betrachtung, wie eingangs bereits angedeutet, unabhängig von der Datenquelle und der Datenstruktur dieser Quelle. So ist es möglich, OLAP auf verschiedenste, heterogene Datenquellen anzuwenden. Sowohl konventionelle relationale Datenbanken, MES- und ERP-Systeme als auch diverse Dateiformate und Webservices können in eine Modellierung einbezogen werden. Die multidimensionale Datenmodellierung ist unabhängig von der tatsächlichen physikalischen Speicherstruktur der aus verschiedenen Quellen integrierten Daten. Die integrierten Daten können zum einen relational abgelegt werden. In diesem Fall wäre nur die

## EMuDig4.0

Modellierung multidimensional. Zum anderen kann auch die tatsächliche Speicherung multidimensional erfolgen. Dadurch würden sich zusätzlich zu den Vorteilen der Datenmodellierung noch erhebliche Geschwindigkeitsvorteile ergeben. Die Dimensionen der multidimensionalen Datenstruktur ergeben sich aus den verschiedenen Sichten auf die Daten. An den Verzweigungen jedes Strukturpfads stehen Attribute, die die logische Beschreibung der Daten vornehmen. Die Granularität der Datenbeschreibung steigt mit der Hierarchietiefe. Die Daten selber können für verschiedene Dimensionen in verschiedenen Granularitäten vorliegen, ohne dass dadurch eine homogene Datenmodellierung verhindert wird. So ist es möglich, stark heterogene Datenstrukturen in einer Struktur zu vereinen. Da sich im Fall von drei Dimensionen bei der Datenvisualisierung ein Würfel ergeben würde, hat sich der Begriff des „Data Cube“ etabliert.

Die Modelle werden aus einer Vielzahl an Dimensionen aufgebaut. Dimensionen können bspw. durch die Zeit-, Orts- oder Produktinformationen aufgespannt werden. Diese enthalten eine interne hierarchische Struktur, in der jede Verzweigung durch ein Element beschrieben wird. Basis-Elemente sind Vektoren von Prozessdaten zugeordnet und markieren die innerste Schicht des Datenmodells. Ein Basiselement könnte z.B. die ID 1234 eines Werkstücks sein. Höhere Schichten kapseln kumulierte Informationen der unteren Schichten. Eine höhere Hierarchieebene der Dimension ‚Produktinformation‘ würde bspw. die Chargeninformation beinhalten. Der ETL-Stack bildet neue Daten auf dieses hierarchisch strukturierte Modell ab. Wird bspw. ein neues Produkt eines neuen Produkttypen produziert oder Chargen anstelle von einzelnen Bauteilen verfolgt, können diese Informationen auf die Dimensionsstruktur projiziert und nach einem Abgleich hinzugefügt werden. Diese Projektion bewirkt die Erstellung von Elementen innerhalb der Hierarchie wie zum Beispiel die Erstellung höherer Schichten für die Chargenverfolgung. Folglich können unterschiedliche Prozessdatenstrukturen unterschiedlicher Produkte und Mengen angepasst werden. Die gleiche Flexibilität gilt für jedes andere Element jeder anderen Ebene in der hierarchischen Dimensionsstruktur. Die zugehörigen Daten werden durch die Anwendung geeigneter Aggregationsverfahren berechnet. Somit können heterogene Daten in ein homogenes Modell integriert werden, da es möglich ist, verschiedene Schichten mit unterschiedlichen Granularitäten zu assoziieren. Die Rückverfolgung und Überwachung einer beliebigen Losgröße kann dadurch bereits durch das einfache Navigieren durch verschiedene Ebenen des hierarchischen Metadatenkonstrukts erreicht werden. Der Modellierungsansatz ist in Abbildung 5.3 als Würfel visualisiert.

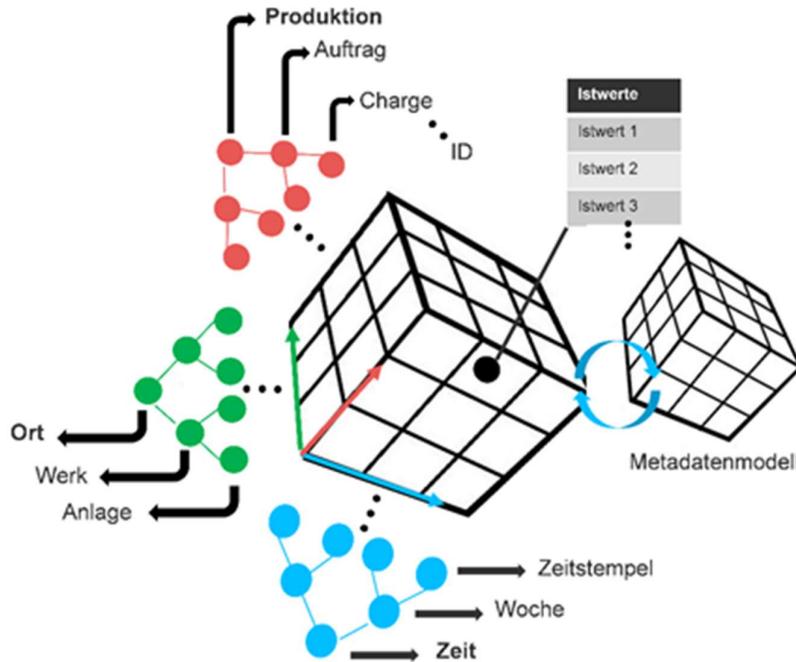


Abbildung 5.3: Multidimensionale Modellierung von Prozessdaten und Bauteileigenschaften

N-dimensionale Modellstrukturen zeichnen sich dadurch aus, dass sie beliebig zusammengesetzt und reduziert werden können. So können die Datenmodelle einzelner Prozessschritte zum Gesamtmodell der Prozesskette komponiert werden. Darüber hinaus werden zu Modellen Metamodelle erzeugt, die das Expertenwissen kapseln indem sie Schwellwerte und Regeln für Prozessparameter beschreiben und in einem formalen Modell abbilden. So kann eine effektive Prozessüberwachung und –diagnose umgesetzt werden. Im Folgenden wird ein Beispiel für die Darstellung von Expertenwissen gegeben: Es sei angenommen, dass die Umgebungsbedingungen entlang der betrachteten Prozesskette zu einer Veränderung hinsichtlich der Qualität des verwendeten Kühlmittels führen, das wiederum die Qualität von Prozess und Produkt beeinflusst. Die Reduktion in der Qualität der Kühlmittel hängt von zahlreichen Faktoren ab, die durch den Experten beschrieben werden können. Dazu gehören unter anderem die Verwendungsdauer (Zeitdimension) sowie der Maschinentyp (Ortsdimension), auf der das Kühlmittel eingesetzt wurde. Folglich können diese Nebenwirkungen formalisiert, auf die Dimensionsstruktur abgebildet und in Form eines Metamodells umgesetzt werden. Eine automatisierte Prozessüberwachung auf Basis von Expertenwissen ist damit möglich.

Das aufgespannte Datenmodell ist die Grundlage für weiterführende Analysen. Im Sinne einer Vorverarbeitung kann eine Selektion bezüglich einzelner oder mehrerer Dimensionen erfolgen. Mit einfachen Operationen kann die hierarchische Struktur auf eine Teilstruktur eingeschränkt werden, um beispielsweise ein bestimmtes Teilproblem zu betrachten. Diese Einschränkungen können flexibel gestaltet und in allen Dimensionen umgesetzt werden. Dadurch ist eine angepasste Betrachtung der Daten und eine darauf zugeschnittene Analyse möglich. Die flexible Datenzusammenführung heterogener Datenstrukturen wird in einem vorverarbeitenden Schritt als Grundlage für die Datenaggregation genutzt. Das beschriebene Konzept soll mit einer tiefgreifenden Datenanalyse kombiniert werden, indem es in ein Analyseframework eingebunden wird. Im Hinblick auf den in der Umformtechnik bestehenden Datenraum ergeben sich durch die Anwendung von OLAP diverse Vorteile. Das vorgestellte Konzept erlaubt es, die Daten der einzelnen Prozessschritte in einem

# EMuDig4.0

separaten oder in einem gemeinsamen Modell zu verwalten. So ist es möglich, auf einzelne Prozessstationen spezifisch zugeschnittene Analysen durchzuführen, aber auch Zusammenhänge entlang der gesamten Prozesskette abzubilden. Durch die flexible Navigation im erstellten Datenmodell können auch bauteil- oder chargenbezogene Analysen realisiert werden. Durch Ad-hoc-Analysen können unverzüglich fehlerbehaftete Verläufe und Anomalien erkannt und Gegenmaßnahmen eingeleitet werden. Abhängig von den Ergebnissen der Analysen werden Steuerdaten generiert, die über die zuvor beschriebene Infrastruktur direkt Einfluss auf das Anlagenverhalten nehmen können. Eine Steuerung aus der Cloud ist dadurch möglich.

## 5.3 Prozessregelung/Mechanismen für die Umsetzung von präventiven Maßnahmen zur Einhaltung der Qualitätsanforderungen

### 5.3.1 Einführung

Es gibt verschiedene gut erforschte Kontrollansätze in der diskreten Fertigung, die auf einzelne Prozesse angewendet werden. Ein wesentlicher Beitrag dieses Teilprojekts ist das Konzept und die Realisierung einer prozessübergreifenden Regelung, die auf eine Kette von miteinander verbundenen Prozessen angewendet wird. Um eine genaue Beeinflussung zu erreichen, werden Expertenwissen und Verfahren der Datenanalyse in einem System im Sinne eines Grey-Box-Ansatzes kombiniert. Dabei wird das Expertenwissen in Form von numerischen Simulationen in die Betrachtung integriert. Neuronale Netzwerke werden verwendet, um neues Wissen aus aufgezeichneten Daten zu extrahieren. Die Regelung wird als bauteilbezogene Regelung („Run-to-Run-Control“) unter Berücksichtigung der noch ausstehenden Prozessschritte für das jeweilige Bauteil realisiert.

### 5.3.2 Ansatz

Der im Folgenden erläuterte Ansatz hat die Prävention von Anomalien zur Reduktion von Qualitätsschwankungen im Sinne einer Regelung der Produktqualität zum Ziel. Dazu soll eine Modifikation der Stellgrößen durch eine Kompensation vorgenommen werden, die adaptiv während des Betriebs durch ein Neuronales Netz generiert wird. Abbildung 5.4 illustriert diese Idee mit einem vereinfachten Blockschaltbild. Es wird davon ausgegangen, dass ein initiales Modellverständnis des Prozesses vorhanden ist oder generiert wurde und damit eine Einstellung des Prozesses durch den Werker oder einen Regler zur Verfügung steht.

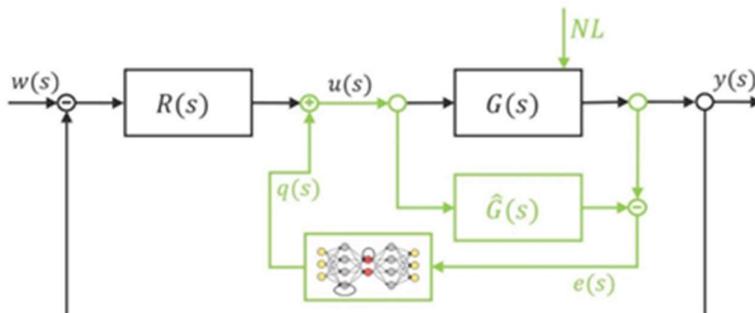


Abbildung 5.4: Vereinfachtes Regelkreisschema zur Regelung der Produktqualität

# EMuDig4.0

Die stabile Einstellung des Prozessmodells durch den Regler wird in der Illustration durch die jeweils schwarzen Koordinatenpunkte repräsentiert. Ein Problem fertigungstechnischer Prozesse hinsichtlich der Modellbildung ist der Umstand, dass das Prozessverhalten örtlich und zeitlich variiert sowie hochgradig nichtlinear ist. Ein auf Basis numerischer Simulationen gewonnenes Modell kapselt nicht alle Zusammenhänge adäquat und weist daher Abweichungen zum realen Prozessverhalten auf. Diese durch Modellunsicherheiten bzw. unbekannte Nichtlinearitäten entstehende Diskrepanz aus gemäß Prozessmodell  $G(s)_{th}$  erwartetem Ausgang  $y_1$  und tatsächlichem Ausgang  $y_2$  wird zur Anpassung der Stellgröße genutzt. Die Stellgröße wird zu  $u_2$  angepasst, sodass auch der reale Prozess den erwarteten Ausgang liefert. Um Anpassungen dieser Art für jedes Bauteil realisieren zu können, wird die gemessene Abweichung  $e(s)$  wie in Abbildung 5.5 zur Generierung des Lernsignals für ein Neuronales Netz genutzt. Dieses Netz wird darauf trainiert, der Stellgröße das Kompensationssignal  $q(s)$  aufzuschalten. In diesem Fall wird nicht das tatsächliche Prozessverhalten  $G(s)_{tat}$  oder dessen Abweichung zum Modell erlernt, sondern die Kompensation  $R(s)_{komp}$ , die aufgebracht werden muss, um eine Abweichung zu vermeiden. Der ursprünglich entworfene Regler bzw. das Verhalten des Werkers  $R(s)_{th}$  kann dadurch unverändert bleiben. Die initial entworfene Regelung wird entsprechend durch die in der Abbildung grün dargestellten Glieder erweitert. Zu verdeutlichen ist an dieser Stelle, dass alle Stellgrößen bauteilbezogen, d.h. mit jedem am Prozess ankommenden Bauteil und damit in diskreter Zeit neu eingestellt werden. Eine Online-Regelung der Qualität während des Prozesses ist nicht vorgesehen. Durch dieses bauteilbezogene Eingreifen sollen bisher unbekannte, die Produktqualität beeinflussende Effekte erlernt und kompensiert

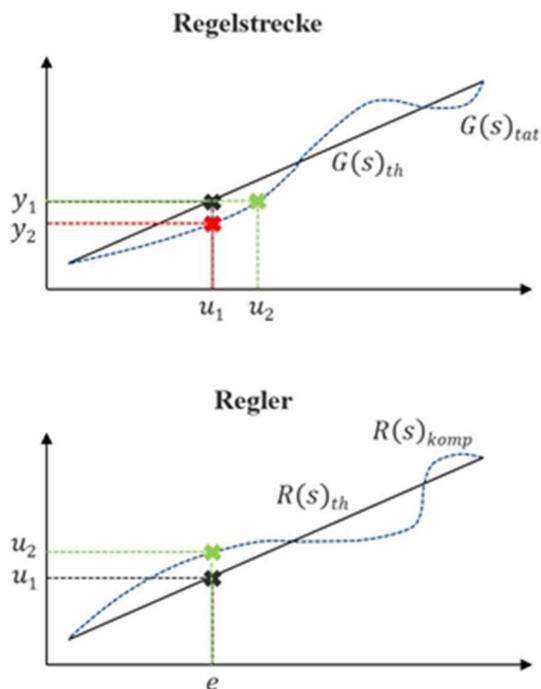


Abbildung 5.5: Vereinfachte Darstellung des grundlegenden Ansatzes zur Kompensation von Anomalien

## 5.4 Anomalie-Charakterisierung

### 5.4.1 Entwickelte Netzarchitektur

Im Folgenden wird die zur Prädiktion und Kompensation entwickelte Neuronale Netzstruktur beschrieben. Zur Erzeugung der Netzeingänge wird eine Diskrete-Wavelet-Transformation (DWT) über die Verläufe der Prozessparameter gebildet, sodass die Daten in einen niederdimensionalen Unterraum projiziert werden. Dies geschieht mit einer definierten Fenstergröße, um eine maximale Reduktion der Datenmenge unter minimalem Informationsverlust zu realisieren. Die hochfrequenten Koeffizienten der DWT werden verworfen, um eine Rauschunterdrückung zu erreichen. Durch diese Abstraktion der Informationen in den Wavelets erhöht sich die Übertragbarkeit des Ansatzes, da die Netzarchitektur mit abstrakten Features trainiert wird. Zur Dimensionsreduktion wird ein Autoencoder genutzt, der die Koeffizienten der DWT als Eingänge  $m$  erhält. Der Autoencoder reduziert diese auf die Anzahl an Neuronen seiner verdeckten Schicht  $n$  und versucht diese anschließend möglichst exakt zu rekonstruieren.

Die auf die Neuronen der verdeckten Schicht reduzierten Features werden zur weiteren Verarbeitung im Sequence-to-Sequence Netz genutzt, das zur Vorhersage eingesetzt wird. Das Netz ist in einen den Eingang verarbeitenden Encoder- und einen den Ausgang erzeugenden Decoder-Teil gegliedert und in Bild 6 dargestellt. Dazu werden zwei Netze mit inverser Struktur aufgebaut, um die Gewichte getrennt voneinander trainieren zu können. Der Encoder-Teil verarbeitet die Eingangsdaten und extrahiert abnormales Prozessverhalten. Um sich das abnormale Verhalten „merken“ zu können, wird über das Netz iteriert, bis die Eingangssequenz vollständig im Encoder-Teil verarbeitet wurde. Die Größe des Netzes ist auf die Größe der Eingabesequenz abgestimmt, um ein möglichst gutes Gedächtnisverhalten zu realisieren. Die Eingangssequenz wird dann in den Decoder-Teil kopiert. Dieser generiert auf Basis des bezüglich der Anomalien erzeugten Gedächtnisverhaltens ein neues Signal zur Vorhersage des erwarteten Ausgangs. Dies ist in der Abbildung anhand eines einfachen Beispiels aus Frage und zu erwartender Antwort dargestellt. Zusätzlich ist an Ein- und Ausgang die Vorhersage von Anomalien am Beispiel realer Zeitreihen zu sehen. Die Grundlage des Netzes sind LSTM Zellen, die in der Abbildung grau angedeutet sind. Sie erzeugen aus den Eingängen einen Zustand, der mit jedem Zeitschritt propagiert wird. Die Vorhersage geschieht über ein definiertes Zeitintervall wie bspw. einen Tag, eine Stunde oder eine Sekunde. Deshalb müssen die Zieldaten den entsprechenden zeitlichen Abstand zu den Eingangsdaten besitzen. Durch wiederholtes Ausführen des Sequence-to-Sequence Netzes kann über größere Zeithorizonte vorhergesagt werden. Es werden die Dimensionsreduktion auf der einen sowie die Unterbindung des Vanishing-Gradient-Effektes auf der anderen Seite in einem Netz zusammengeführt.

# EMuDig4.0

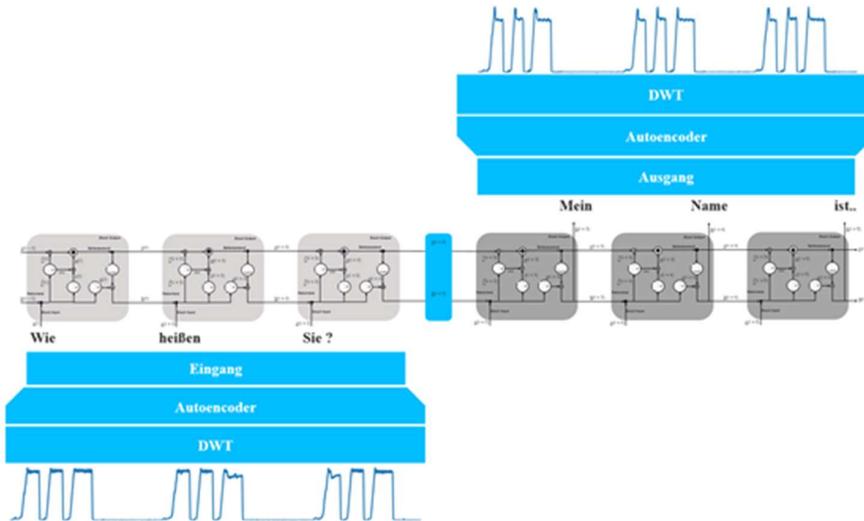


Abbildung 5.6: Sequence-to-Sequence-Netz zur Prädiktion mit reduzierten Features als Ein- und Ausgänge

Um die Systemarchitektur zu testen, wird der gesamte Datenintegrationsprozess zyklisch und GPU-unterstützt ausgeführt, um paralleles Rechnen zu ermöglichen. Dadurch kann eine effiziente Online-Überwachung realisiert werden. Detaillierte Informationen zur verwendeten Hardware sind in Tabelle 5.1 zu finden.

Tabelle 5.1: Verwendete Hardware

OS	GPU	CPU	RAM	Netzwerk
CentOS Server Version 7.3	2x Tesla K80 24 GB GDDR5	18-Core @ Xeon 2.10 GHz	512 GB DDR4 LRDIMM	1 Gigabit Ethernet

Der Einsatz einer leistungsfähigen Grafikkarte ermöglicht neben den Vorteilen für die Datenmodellierung auch eine erhebliche Verkürzung der Trainingsdauer der Netze. Das liegt daran, dass sich der Algorithmus aus einer Vielzahl einfacher und größtenteils unabhängiger Rechenoperationen zusammensetzt, sodass mit Hilfe einer GPU eine effiziente Parallelisierung der Rechenoperationen umgesetzt werden kann. Zu Beginn der Trainingszeit werden die vorverarbeiteten Datensätze in den Arbeitsspeicher geladen. Durch eine Vergrößerung der Merkmalsanzahl konnte der Geschwindigkeitsvorteil weiter erhöht werden. Im Hinblick auf die betrachtete Prozesskette kann die Anzahl der Merkmale jedoch nicht beliebig vergrößert werden, da die Größe des Sliding-Window von dem zu analysierenden Szenario abhängig ist. Darüber hinaus wurde bei einigen Operationen ein starker Performanceeinbruch beobachtet, der auf eine erhöhte Anzahl an Speicherzugriffen zurückzuführen ist.

## 5.5 Darstellung der Assistenzfunktionen und Integration in ein Assistenzsystem

Die in den Modulen 5.1 bis 5.4 beschriebenen Funktionalitäten wurden in Form eines web-basierten Assistenzsystem realisiert, deren wesentliche Ansichten in diesem Modul kurz vorgestellt werden sollen.

### 5.5.1 Systemarchitektur

Der Aufbau der Systemarchitektur ist in Abbildung 5.7 dargestellt. Die Interaktion mit technischen Prozessen erfolgt über eine OPC-UA-Schnittstelle. Daten werden nach einem Verarbeitungsschritt in ein Datenbank-System integriert. Die langfristige Verwaltung der großen Datenmengen erfolgt mittels des verteilten Dateisystems Hadoop

Das Assistenzsystem ist über das in Web-Interface erreichbar. Der Anwender kann dabei auf die drei wesentlichen Funktionen "Datenmonitoring", "Big-Data-Analyse" und "Control" zugreifen. In der Modellfabrik wurde das System als Private Cloud innerhalb der Universität Stuttgart realisiert. Der Zugriff kann sowohl über nahezu beliebige mobile Endgeräte als auch über Smart Glasses erfolgen.

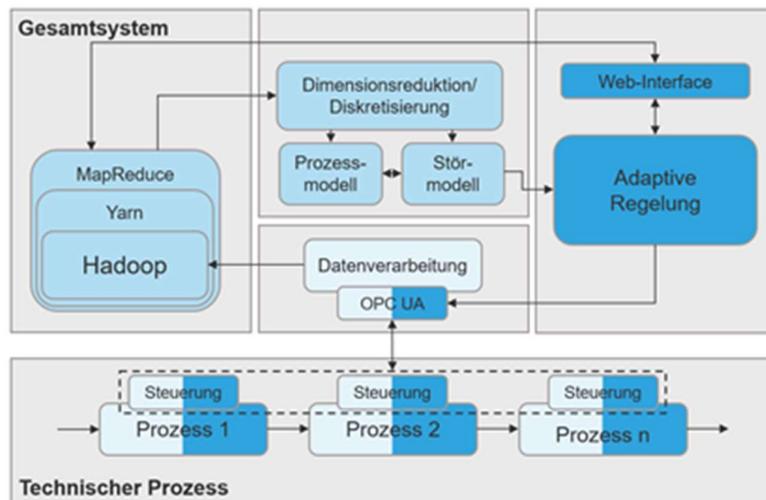


Abbildung 5.7: Systemarchitektur des Gesamtsystems inkl. Datenerfassung, Datenhaltung und Analyse sowie Regelung gekennzeichnet durch abgestufte Blautöne. Die Datenverwaltung erfolgt in Hadoop.

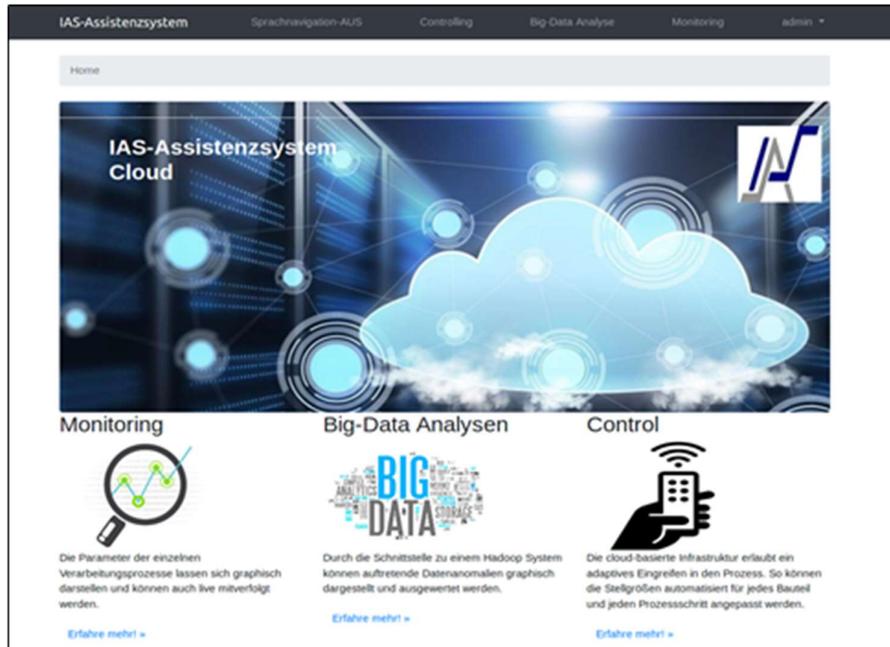


Abbildung 5.8: Startseite des web-basierten Assistenzsystems mit den drei wesentlichen Funktionalitäten zum Monitoring der Daten, zur Analyse und zur Beeinflussung des Prozessgeschehens. Das System ist als Private Cloud innerhalb der Universität Stuttgart realisiert und kann auch von mobilen Endgeräten sowie Smart Glasses abgerufen werden

## 5.5.2 Datenmonitoring

Für dieses Modul werden die Metadaten über das in Abbildung 5.10 dargestellte OPC-UA-Informationsmodell erfasst. Dieses Modell ist als Baumdarstellung in das Web-Interface integriert und stellt die semantische Verknüpfung zu SPS-Steuerungen der angeschlossenen Prozesse her. Im Modul selbst können Prozessdatenverläufe direkt angewählt und dann mit direkter Zuordnung der Kurven zu Bauteilidentifikationen angezeigt werden. Eine exemplarische Ansicht findet sich in Abbildung 5.9.

# EMuDig4.0



Abbildung 5.9: Darstellung ausgewählter Prozessdatenverläufe mit direkter Zuordnung der Kurven zu Bauteilidentifikationen

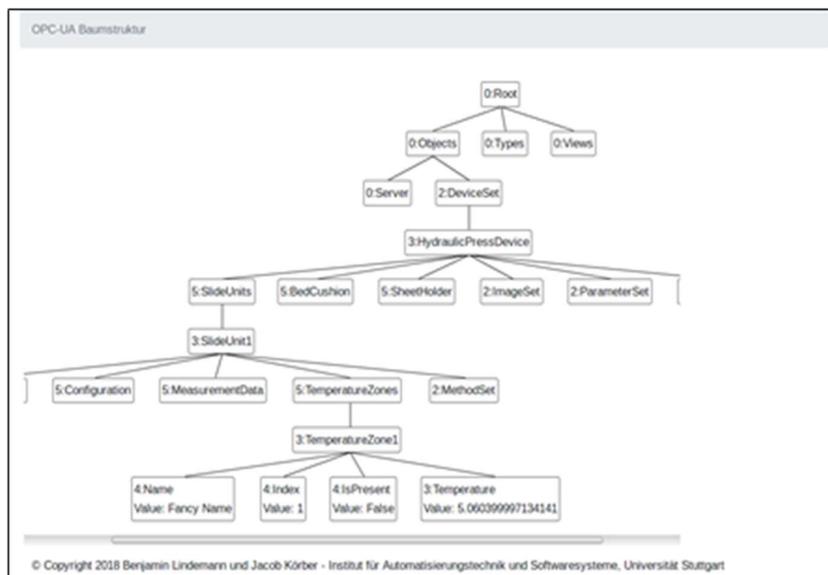


Abbildung 5.10: Darstellung der erfassten Metadaten hier in Form des OPC UA Informationsmodells. Dieses ist als Baumdarstellung in das Web-Interface integriert und stellt so die semantische Verknüpfung zur SPS her.

## 5.5.3 Rückverfolgung

Auch die Rückverfolgung sämtlicher Werkstücke über die ausführlich in Kapitel 2 beschriebenen Kennzeichnungsansätze ist im Webinterface möglich. Die in Abbildung 5.11 dargestellte Ansicht ermöglicht die Nachverfolgung einzelner Bauteile – in diesem Falle des Produkts mit der ID „70“ – durch die komplette Prozesskette. Die Darstellung aktualisiert sich im Betrieb in Echtzeit.

# EMuDig4.0

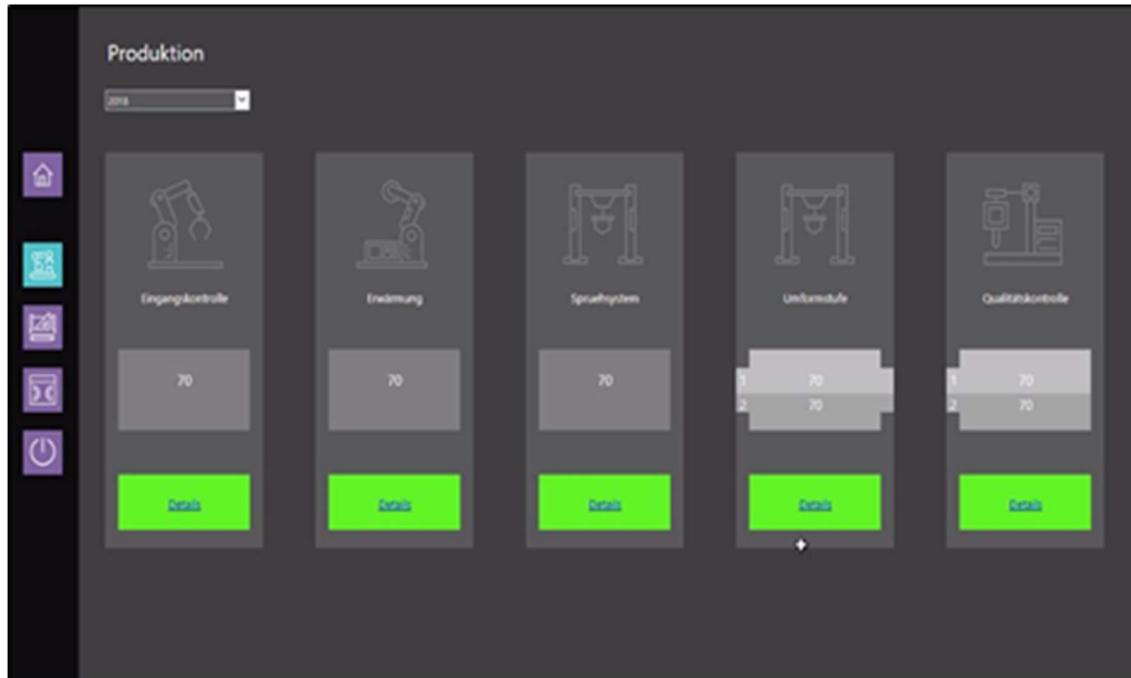


Abbildung 5.11: Bauteilnachverfolgung über einzelne Prozessschritte. Produkt „ID 70“ durchläuft Prozesskette

# EMuDig4.0

Die in Abbildung 5.12 gezeigte Übersicht ermöglicht die schnelle und intuitive Überwachung der Daten. Dafür wird der Status der einzelnen überwachten Chargen oder Werkstücke über eine Ampelcodierung direkt ersichtlich. Für die Bewertung des jeweiligen Status können unterschiedliche Regeln aktiviert werden. Über die zu Grunde liegende Datenmodellierung kann der Benutzer in der Ansicht in den drei Hauptdimensionen „Zeit“, „Ort“ und „Produkt“ navigieren. Das ermöglicht es, einzelne Werkstücke oder Chargen zu bestimmten Zeitpunkten und für bestimmte (Teil-)Prozessschritte zu überwachen und damit die Produktionsqualität insgesamt zu gewährleisten.



Abbildung 5.12:Übersichtsdarstellung mit Ampelcodierung

## 5.5.4 Analyse

Sämtliche Analysewerkzeuge wurden in Tensorflow über Python realisiert. Sie sind über ein in Abbildung 5.13 gezeigtes Dashboard des in Abbildung 5.7 dargestellten Hadoop-Systems direkt im Assistenzsystem verfügbar.

# EMuDig4.0

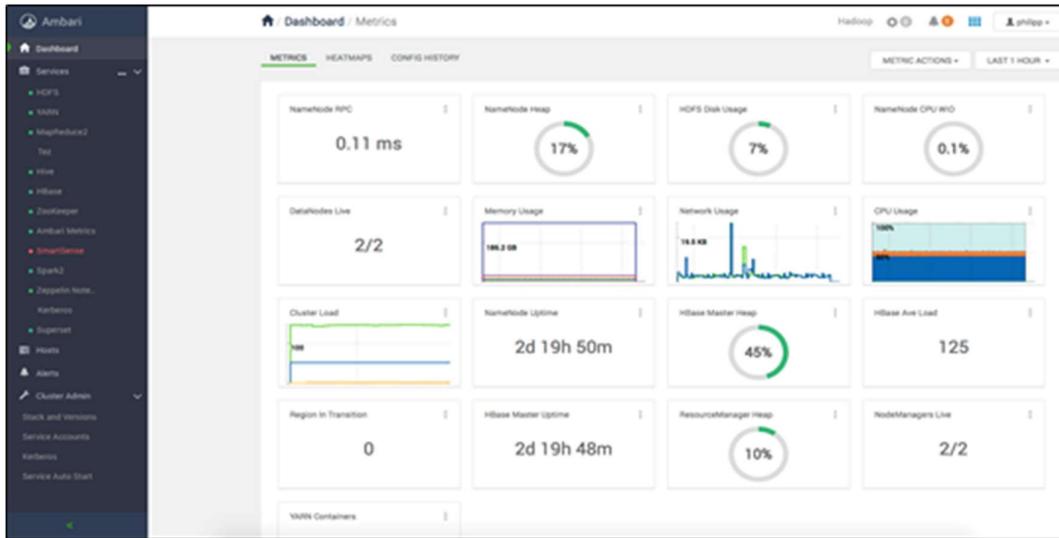


Abbildung 5.13: Dashboard des Hadoop-Systems für Zugriff auf Analysen

Nach der Tensorflow-basierten Datenanalyse erfolgt die Visualisierung über die in Abbildung 5.14 gezeigte Visualisierungs-App von Google. So können Anwendern die Zusammenhänge schon während des Lernvorgangs direkt sichtbar gemacht und visuell zurückgemeldet werden. Die per Klick auswählbaren Daten entsprechen direkt den umformtechnischen Prozessdaten.

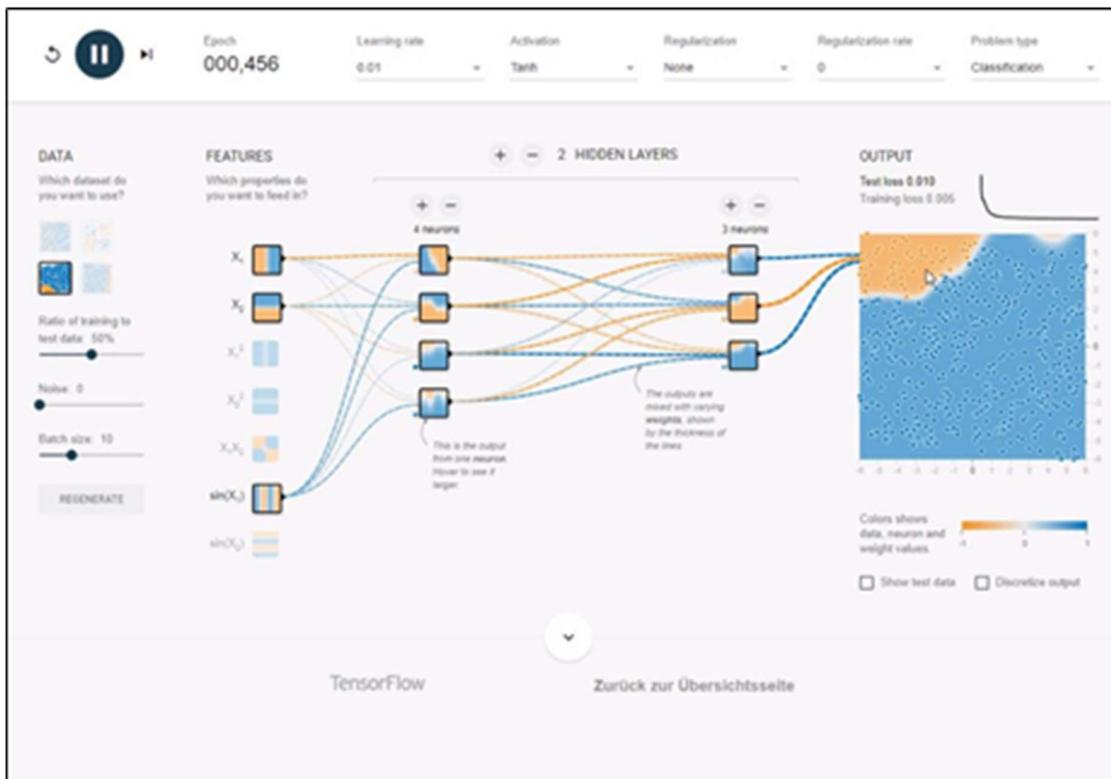


Abbildung 5.14: Visualisierung der übergeordneten Zusammenhänge nach Datenanalyse mittels Tensorflow

# EMuDig4.0

## 5.5.5 Regelung der Qualität

Auch auf sämtliche Stellgrößen kann über das Assistenz-System und das in Abbildung 5.15 exemplarisch dargestellte Interface zugegriffen werden. Allgemeine Handlungsempfehlungen sowie konkrete Empfehlungen zur Einstellung der Stellgrößen, die über die Analyse mittels Neuronaler Netze identifiziert werden, können über Benutzerschnittstellen weitergeleitet werden. Generell kann dabei über eine höhere Entscheidungsebene – in der Praxis: der Werker an der Anlage – Einfluss auf die ermittelten Parameter gewonnen werden. Änderungen der Einstellungen können entweder global oder spezifisch akzeptiert bzw. Verworfen werden. Das ermöglicht eine vollautomatisierte Einflussnahme auf das Prozessgeschehen; aber auch eine Prozessgestaltung mit dem Menschen als letzte Kontrollinstanz.

Die Ergebnisse der Prozessgestaltung können dann über die produzierten Bauteile und Chargen hinweg ebenfalls überwacht und visualisiert werden. Abbildung 5.16 zeigt eine exemplarische Darstellung für das Beispiel der Überwachung der Ausbildung eines Grats.

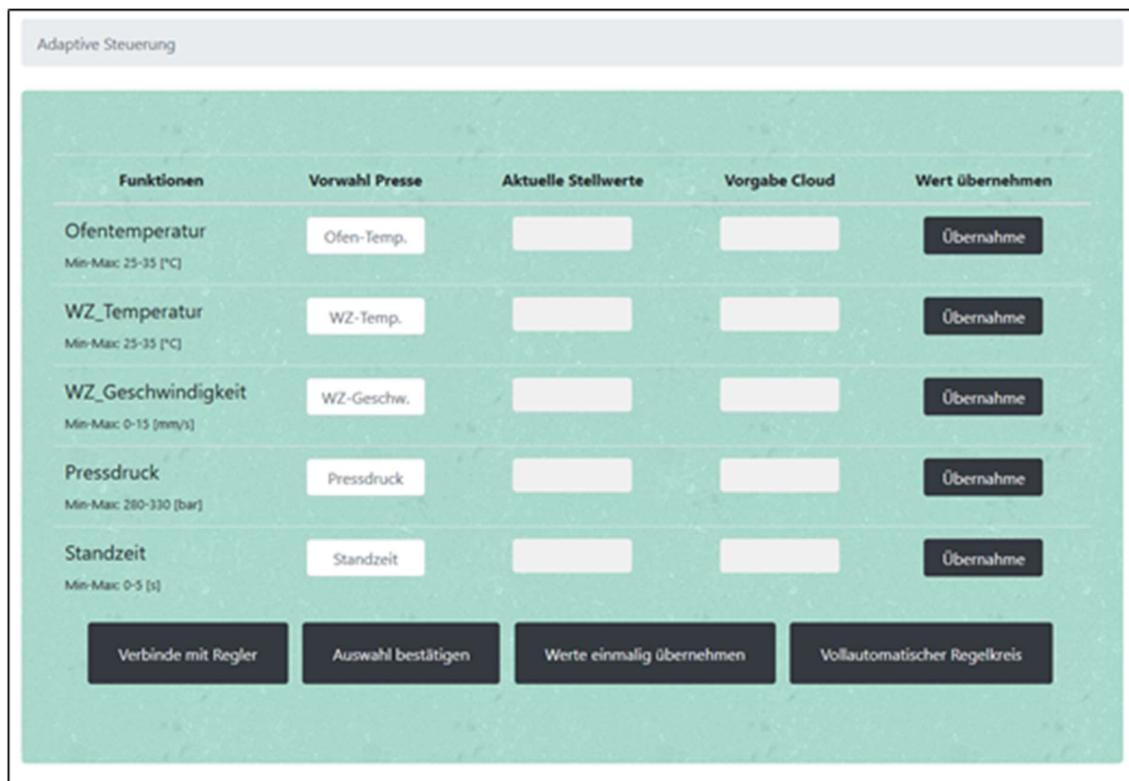


Abbildung 5.15: Ansicht zur Einstellung der relevanten Stellgrößen.

# EMuDig4.0



Abbildung 5.16: Darstellung der Ergebnisse zur Produktqualität der produzierten Bauteile, hier am Beispiel der Überwachung der Ausbildung von Grat

## 6 Langzeitanalyse

### 6.1 Skalierbare Infrastruktur für die Datenanalyse

Um das Verständnis von Produktionsprozessen innerhalb eines Unternehmens zu verbessern, sind detaillierte Analysen von Prozessdaten notwendig. Dafür ist eine leistungsfähige IT-Infrastruktur erforderlich, die es erlaubt, Datenanalysen schnell und effizient durchführen zu können. Da die anfallende Datenmenge von Produkt zu Produkt und von Unternehmen zu Unternehmen variiert, sollte aus Effizienzgründen eine Infrastruktur verwendet werden, die skaliert, also bei einer größeren Datenmenge mehr Rechenleistung, bei einer kleineren Datenmenge weniger Rechenleistung bereitstellt. Eine Cloud-Infrastruktur erfüllt diese Anforderung: Ressourcen in der Cloud können durch Virtualisierung einfach an veränderte Bedürfnisse angepasst werden. Die Integration von Datenanalysemethoden in eine skalierbare Datenanalyse-Infrastruktur erlaubt, das erlernte Wissen in einfacher Weise auch auf andere Prozesse in der Massivumformungsbranche übertragen zu können.

Das folgende Modul des Leitfadens gibt einen Überblick und eine Beschreibung der wichtigsten Schritte beim Aufbau einer skalierbaren Datenanalyse-Infrastruktur als Cloud für die Massivumformungsbranche am Beispiel einer Open-Source Software-Lösung. Dazu gehören:

1. Wahl der Cloud-Infrastruktur (Anmietung oder Aufbau einer eigenen Cloud) zur Bereitstellung einer Umgebung, in der Datenanalyse-Dienste aufgebaut und skalierbar betrieben werden können (Abschnitt 6.1.1),
2. Beschreibung der Architektur der Datenanalyse-Infrastruktur, in der Dienste zur Datenanalyse skalierbar betrieben werden können (Abschnitt 6.2).

#### 6.1.1 Cloud-Infrastruktur

Zu Beginn ist zu entscheiden, ob die Cloud-Infrastruktur angemietet (öffentliche Cloud) oder selbst im Unternehmen (private Cloud) aufgebaut werden soll. Etablierte kommerzielle IaaS-Anbieter (Infrastructure-as-a-Service) sind Amazon mit dem Web-Service „Amazon Elastic Compute Cloud“ (Amazon EC2) [Amazon Web Services (2020)] und Microsoft mit dem Service „Microsoft Azure“ [Microsoft (2020)]. Das Zahlungsmodell ist meist pay-per-use.

Bei der Verwendung kommerzieller Anbieter aus dem Ausland ist zu berücksichtigen, dass die in der Cloud gespeicherten Daten üblicherweise nicht auf deutschen / (EU-)europäischen Servern gehostet werden, sondern auf Servern im nicht-europäischen Ausland. Das dort geltende IT-Recht könnte es erlauben, dass staatliche Organisationen auf die gespeicherten Unternehmensdaten in der Cloud zugreifen dürfen. Es wird daher empfohlen, dass die Cloud-Server im Gültigkeitsbereich der DSGVO gehostet werden und der Netzwerkverkehr vom Unternehmen zum Cloud-Server ausschließlich verschlüsselt läuft.

Bei Nutzung einer Cloud-Infrastruktur entfällt die Wartung der Hardware-Komponenten (z.B. der Austausch bei Defekten). Die Wartung / Administration reduziert sich dann auf die Wartung der in der Cloud laufenden Softwaretools. Generell hängt der Schwierigkeitsgrad der Wartungsarbeiten stark von der Größe der aufgebauten Cloud-Infrastruktur und von den auftretenden Problemen ab. Um ein möglichst reibungsloses Funktionieren der Cloud sicherzustellen, wird der Einsatz eines IT-Administrators empfohlen.

# EMuDig4.0

Kosten für die Unterbringung der Hardware-Komponenten, den Betrieb inklusive Kühlung sind bei einer Anmietung durch die Miete bereits abgegolten. Beim Aufbau der Hardware-Infrastruktur im eigenen Unternehmen müssen entsprechende Personal- und Betriebskosten eingeplant werden.

Generell wird empfohlen, die Datenanalyse-Infrastruktur nicht als öffentliche Cloud, sondern als private Cloud innerhalb des firmeninternen Netzwerks zu betreiben, um unerlaubten Zugriff auf die Daten zu unterbinden und eine volle Kontrolle über die in der Cloud verarbeiteten Daten zu haben. Dies hängt natürlich auch vom Schutzbedarf der Daten ab.

## 6.1.1.1.1 IT-Sicherheit

Damit sichergestellt ist, dass die Cloud-Infrastruktur grundlegende Voraussetzungen an die IT-Sicherheit erfüllt, sollten die vom BSI angegebenen Basischecks zum IT-Grundschutz [Bundesamt für Sicherheit in der Informationstechnik 2020] zu den folgenden Bausteinen / Modulen (Edition 2020) durchgeführt und die darin empfohlenen organisatorischen und technischen Maßnahmen umgesetzt werden.

Alternativ kann die Cloud-Infrastruktur auch nach der Norm ISO/IEC 27001 durch ein Unternehmen zertifiziert werden. Die Norm spezifiziert Anforderungen, um ein Informationssicherheits-Managementssystem einzurichten und umzusetzen. Die Zertifizierung kann jedoch mit nicht unerheblichen Kosten verbunden sein.

## 6.1.1.1.2 Hardware

Für ein stabiles Setup sollte die Cloud-Infrastruktur auf existierenden und etablierten Hardware- und Software-Lösungen aufbauen.

Um die Hardware geeignet zu dimensionieren, ist vorab zu klären, wie groß das zu speichernde Datenvolumen pro Zeiteinheit ist. Als Rechenbeispiel dient dazu im weiteren Verlauf des Leitfadens die folgende Konfiguration: Signale von 100 pro Maschine installierten Sensoren werden für zehn Maschinen in einem Takt von 0.01 Sekunden aufgezeichnet.

Damit aufgezeichnete Werte eindeutig dem Zeitpunkt der Aufzeichnung zugeordnet werden können, sind alle Sensorsignale als Zeitstempel-Wert-Paare aufzuzeichnen und zu speichern. Bei Speicherung des Zeitpunkts als Unix-Zeitstempel (Anzahl der Sekunden seit dem 1. Januar 1970) werden 64 Bit Speicherplatz benötigt, für den aufzuzeichnenden Wert des Sensors weiter 64 Bit bei Speicherung als Gleitkommazahl. Dies führt zu dem folgenden Datenvolumen pro Sekunde:

$$10 \times 2 \times 64 \text{ Bit} / 0.01 \text{ s} \times 100 = 10 \times 128 \times 10.000 \text{ Bit/s} = 1.28 \text{ MBit/s} = 0.16 \text{ MB/s}$$

Zum Aufzeichnen von 100 Signalen werden damit 0.16 Megabyte pro Sekunde benötigt. Dies führt zu einem täglichen Datenaufkommen von:

$$0.16 \text{ MB/s} = 576 \text{ MB/h} = 13.824 \text{ GB/day} = 13,824 \text{ TB/day}$$

Das entspricht dem Füllen von ca. 14 Festplatten mit einer Speicherkapazität von jeweils 1 Terabyte. Um das Datenaufkommen zu reduzieren, ist es z.B. sinnvoll, bei Sensorsignalen, deren Werte sich selten ändern, nur den Zeitpunkt der Änderung aufzuzeichnen bzw. Signale mit einer zur Änderungsrate passenden Frequenz aufzuzeichnen. Ebenso sollten nur Signale aufgezeichnet werden,

# EMuDig4.0

die zur Beantwortung der Fragestellungen notwendig sind. Die Auswahl von Sensorsignalen hängt damit eng mit den Fragestellungen eines Unternehmens zusammen.

Aufgrund der Cloud-Eigenschaft, die Rechenleistung des Rechnersystems durch einfaches Hinzufügen weiterer CPUs zu vorhandenen virtuellen Maschinen (VM) zu erhöhen, sind keine speziellen Prozessoren notwendig und aktuelle Standardprozessoren können zum Einsatz kommen. Empfohlen wird die Verwendung von Prozessoren vom Typ x86-64, da dieser CPU-Typ am weitesten verbreitet ist und von den Betriebssystemen Linux und Windows unterstützt werden. Da Cloud-Managementsysteme einer virtuellen Maschine einzelne Kerne einer CPU zuweisen können, sind bei gleichen Kosten und gleicher Gesamtleistung CPUs mit vielen Kernen tendenziell vorzuziehen.

Bei Verwendung von Algorithmen des maschinellen Lernens im Rahmen von Datenanalysen kann auch der zusätzliche Einsatz von GPUs sinnvoll sein, da entsprechende Softwaretools wie z.B. Keras und Tensorflow zunehmend die Nutzung von GPUs unterstützen und damit deutlich schneller Rechnungen durchführen können.

Da es sich um bei dem aufzubauenden Rechnersystem um eine Cloud zur Datenanalyse handelt, sind datenintensive Rechnungen der Standardfall. Um diese Rechnungen nicht durch Auslagern auf Festplatte zu bremsen, sollte die Größe des Arbeitsspeichers großzügig bemessen sein.

Als Massenspeicher sind aktuelle SSD-Platten aufgrund ihrer Performance herkömmlichen Magnetfestplatten vorzuziehen. Für die oben angegebene Beispielkonfiguration ergibt sich ein tägliches Datenaufkommen von ca. 14 Festplatten mit einer Speicherkapazität von 1 Terabyte. Jeder VM ist zudem weiterer Massenspeicher zuzuordnen, der zum Speichern des Betriebssystems genutzt wird. Der Speicher ist entsprechend zu dimensionieren.

Aktuelle kabelgebundene Ethernet-Netzwerke unterstützen Datenübertragungsraten von 100 MBit/s bis 1 GBit/s. Wird die obige Beispielkonfiguration zugrunde gelegt, dann kann bei einer durchschnittlichen Auslastung des Netzwerkes die evtl. schon vorhandene Netzwerkinfrastruktur weiter genutzt werden. Insbesondere ist jedoch keine spezielle Netzwerktechnik nötig. Eine kabelgebundene Lösung ist zu empfehlen, um eine möglichst störungsfreie und stabile Datenübertragung sicherzustellen. Sollte dies nicht möglich sein (z.B. aufgrund der Position der Maschinen in der Produktionshalle), dann ist die Verwendung eines firmeninternen WLANs eine weitere Option.

## 6.1.1.1.3 Architektur

Beim Aufbau einer eigenen Cloud-Infrastruktur sind verschiedene Software-Werkzeuge notwendig:

- Ein Cloud-Verwaltungswerkzeug zur flexiblen und zuverlässigen Vereinigung von virtuellen Maschinen zu einem Cluster (Abschnitt 6.1.1.1.4),
- ein Ressourcenmanager zur Orchestrierung des Deployments von Anwendungen, deren Überwachung und Skalierung (Abschnitt 6.1.1.1.5),
- ein Job-Scheduler zum Planen der Verteilung von Services in der Cloud-Umgebung. Der Scheduler überwacht die Ausführung von Services und startet sie im Fehlerfall neu (Abschnitt 6.1.1.1.6),
- optional: ein Software-Werkzeug zur Containervirtualisierung, mit dem die wichtigsten Services zur Datenanalyse in Containern innerhalb der virtuellen Maschinen gekapselt werden können (Abschnitt 6.1.1.1.7).

# EMuDig4.0

Im Projekt EMuDig 4.0 wurde die Infrastruktur auf der Basis von Linux (Ubuntu Server) aufgebaut und betrieben. Letztendlich hängt die Wahl des Betriebssystems von den Anforderungen des Cloud-Verwaltungswerkzeugs ab.

Abbildung 6.1 zeigt die grundlegende Architektur der Cloud-Infrastruktur und die wesentlichen, darin laufenden Dienste.

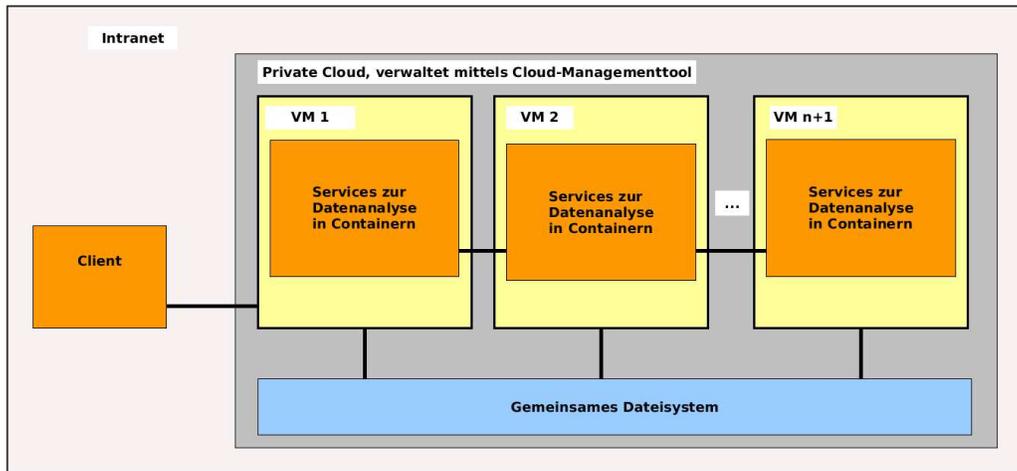


Abbildung 6.1: Grundlegende Architektur der Cloud-Infrastruktur.

## 6.1.1.1.4 Cloud-Verwaltungswerkzeuge

Ausgehend von der Virtualisierung von Servern und der zugehörigen Managementtools haben sich mehrere Cloud-Verwaltungswerkzeuge auf dem Markt etabliert, die im Rahmen des Projektes EMuDig 4.0 evaluiert wurden. Mit diesen Werkzeugen lässt sich u.a. die Größe die Anzahl der CPU-Kerne sowie die Größe des Arbeitsspeichers und des Massenspeichers von virtuellen Maschinen (VM) in der Cloud komfortabel konfigurieren. Sie unterscheiden sich in Punkten der Erweiterbarkeit, der Tiefe des Nutzermanagements und der Unterstützung von Cloud-Standards.

Die verwendete Cloud-Managementsoftware sollte auf dem Markt etabliert sein, sie sollte ausreichend Flexibilität für die Administration und für die Benutzer:innen bieten, um die vorhandenen Ressourcen in heterogene Clouds integrieren zu können, und sie sollte nicht auf spezieller Hardware aufbauen und mit vertretbarem Aufwand installierbar und betreibbar sein.

OpenStack [Contributors to Openstack 2020b], Cloudstack [Contributors to CloudStack 2020a], OpenNebula [OpenNebulaSystems2020a] oder Eucalyptus [Contributors to Eucalyptus 2020] sind etablierte Open-Source-Werkzeuge für IaaS-Lösungen. Der Einsatz von kommerziellen Cloud-Management-Lösungen wie z.B. VMware [VMware Inc. 2020] ist aber auch möglich.

OpenStack wurde von Rackspace und der NASA im Jahr 2010 gegründet und ist mittlerweile das größte und aktivste quelloffene Projekt im IaaS-Umfeld. Es wird als ein vollwertiges Cloud-Betriebssystem betrachtet, das große Mengen an Rechen-, Speicher- und Netzwerkressourcen effektiv verwalten kann. Die Verwaltung der Cloud-Ressourcen ist dabei über hochentwickelte Kommandozeilentools und ein benutzerfreundliches, grafisches Webinterface möglich. Mandantenfähigkeit und eine feingranulare Rechteverwaltung sichern den Betrieb auch mit mehreren parallelen Benutzer:innen ab. OpenStack ist in funktionale Module unterteilt, die jeweils ein eigenes Unterprojekt bilden und von

## EMuDig4.0

einer eigenen Entwicklergemeinde unterhalten werden. OpenStack unterstützt eine breite Palette Software-definierter Netzwerke (z.B. VLAN, VXLAN) basierend auf OpenvSwitch oder Linux Bridge. Neben den klassischen Blockspeichern kann auch Objektspeicher durch das Unterprojekt swift genutzt werden. Es unterstützt viele Hypervisoren zur Virtualisierung wie z.B. KVM, Xen und Vmware.

Als Cloud-Managementsoftware wird OpenStack empfohlen, insbesondere aufgrund seiner sehr guten Unterstützung von Virtualisierungslösungen in den Bereichen Hypervisor, Netzwerk und Speicher, der flexiblen Architektur, der Unterstützung von Komponenten des Hochleistungsrechnens und des Community-Supports.

Zur Erhöhung der IT-Sicherheit sind für jede virtuelle Maschine Sicherheitsregeln über das Cloud-Managementtool anzulegen, um nur die Ports zu öffnen, die für den Betrieb der Datenanalyse-Infrastruktur notwendig sind.

### 6.1.1.1.5 Software-Werkzeuge für das Ressourcenmanagement

Mit Hilfe eines Cloud-Managementsystems können flexibel und zuverlässig Ressourcen in einem Cloud-Cluster vereinigt werden. Die Aufgabe eines Ressourcenmanagers in einem Cloud-Cluster ist die Orchestrierung des Deployments von Anwendungen, deren Überwachung und Skalierung.

Neben spezifischen Ressourcenmanagern wie z.B. Apache YARN für das Hadoop-Ökosystem haben sich mehrere generische quelloffene Ressourcenmanager etabliert. Mesos (The Apache Software Foundation 2020a) der Apache Foundation, Kubernetes (The Kubernetes Authors 2020) von Google bzw. der Cloud Native Computing Foundation sind die verbreitetsten Vertreter. Mesos und Kubernetes bauen beide auf einer Manager-Worker-Architektur auf. Kubernetes beschränkt sich auf die Ausführung von so genannten Pods (einem Set von Docker-Containern), Speichervolumen und Konfigurationen. Obwohl es erweiterbar ist und über eine Web-API für die Orchestrierung verfügt, liegt seine Stärke insbesondere in der Koordinierung der Services und der Plugin-Netzwerke. Mit cAdvisor bringt Kubernetes eine Monitoring-Komponente mit, die CPU-, Speicher- und Netzwerknutzung der Container überwacht und aufzeichnet.

Mit Hilfe von Docker Swarm (Docker Docs. 2020) kann ein Cluster / „Swarm“ von Docker-Maschinen auf den virtuellen Maschinen der Cloud erstellt und verwaltet werden. Über die Docker-Maschinen können dann Dienste ausgeführt werden. Es ist keine weitere Orchestrierungssoftware zum Erzeugen und Verwalten des Swarms notwendig. Docker Swarm kann also entweder selbst als Ressourcenmanager verwendet werden oder Ressourcen über Apache Mesos reservieren und darin Docker-Container ausführen, die mit einem Host-übergeifenden privaten Softwarenetzwerk (SDN) verbunden sind. Swarm-Manager (Knoten im Swarm-Cluster) überwachen den Zustand des Swarm-Clusters und starten Services im Cluster – bei Bedarf reskaliert – automatisch neu. Standardmäßig erfolgt die Kommunikation zwischen den Knoten im Swarm-Cluster verschlüsselt. Lastbalancierung und Service-Discovery (Automatische Erkennung von Diensten in einem Rechnernetz) sind zwei weitere Features, die von Docker Swarm unterstützt werden.

Ohne weitere Software lässt sich Docker Swarm nur über die Kommandozeile verwenden. Mit Hilfe von Swarmpit [TopmonksSwarmpit2020a] können Swarm-Cluster auch komfortabel über eine Webschnittstelle verwaltet werden. Swarmpit läuft im Mehrbenutzermodus und verfügt über eine Monitoringkomponente, mit der grundlegende Hardwaregrößen wie CPU-, Speicher- und Festplattennutzung überwacht werden können. Über die Webschnittstelle von Swarmpit kann in privaten und öffentlichen Docker-Repositories nach Container-Images gesucht werden.

# EMuDig4.0

Durch die Möglichkeit, Docker Swarm gleichzeitig als Ressourcenmanager und als Job-Scheduler verwenden zu können, und durch die gute Integration von Docker-Containern in Docker Swarm, ist Docker Swarm die bevorzugte Wahl bei einer, auf Docker-Containern basierenden Cloud.

## 6.1.1.1.6 Job-Scheduler

Job-Scheduler planen die Verteilung von Services in der Cloud-Umgebung, überwachen deren Ausführung und starten sie im Fehlerfall neu, um ihre Hochverfügbarkeit sicherzustellen. Scheduler verfügen über eine geeignete Sicherheitsstruktur, um die Cloud auch über das Internet steuerbar zu machen. Ressourcenanfragen werden vom Job-Scheduler an den Ressourcenmanager weitergeleitet.

Für Docker-basierte Services eignen sich insbesondere die Job-Scheduler Marathon, Chronos, Aurora und Docker Swarm.

Marathon und Chronos besitzen ein Web-Frontend und bieten eine reichhaltige REST-API, um den Zustand auslesen und den jeweiligen Scheduler steuern zu können. Marathon ist speziell als Initsystem konzipiert und steuert Langzeitjobs und deren Abhängigkeiten. Chronos ist auf wiederkehrende, zeitgesteuerte Jobs spezialisiert. Im Aurora-Scheduler lassen sich komplexe Abhängigkeiten zwischen Jobs am besten abbilden und in großen Umgebungen ausführen. Auch integratives Scheduling über mehrere Cluster hinweg ist mittels Aurora möglich. Die Ausführungseinheit Thermos bietet dafür aber eine relative schwergewichtige Sandbox, die für Docker-Anwendungen nicht notwendig ist.

Im Rahmen von EMuDig 4.0 wurde Docker Swarm sowohl als Ressourcenmanager als auch als Job-Manager gewählt (siehe Abschnitt 6.1.1.1.5).

## 6.1.1.1.7 Containervirtualisierung

Durch die Kapselung wichtiger Services zur Datenanalyse in Containern innerhalb der virtuellen Maschinen wird das Service-Deployment vereinfacht: Verschiedene Versionen (z.B. zum Testen) oder Updates der Services können dann einfach, nur durch Bereitstellung eines neuen Container-Images eingespielt werden

Werden Container selbst entwickelt, dann ist sicherzustellen, dass die Container ohne Administrator-Rechte betrieben werden, um zu verhindern, dass Fremde Zugriff auf das Host-System erhalten [Marius Dorlöchter 2017]. Wenn verwendete Container nicht selbst erstellt wurden, dann sollte für die genutzten Images immer auch das Rezept zum Bauen des Images verfügbar sein. Damit lässt sich überprüfen, ob die Software im Container Schadcode enthält.

Docker [Docker Docs 2020] ist eine populäre Software zur Containervirtualisierung und wurde in EMuDig 4.0 ausgewählt. Docker erlaubt seit Version 1.11, mit Hilfe von "Control groups" die Ressourcennutzung von Containern zu beschränken. Diese Möglichkeit sollte genutzt werden, damit sich Docker-Container nicht untereinander und das Host-System beeinflussen [Marius Dorlöchter 2017]. Weitere Hinweise zum (IT-sicheren) Umgang mit Docker-Containern finden sich bei [Marius Dorlöchter 2017].

Eine benutzerfreundliche Bereitstellung von Docker-Containern innerhalb der Cloud kann durch die Nutzung einer zentralen Docker-Registry erreicht werden. Eine private Docker-Registry ist zu empfehlen, da diese volle Kontrolle über den Speicherort der Container-Images bietet und darüber hinaus keine Verbindung mit dem Internet zu öffentlichen Registries aufgebaut werden muss. (Marius

# EMuDig4.0

Dorlöchter 2017) Als private Registry kann z.B. die frei verfügbare Docker-Registry in einem eigenen Docker-Container innerhalb der Cloud betrieben werden.

Um die IT-Sicherheit zu erhöhen, sollte beim Bau eigener Container-Images bevorzugt das Betriebssystem Alpine Linux als Basis genutzt werden. Alpine Linux ist eine Linux-Distribution, die sich insbesondere durch eine erhöhte IT-Sicherheit und Ressourceneffizienz von anderen Linux-Distributionen unterscheidet. Alpine-basierte Container-Images sind üblicherweise deutlich kleiner als Ubuntu-basierte Container-Images. Zum Vergleich: Das Basis-Image von Alpine Linux 3.11 ist ca. 5 MB groß, das von Ubuntu 18.04 ca. 90 MB.

## 6.1.1.1.8 Netzwerk

Das Netzwerk zwischen den virtuellen Maschinen der Cloud kann in OpenStack als VXLAN auf der Basis von OpenvSwitch ausgeführt werden.

Bei der Wahl von Docker Swarm als Container-Orchestrierungstool erfolgt die Netzwerkkommunikation zwischen Containern mit Hilfe eines Overlay-Netzwerkes: Die Container haben keinen direkten Zugriff auf das Netzwerkinterface des Hosts, sondern benötigen Portweiterleitungen, um ihre Services nach außen anzubieten. Ihr privates Netzwerk erstreckt sich mittels Docker Swarm über alle beteiligten Knoten, sodass auch mit allen Docker-Containern in diesem Netzwerk direkt und sicher kommuniziert werden kann, auch wenn sie auf verschiedenen Hosts laufen.

## 6.1.1.1.9 Dateisysteme

Da auf das Betriebssystem einer VM nur von der VM selbst zugegriffen werden muss, ist ein lokales Dateisystem (z.B. bei Linux-basierten Betriebssystemen ext4) zum Speichern des Betriebssystems ausreichend.

Für eine einheitliche Sicht und Zugriff auf die Prozessdaten sollte ein verteiltes Dateisystem bereitgestellt werden. Die Nutzung eines verteilten Dateisystems hat gegenüber eines lokalem Dateisystems den Vorteil, dass die vorhandenen Speicherressourcen transparent allen VMs in der Cloud unter einem einheitlichen Dateizugriff zugänglich gemacht werden können. Zwei etablierte verteilte Cluster-Dateisysteme sind HDFS [The Apache Software Foundation 2020c] und GlusterFS [Red Hat Inc. 2020]. Aufgrund der guten Integration von HDFS in das Hadoop-Ökosystem wurde HDFS in EMuDig 4.0 ausgewählt.

Die Cloud-Managementsoftware OpenStack bietet die Möglichkeit, ceph-Block-Storage den VMs zuzuordnen. Bei Verwendung von externem ceph-Blockspeicher (ceph ist eine verteilte Speicherlösung) in OpenStack kann der Speicher in Blocks aufgeteilt werden und den VMs zugeordnet werden. Bei Bedarf kann daher über die Cloud-Managementsoftware weiterer ceph-Blockspeicher eingebunden werden.

## 6.1.1.1.10 Hochverfügbarkeit

Hochverfügbarkeit der virtuellen Maschinen in der Cloud kann z.B. durch Erstellen von Snapshots der VMs oder durch Verwendung von Softwaretools zur Überwachung der VMs erreicht werden.

# EMuDig4.0

Eine manuelle Lösung ist die Erstellung von Snapshots der VMs in regelmäßigen Abständen. Bei Ausfall einer virtuellen Maschine kann der zugehörige Snapshot auf eine neue Maschine überspielt werden und diese VM dann gestartet werden. Die Lösung ist relativ einfach, erfordert aber im Regelfall den Eingriff durch den Cloud-Administrator.

Eine automatisierte Lösung ist die Spiegelung von virtuellen Maschinen und Verwendung von Monitoringtools. Für jede VM können mehrere gespiegelte VMs mit lokalen IP-Adressen erstellt werden. Mit Hilfe einer Software (z.B. keepalived unter Ubuntu) kann überprüft werden, ob die VMs ansprechbar sind. Falls eine VM ausfällt, dann kann die entsprechende gespiegelte VM zum Einsatz kommen. Das Cloud-Managementsystem (hier: OpenStack) sorgt dafür, dass die VM für Benutzer:innen stets unter der gleichen IP-Adresse erreichbar ist, d.h., Benutzer:innen merken einen evtl. Ausfall der VM nicht.

Hochverfügbarkeit der in den Containern laufenden Dienste lässt sich über Docker Swarm umsetzen [Docker Inc. 2020].

## 6.1.1.1.11 Zugang

Der administrative Zugang auf die Cloud-Umgebung sollte ausschließlich über das sichere Transportprotokoll SSH aus dem firmeninternen Netzwerk über einen Kontrollknoten ermöglicht werden. Wird ein SOCKS-Proxy verwendet, dann ist trotzdem ein komfortabler Zugriff auf Webschnittstellen der Cloud über den Webbrowser möglich.

## 6.1.1.1.12 Umsetzung in EMuDig 4.0

Zusammengefasst wurde in EMuDig 4.0 die in Abbildung 6.2 dargestellte Cloud-Architektur umgesetzt:

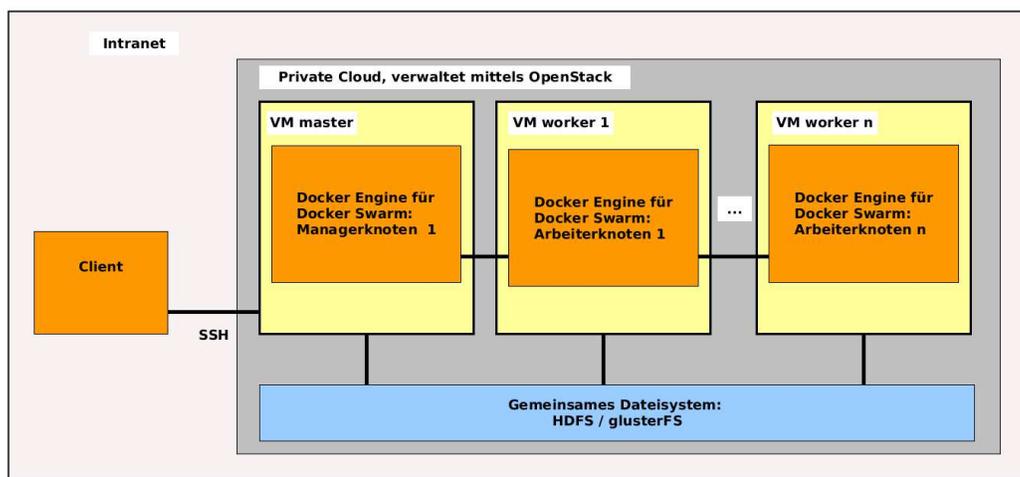


Abbildung 6.2: In EMuDig 4.0 umgesetzte Architektur der Cloud-Infrastruktur.

# EMuDig4.0

## 6.2 Datenanalyse-Infrastruktur

Die Cloud-Infrastruktur liefert eine Umgebung, in der Datenanalyse-Dienste aufgebaut und skalierbar betrieben werden können. In EMuDig 4.0 wurden diese Dienste auf der Basis von Open-Source-Software-Komponenten konzipiert und umgesetzt. Diese Komponenten wurden für den spezifischen Anwendungsfall der Analyse von Sensordaten aus der Massivumformung zunächst getestet und evaluiert sowie erweitert, angepasst und teilweise neu entwickelt.

Damit Datenanalysen von Benutzern in der Cloud-Umgebung schnell beantwortet werden können, sollten Ergebnisse der Anfragen schnell zu berechnen sein und bestenfalls schon vorberechnet vorliegen. Die Lambda-Architektur ist ein Konzept zum Verarbeiten großer Datenmengen mittels Batch- und Streamverarbeitung. [Marz and Warren 2015] Die Streamverarbeitungsschicht ist für Analysen in Echtzeit vorgesehen und berücksichtigt, abhängig von der Fragestellung, nur Daten der letzten Minuten oder Stunden. Die Batchverarbeitungsschicht ist für Langzeitanalysen konzipiert und berücksichtigt den gesamten vorliegenden Datensatz. Die Lambda-Architektur wurde entwickelt, um langandauernde, datenintensive Berechnungen von der Benutzerschnittstelle zu entkoppeln, indem diese vorab berechnet werden.

Da in der aufzubauenden Cloud nur Langzeitanalysen durchgeführt werden sollen, ist die Streamverarbeitungsschicht nicht relevant und der Aufbau der Software-Infrastruktur lässt sich damit auf die drei folgenden Schichten reduzieren (siehe Abbildung 6.3):

- Ingestion-Layer für die Verarbeitung ankommender Datenströme
- Batch-Layer für die Batchverarbeitung der Prozessdaten und Speicherung des Masterdatensatzes
- Serving-Layer für die Beantwortung von Anfragen und zur Speicherung und Visualisierung vorberechneter Ergebnisse

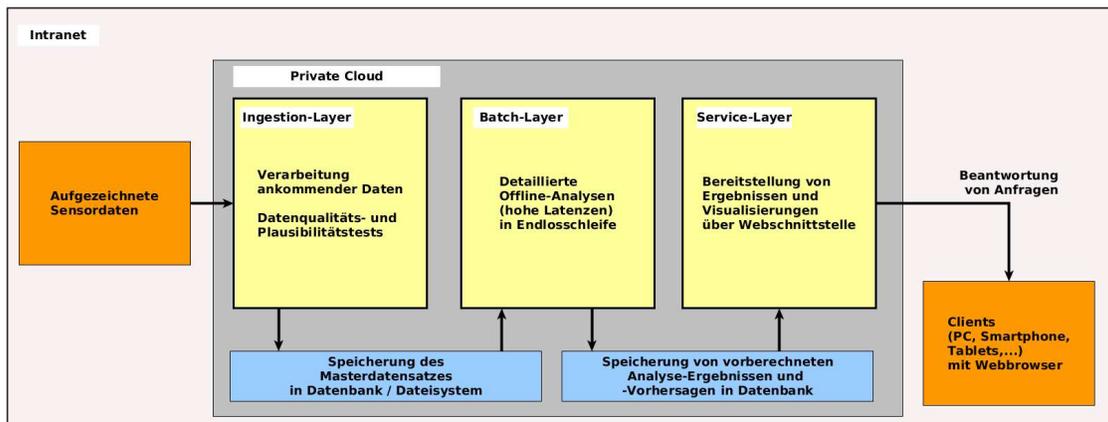


Abbildung 6.3: Datenfluss in Datenanalyse-Infrastruktur

### 6.2.1 Ingestion-Layer

Es wird vorausgesetzt, dass alle aufgezeichneten Sensordaten als Datenstrom in Form von Zeitstempel-Wert-Paaren an die Cloud gesendet. Das Senden der Daten an die REST-API des Apache Kafka-Clusters sollte dabei aus Sicherheitsgründen ausschließlich über das sichere Protokoll HTTPS erfolgen.

# EMuDig4.0

Um sicherzustellen, dass nur strukturell korrekte und inhaltliche plausible Daten in der Cloud gespeichert werden, sind die ankommenden Daten vor der Speicherung strukturellen und inhaltlichen Plausibilitätstests zu unterziehen (siehe Abschnitt 1.1.1), z.B. mit Apache Avro oder Apache Thrift. Beides sind Frameworks zur Datenserialisierung auf Basis zuvor festgelegter Datenschemas. Geprüfte Daten sind dann in einer Datenbank, dem sogenannten „Masterdatensatz“, zu speichern.

## 6.2.2 Batch-Layer

Die Hauptaufgabe des Batch-Layers ist, Ergebnisse zu Analyse-Anfragen auf Basis des gesamten und immer größer werdenden Masterdatensatzes in regelmäßigen Abständen vorzuberechnen. Damit alle VMs Zugriff auf den Masterdatensatz haben, ist der Datensatz im gemeinsamen Dateisystem zu speichern.

Für die Vorberechnung und Analyse der Prozessdaten wurde in EMuDig 4.0 die Statistik-Software R verwendet. R ist ein etabliertes, frei verfügbares Softwaretool für statistische Berechnungen. [R Foundation 2020] Die schon in der Basisversion von R reichhaltige Funktionalität lässt sich durch Verwendung von R-Paketen noch erweitern. Aufgrund der großen Nutzergemeinde gibt es dabei Erweiterungspakete für sehr viele Gebiete der klassischen und modernen Statistik. Aus Sicherheitsgründen sollten R-Pakete nur aus dem offiziellen CRAN-Repository installiert werden. [Contributors to R 2020] Für die Erstellung interaktiver Webapplikationen können die R-Pakete shiny und shinydashboard genutzt werden. Neben R ist Python [Python Software Foundation 2020] eine etablierte Programmiersprache im Data-Analytics-Umfeld. Sehr viele Softwaretools wie z.B. Apache Kafka und Apache Spark besitzen Schnittstellen sowohl zu Python als auch R. Aufgrund der einfachen Verknüpfung von Datenanalysemethoden in modernen, interaktiven Webapplikationen mit Hilfe von shiny und der Ausrichtung auf das Gebiet der Datenanalyse war R in EMuDig 4.0 die bevorzugte Wahl.

Große Datenmengen lassen sich mittels verschiedener Big-Data-Frameworks wie Hadoop, Apache Flink, oder Apache Spark verarbeiten. Alle Frameworks wurden entworfen, um auf verteilten Rechnersystemen zu arbeiten, und sind dementsprechend Cluster-fähig. Flink und Spark erweitern den Map/Reduce-Ansatz von Hadoop durch angepasste Operatoren zur Datenmanipulation und stellen Frameworkfunktionalitäten für eine Batch-basierende Datenverarbeitung zur Verfügung. Durch die Datenhaltung im Speicher kann Apache Spark i.d.R. deutlich schneller Daten verarbeiten als Hadoop. Apache Spark ist ein quelloffenes Toplevel-Projekt der Apache Software Foundation. [The Apache Software Foundation 2020d] Es besitzt Schnittstellen zu den Programmiersprachen Java, Scala, Python und R, in denen schnell Anwendungen erstellt werden können. [The Apache Software Foundation 2020b]

Spark unterstützt dabei sowohl das Einlesen als auch das Speichern der verarbeiteten Daten in einer Vielzahl von Datenbanken und Datenformaten. [Karau et al. 2015] Aufgrund der guten Integration von Spark in die Statistik-Software R über R-Pakete wie sparklyr und sparkR und der umfangreichen Machine-Learning-Bibliothek wurde Apache Spark für die Umsetzung in EMuDig 4.0 ausgewählt. Apache Spark sollte dabei aus Sicherheitsgründen so konfiguriert werden, dass der Datenaustausch im Spark-Cluster über das sichere HTTPS-Protokoll erfolgt.

## 6.2.3 Serving-Layer

Der Serving-Layer hat die Aufgabe, die im Batch-Layer vorberechneten Ergebnisse im gemeinsamen Dateisystem in einer Datenbank zu speichern und Analyseergebnisse zu Anfragen schnell

# EMuDig4.0

Anwender:innen bereitzustellen. Im Gegensatz zum Masterdatensatz darf die Datenbank zum Speichern der vorberechneten Ergebnisse durchaus Redundanzen enthalten, um Anfragen von Benutzer:innen möglichst schnell beantworten zu können. Geeignete Systeme zur Speicherung sind Zeitreihen-basierte Datenbankmanagementsysteme (DBMS) wie z.B. influxDB [InfluxData Inc. 2020]. influxDB ist unter einer Open-Source-Lizenz verfügbar und Cluster-fähig.

Damit Anwender:innen selbst Analysen anstoßen können, sollte Datenanalysemethoden über eine oder mehrere Applikationen zugänglich gemacht werden. Anfragen an die Datenanalyse-Infrastruktur und Visualisierungen der Analyseergebnisse sollten dabei bevorzugt über einen Webbrowser möglich sein, um unabhängig vom verwendeten Endgerät zugreifen zu können. Des Weiteren sollten die wichtigsten Prozessgrößen interaktiv über eine Webschnittstelle einstellbar sein.

Dabei ist zu beachten, dass eine interaktive Webschnittstelle zur Datenanalyse nur zu einem bestimmten Teil generisch sein kann (z.B. für die Erstellung beschreibender Statistiken), da Anfragen und Ergebnisse stets von den Prozessdaten und den Fragestellungen des Unternehmens abhängen. Zum deutlich größeren Teil ist eine Webapplikation daher an die spezifischen Fragestellungen und Prozessdaten des Unternehmens anzupassen. Es empfiehlt sich daher, die Webapplikationen möglichst modular zu erstellen. Werden die gewünschten Analysemethoden in jeweils eigenen Visualisierungs- bzw. Ergebnisboxen gekapselt, die über ein Auswahlménú hinzugefügt oder entfernt werden können, dann können die Webapplikationen den zugrunde liegenden Fragestellungen entsprechend zusammengestellt und flexibel angepasst werden.

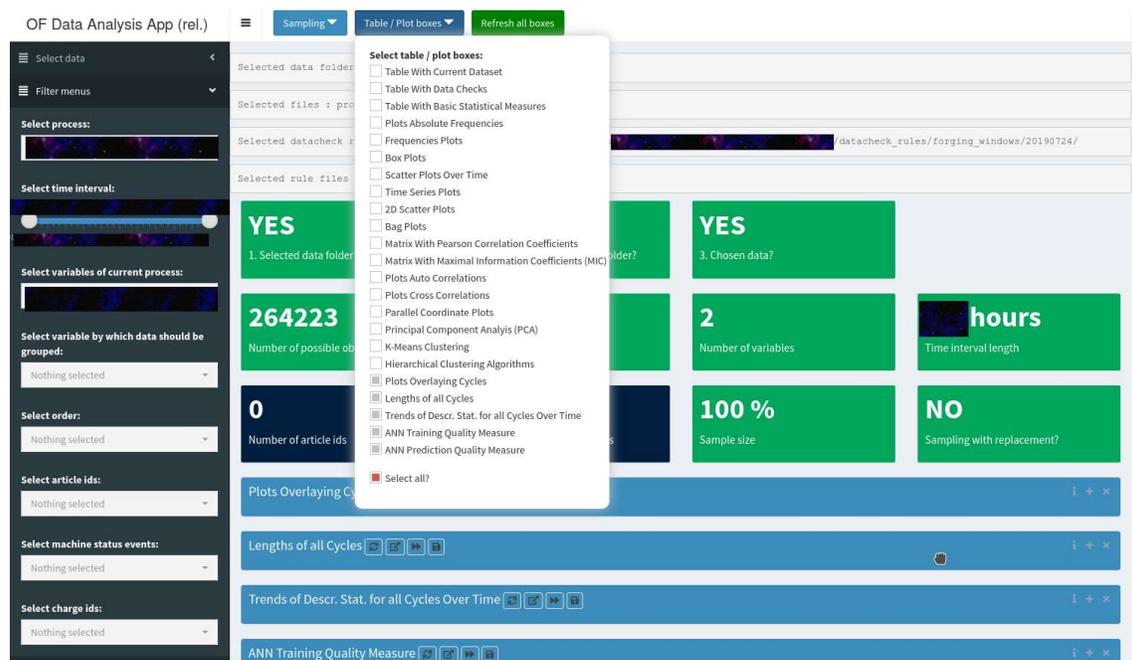


Abbildung 6.4: Beispiel-Webapplikation zur Datenanalyse (Sensible Daten wurden geschwärzt). Visualisierungs- und Ergebnisboxen können über ein Auswahlménú einfach entfernt oder hinzugefügt werden.

Neben den Web-Applikationen zur Langzeitdatenanalyse ist es sinnvoll, eine weitere interaktive Web-App zur Datenvorverarbeitung zu erstellen. Diese soll Benutzer:innen helfen, die zeitintensive und aufwändige Vorverarbeitung von Produktionsdaten automatisiert und komfortabel über eine grafische Webschnittstelle anhand kleinerer, dateibasierter Datensätze direkt durchzuführen.

# EMuDig4.0

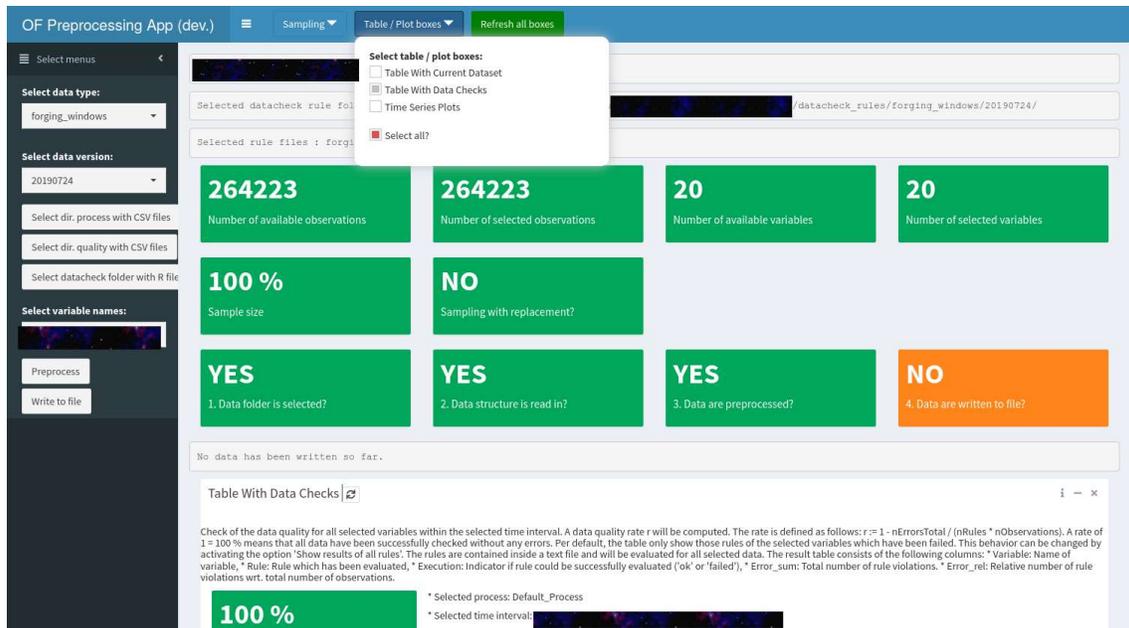


Abbildung 6.5: Webapplikation zur Datenvorverarbeitung (Sensible Daten wurden geschwärtzt).

Für die Erstellung interaktiver Webapplikationen in EMuDig 4.0 wurden u.a. die R-Pakete `shiny` und `shinydashboard` genutzt.

## 6.2.4 Umsetzung in EMuDig 4.0

Zusammengefasst wurde in EMuDig 4.0 die in Abbildung 6.6 angegebene Datenanalyse-Infrastruktur umgesetzt.

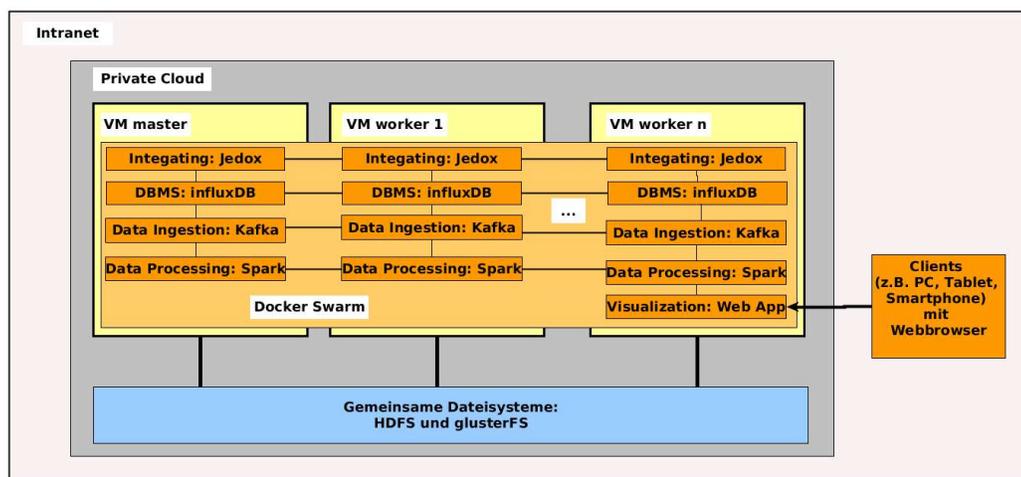


Abbildung 6.6: In EMuDig 4.0 umgesetzte Datenanalyse-Infrastruktur als Cloud.

## 6.3 Datenvorverarbeitung

Hauptziele bei der Analyse von historischen Prozessdaten in der Massivumformungsbranche sind die Untersuchung von Abhängigkeiten zwischen Prozessgrößen und das Finden von Mustern und Abweichungen in den Daten. Um Analysen durchführen zu können, müssen die Daten zuvor in geeigneter, maschinenlesbarer Form vorliegen.

Das folgende Modul des Leitfadens gibt einen Überblick und eine Beschreibung über die wichtigsten Schritte bei der Datenvorverarbeitung von Prozessdaten in der Massivumformungsbranche. Dazu gehören:

1. Ausarbeitung und Umsetzung eines IT-Sicherheitskonzeptes (Abschnitt 6.3.2),
2. Datenintegration, um aufgezeichnete Daten in digitaler Form aus verschiedenen Datenquellen (u. a. aus analogen und digitalen Quellen) zusammenzuführen (Abschnitt 6.3.3),
3. Auswahl von relevanten Daten im Rahmen der Datenbereitstellung (Abschnitt 6.3.4),
4. Erstellung einer Datendokumentation, in der die Datenstrukturen beschrieben werden, um Daten inhaltlich verstehen zu können und um die Qualität der Daten überprüfen zu können (Abschnitt 6.3.5),
5. Überprüfung der Datenqualität anhand der bereitgestellten Datendokumentation (Abschnitt 1.1.1)
6. Bereinigung der Daten von fehlerhaften, doppelten Werten, u.a. auf Basis der durchgeführten Datenvalidierung (Abschnitt 6.3.7),
7. Zusammenführung der Daten für eine einheitliche Sicht auf die Daten (Abschnitt 6.3.8).

Die Datenvorverarbeitung ist im Wesentlichen ein sequentieller Prozess, d.h., ein Teilschritt erfordert die Bearbeitung aller vorherigen Schritte. Mehrere Iterationen über vorangegangene Schritte sind nicht ungewöhnlich.

Als grober Richtwert gilt, dass ca. 70 bis 80 Prozent der Zeit bei einer Datenanalyse für die Datenvorverarbeitung aufgewendet werden, d.h., die Datenvorverarbeitung ist der kritischste und zeitaufwändigste Punkt. Daher sollte ausreichend Zeit dafür eingeplant werden.

### 6.3.1 Ausgangssituation

Allgemein sind bei einer Datenanalyse stets zwei Parteien beteiligt: Auf der einen Seite der Datengeber, d.h. die Partei, die die Daten bereitstellt. Dies wird im Regelfall das Unternehmen aus der Massivumformungsbranche sein; auf der anderen Seite der Datenanalytiker, d.h. die Partei, die die bereitgestellten Daten analysiert.

Typischerweise treffen zu Beginn eine oder mehrere der folgenden Aussagen zu:

- Daten werden bei Unternehmen aufgezeichnet und gesammelt.
- Es gibt viele Produktionsschritte.
- Nicht zu jedem Produktionsschritt werden Daten aufgezeichnet.

# EMuDig4.0

- Viele Prozesse erzeugen große Datenmengen.
- Daten einzelner Prozesse werden individuell gesammelt.
- Nicht alle Daten liegen zwingend in digitaler Form vor.
- Verschiedene Prozesse nutzen oft verschiedene / mehrere Datenbanken, Datenformate und Programme.
- Prozesse oder Datensätze sind oft nur über Zeitstempel und nicht über andere Identifikatoren verknüpft.
- Produktionsdaten haben einen erhöhten Schutzbedarf.

## 6.3.2 IT-Sicherheit

Oft haben die aufgezeichneten Daten eines Unternehmens einen erhöhten Schutzbedarf, d.h., die Weitergabe der Daten und die anschließende Weiterverarbeitung sind an bestimmte Bedingungen geknüpft. Damit ein Unternehmen Daten zur Analyse bereitstellen kann, ist im Vorfeld ein IT-Sicherheitskonzept zu entwerfen und umzusetzen. Je nach Höhe des Schutzbedarfs der Daten, dieser wird durch den Datengeber festgelegt, kann die Erstellung und Umsetzung eines Sicherheitskonzeptes, z.B. nach Standards wie ISO-Normen oder BSI, umfangreich sein und einen längeren Zeitraum in Anspruch nehmen.

### 6.3.2.1 Anforderungsanalyse

Mit den Projektpartnern ist dazu eine Anforderungsanalyse durchzuführen, in der mindestens folgende Fragen geklärt werden müssen:

- Werden personenbezogene Daten verarbeitet?
- Welchen Schutzbedarf haben die Daten? In welche Schutzklassen lassen sich die Daten einteilen?
- Zugriffskontrolle / Rechten-Rollen-Management: Wer darf auf Daten zugreifen und wie?
- Wie / Auf welchem Wege sollen Daten ausgetauscht werden (Verschlüsselung der Daten)?
- Auf welchen Rechnersystemen (IT-Infrastrukturen) dürfen die Daten gespeichert und verarbeitet werden?
- Dürfen Backups der Daten erstellt werden? Auf welchen Systemen?
- Dürfen Daten archiviert werden, während des Projekts und danach?

Die Datenschutz- und Informationssicherheitsbeauftragten aller beteiligten Projektpartner sollten frühzeitig in diesen Prozess einbezogen werden.

# EMuDig4.0

## 6.3.2.2 IT-Infrastrukturskizze

Basierend auf der Anforderungsanalyse ist für den Datengeber eine IT-Infrastrukturskizze zu erstellen, in der festgelegt wird, auf welcher IT-Infrastruktur zu analysierende Daten vom Datenanalysten verarbeitet werden dürfen. Diese Skizze enthält daher u.a. die IT-Infrastruktur und den Workflow zur IT-sicheren Verarbeitung der bereitgestellten Daten. Unter anderem muss die Infrastrukturskizze Angaben zu folgenden Punkten enthalten:

- Server, auf denen Daten gespeichert / analysiert werden
- verwendete Clients, die auf Server Zugriff haben
- räumliche Lage der Server / Serverstandort (z.B. Inland, Ausland)
- Netzwerkverbindungen
- Zugriffskontrollen und Zugriffsmöglichkeiten auf Server und Clients.
- Betriebssysteme auf Servern
- Anwendungen, die auf den Clients und Servern laufen

## 6.3.2.3 Modellierungsart

Nach Erstellung der IT-Infrastrukturskizze müssen sich die Projektpartner auf eine Modellierungsart einigen, nach der die IT-Sicherheit der geplanten IT-Infrastruktur modelliert werden soll. Zwei etablierte Modellierungsarten sind die ISO 27001-Norm und der IT-Grundschutz des Bundesamtes für Sicherheit in der Informationstechnik (BSI). Eine Zertifizierung nach ISO ist mit zusätzlichen Kosten verbunden. Im Folgenden wird beispielhaft eine Modellierung nach BSI-IT-Grundschutz beschrieben. In diesem Fall sind u.a. Bausteine zu verwenden, die den Serverstandort, Server zur Verarbeitung der Daten, das Dateisystem zum Speichern der Daten und das Netzwerk betreffen. Die Wahl der Bausteine ist dabei abhängig von der IT-Infrastruktur.

## 6.3.2.4 Ist-Soll-Vergleich

Die Bausteine bilden die Basis für einen Ist-Soll-Vergleich bezüglich der erforderlichen organisatorischen und technischen Maßnahmen. Dieser Vergleich ist von denjenigen Datenanalysten durchzuführen, die die bereitgestellten Daten üblicherweise auf eigenen IT-Systemen verarbeiten möchten. Gegebenenfalls sind noch nicht umgesetzte Maßnahmen umzusetzen. Abbildung 6.7 zeigt beispielhaft, in welcher Form der Ist-Zustand der Umsetzung der erforderlichen Maßnahmen dokumentiert werden kann.

Maßnahme	Titel	Umsetzung	Erläuterung
M 1.28	Änderung voreingestellter Passwörter	Ja	—
.	.	.	.
.	.	.	.

Abbildung 6.7: Auszug aus einer Maßnahmentabelle

Die Abbildung des Ist-Soll-Vergleichs in einem Software-Tool wie verinice [SerNet 2020] erleichtert dabei die Verwaltung und gemeinsame Bearbeitung innerhalb der Organisation. verinice erlaubt z.B. das Verwalten der Bausteine über ein komfortables Webinterface, verfügt über eine Berichterstattungsfunktion und ist für den Mehrbenutzerbetrieb geeignet.

Abhängig von der IT-Infrastrukturskizze kann sich der Vergleich auf mehrere Abteilungen (z.B. Netzwerk, Services) der Organisation erstrecken und benötigt daher Abstimmung und Zuarbeiten der Abteilungen. Um die Zeit für den Ist-Soll-Vergleich und die Umsetzung von Maßnahmen möglichst kurz zu halten, sollten Verantwortlichkeiten innerhalb der Organisation bereits im Vorfeld eines Projekts geklärt werden. Im besten Fall sind die erforderlichen Maßnahmen der wichtigsten Bausteine bereits im Rahmen eines früheren Projekts geprüft oder bereits umgesetzt worden und können wiederverwendet werden. Es wird empfohlen, ausreichend Zeit für Verhandlungen zwischen den beteiligten Parteien und Umsetzungen der Maßnahmen einzuplanen. Zwei Monate sind durchaus realistisch.

Der durchgeführte Ist-Soll-Vergleich ist dann als Nachweis zu den organisatorischen und technischen Maßnahmen in Form einer Übersicht vorzulegen, welche Maßnahmen nach BSI-IT-Grundschutz umgesetzt wurden. Auf Basis dieser Übersicht kann der Datengeber, im Regelfall das Unternehmen, dann entscheiden, ob Daten bereitgestellt werden können.

Sollten nicht alle geforderten Maßnahmen bereits vollständig umgesetzt worden sein, so können die Datengeber durchaus Auflagen machen, die mit Zeitfenstern versehen sind.

### 6.3.3 Datenintegration

Bei der Datenintegration ist das Ziel, aufgezeichnete Daten in digitaler Form aus verschiedenen Datenquellen (analoge, digitale Quellen, z.B. PDF-Dateien, Papier, Grafiken) zusammenzuführen. Dieser Schritt ist bei der Daten-bereitstellenden Partei durchzuführen. Nach der Datenintegration liegen die Daten i.d.R. noch in vielen verschiedenen Strukturen vor.

Erlauben die Produktionsbedingungen, Bauteile mit einer eindeutigen Seriennummer über die gesamte Produktionskette zu kennzeichnen, dann sind zu jeder aufgezeichneten Beobachtung eines Sensorsignals mindestens die folgenden Informationen zu speichern:

- Zeitstempel
- Sensorwert
- Name des Prozesses
- Seriennummer / Bauteilnummer (wenn Kennzeichnung auf Bauteilebene möglich)

# EMuDig4.0

- Chargen-ID: Zur Zuordnung des Bauteils zu einer Charge
- Auftrags-ID: Zur Zuordnung des Bauteils zu einem Auftrag
- Maschinenstatus-ID: Maschinenstatus des aktuellen Prozesses

## 6.3.4 Datenbereitstellung

Sofern alle geforderten Maßnahmen bzgl. der IT-Sicherheit geprüft und gegebenenfalls umgesetzt worden, können Daten für andere Projektpartner bereitgestellt werden.

### 6.3.4.1 Anforderungsanalyse

Folgende Fragen / Punkte sind dabei von den Projektpartnern zu klären:

- Welche Dateiformate besitzen die bereitgestellten Dateien?
- Wie ist die zeitliche Auflösung der aufgezeichneten Sensordaten?
- Bereitstellung von einzelnen / zusammengeführten Datensätzen
- Dokumentation der Datenstrukturen (Festlegung geeigneter Konventionen)
- Welche Daten (Live- und/oder historische Daten) sollen bereitgestellt werden?
- Wie hoch ist die zu erwartende Datenmenge?
- Sind firmeneigene / proprietäre Software-Tools zum Lesen des verwendeten Dateiformats notwendig?

### 6.3.4.2 Auswahl relevanter Daten

Mutmaßlich relevante Daten sind in Zusammenarbeit mit den Unternehmen auszuwählen. Die Daten sollten dabei zu den Fragestellungen der Unternehmen passen. Wesentlich ist bei der Datenauswahl die Unterstützung der Ingenieure, die mit den zu analysierenden Prozessen vertraut sind. Sollte sich zu einem späteren Zeitpunkt herausstellen, dass die ausgewählten Daten nicht ausreichende Informationen zur Beantwortung der Fragestellungen enthalten, dann ist die Auswahl zu wiederholen. Gegebenenfalls müssen dann weitere Sensoren installiert und Daten aufgezeichnet werden.

### 6.3.4.3 Format bereitgestellter Daten

Für ein erstes Kennenlernen der Daten sollten die Daten strukturiert in einem gängigen Dateiformat bereitgestellt werden. CSV oder SQLite sind dafür zwei weit verbreitete Dateiformate, wobei das CSV-Format u.a. aus den folgenden Gründen zu bevorzugen ist: In der Regel liegen die Daten schon im CSV-

# EMuDig4.0

Format vor oder können leicht in dieses Dateiformat überführt werden. Des Weiteren wird das CSV-Format von etablierten Software-Werkzeugen zur Datenanalyse wie R und Python unterstützt.

## 6.3.4.4 Austausch bereitgestellter Daten

Für einen schnellen und IT-sicheren Austausch der Daten sollte eine Infrastruktur mit einem Rollen- und Rechtesystem vorhanden sein, über die auf die Daten zugegriffen werden kann. Damit Unternehmen volle Kontrolle über die bereitgestellten Daten behalten, sollten die Daten über firmeneigene Server bereitgestellt werden. So kann z.B. ein Server mit Rollen- und Rechtesystem, zugänglich über HTTPS, im Firmenintranet eingerichtet werden, zu dem die Datenanalysten Zugang erhalten. Es wird empfohlen, die bereitgestellten Daten auf dem Server zusätzlich abzusichern. Eine Möglichkeit ist, die Daten verschlüsselt abzulegen und den Schlüssel dann über ein anderes Kommunikationsmittel bereitzustellen.

## 6.3.5 Datendokumentation

Damit Datenanalysten die bereitgestellten Daten inhaltlich verstehen können, ist eine Beschreibung der Datenstrukturen notwendig. Eine Dokumentation der Daten ist auch wichtig, um die Qualität der Daten zu überprüfen (siehe Abschnitt 6.3.5). Das Ziel ist daher eine möglichst genaue Beschreibung der Daten, insbesondere ihrer Struktur. In Zusammenarbeit der Projektpartner ist dazu eine Datendokumentation zu erstellen, die Antworten auf die folgenden Fragen liefert:

- Welche Merkmale (inhaltlich, physikalisch) sind in den Daten enthalten?
- Wie werden die enthaltenen Merkmale erhoben?
- Wie häufig werden die Merkmale erhoben?
- Welche Einheiten haben die Merkmale?
- Was ist der Wertebereich der Merkmale?
- Wie werden atypische Werte, z. B. fehlende Werte gekennzeichnet?
- Welche Zeichen werden als Dezimal- und Tausender-Trennzeichen verwendet?
- Wie werden fehlende Werte der Sensoren gekennzeichnet?
- In welchem Wertebereich können die Sensoren aufzeichnen? Was passiert mit Werten, die über den zulässigen Wertebereich der Sensoren hinausgehen?

In der Datendokumentation sollte auch festgehalten werden, wie die Daten aussehen:

- Sind die Daten „zensiert“ (Nicht alle Werte eines Merkmals sind bekannt) oder „winsorisiert“ (Werte über / unter eines bestimmten Grenze werden auf eine gegebene Ober- / Untergrenze transformiert)?
- Wird alles kleiner als einem gegebenen Schwellwert / Threshold auf null transformiert?
- Folgen die Daten einer gestutzten Verteilung?

# EMuDig4.0

## 6.3.6 Datenvalidierung

Folgeschritte hängen bei der Datenanalyse von Ergebnissen aus vorherigen Schritten ab. Die Qualität von Analyseergebnissen ist maßgeblich bestimmt durch die Datenqualität. Oft treten Inkonsistenzen bei der Datenspeicherung in unterschiedlichen Programmen und in unterschiedlichen Dateiformaten auf. Durch Plausibilitätschecks können Fehler frühzeitig erkannt und gegebenenfalls korrigiert werden.

Das Ziel der Datenvalidierung ist daher eine Überprüfung der Datenqualität anhand der bereitgestellten Datendokumentation und Sicherstellen einer hohen Datenqualität. Insbesondere sollen im Rahmen der Datenvalidierung fehlerhafte, unvollständige oder doppelt erfasste Datensätze entdeckt werden.

Basischecks sind Überprüfungen hinsichtlich der Vollständigkeit der Werte und des in der Datendokumentation angegebenen Datentypformats. Des Weiteren sind auch inhaltliche Prüfungen sinnvoll, wenn entsprechende Bedingungen in der Datendokumentation beschrieben wurden (z.B. Kraftkurvenverlauf, Stößelweg wird mit fortschreitendem Pressvorgang kleiner). Für jedes in der Datendokumentation beschriebene Merkmal sind dabei Regeln zu definieren, die als Ergebnis „wahr“ oder „falsch“ liefern. Jede Regel ist dann auf alle Beobachtungen des Merkmals anzuwenden und die Anzahl der Beobachtungen mit Ergebnis „falsch“ zu zählen. Auf diese Weise kann eine Datenqualitätsquote ermittelt werden.

## 6.3.7 Datenbereinigung

Die im vorherigen Abschnitt beschriebene Datenvalidierung gibt einen Überblick über die Datenqualität der bereitgestellten Daten. Wurden im Rahmen der Datenvalidierung fehlerhafte Werte gefunden, so sind die Daten von diesen Werten zu bereinigen. Dabei ist zu entscheiden, wie mit fehlenden Werten, doppelten Werten und verschiedenen Datenformaten umgegangen wird.

Des Weiteren sollten für eine möglichst große Kompatibilität mit Software-Tools bei der weiteren Verarbeitung alle Zeichenketten (insbesondere in Kopfzeilen von Tabellen) in ein ASCII-Format umgewandelt werden. Beginnt z.B. eine Zeichenkette in der Kopfzeile mit einer Ziffer, dann sollte der Ziffer ein Buchstabe vorangestellt werden. Die Umwandlung der Zeichenketten kann unter Linux z.B. mit Kommandozeilentools wie dem Stream-Editor sed durchgeführt werden.

## 6.3.8 Datenaggregation

Die Datenaggregation hat zum Ziel, Daten für eine einheitliche Sicht zusammenzuführen. Dafür sollten alle Daten in Tabellen bereitgestellt werden, die über ein eindeutiges Merkmal verknüpft sind. Insbesondere sollte vermieden werden, dass verschiedene Merkmale nur über Zeitstempel zusammengeführt werden können, da im Normalfall Daten von verschiedenen Maschinen gesammelt werden, die die Daten mit Zeitstempeln einer intern eingebauten Uhr versehen. Die Zeit ist damit in der Regel nicht synchron, was zu einem Versatz bei der Zusammenführung der Daten führen wird.

Unterschiedliche Datenquellen müssen z.B. dann zusammengeführt werden, wenn Sensoren diese unabhängig voneinander liefern. Es ist daher gemeinsam mit dem Datengeber zu klären, wie unterschiedliche Datenquellen zusammengeführt werden sollen. Möglich wäre z.B., den Durchschnitt zu bilden, das Minimum oder Maximum zu wählen, Daten hinzuzufügen oder wegzulassen.

# EMuDig4.0

Generell sollte die Datenaggregation durch den Datengeber erfolgen, da er das umfassendere inhaltliche Verständnis der Daten besitzt. Insbesondere bei der Datenaggregation kommt dem Aspekt des Informationsverlustes eine hohe Bedeutung zu (z.B. Durchschnittsbildung, Beschränkung auf Extremwerte). Deshalb sollten derartige Transformationen nur nach Rücksprache mit Prozessexperten und Datenanalysten durchgeführt werden.

## 6.4 Datenanalysemethoden

Ein Ziel im Forschungsprojekt EMuDig 4.0 ist es, Prozesse in der Massivumformungsbranche besser zu verstehen, um Prozesse zielgerecht optimieren zu können. Eine Möglichkeit, das Prozessverständnis zu erhöhen, sind detaillierte Analysen von Prozessdaten, insbesondere die Suche nach Abhängigkeiten, Mustern, Abweichungen in den Daten.

Das folgende Modul des Leitfadens gibt einen Überblick und eine Beschreibung der wichtigsten Schritte bei der Datenanalyse von Prozessdaten in der Massivumformungsbranche. Zu den wichtigsten Schritten gehören:

1. Eine systematische Vorbereitung / Vorverarbeitung der Daten (siehe Leitfaden zu Modul Datenvorverarbeitung und Abschnitt 5.3.0)
2. Die Entwicklung inhaltlicher Fragestellungen, die durch die Datenanalyse beantwortet werden sollen: Was sind die konkreten inhaltlichen Analyseziele? (Abschnitt 6.4.2)
3. Eine deskriptiv-statistische Analyse der Daten für ein besseres Verständnis der Daten (Abschnitt 6.4.3)
4. Die Auswahl der für die Fragestellungen relevanten Daten und Aufbereitung der Daten für die vorgesehenen Analysemethoden (Abschnitt 6.4.4)
5. Verwendung geeigneter Analysemethoden passend zu den vorliegenden Fragestellungen (Abschnitt 5.3.0). Dazu gehört auch eine Visualisierung der Ergebnisse zum besseren Verständnis der Ergebnisse und als Teil einer graphisch-basierten Datenanalyse.
6. Die (prototypische) Umsetzung der Analysemethoden in Softwarewerkzeugen (Abschnitt 6.4.6)
7. Die Aufbereitung und Bereitstellung der Analyseergebnisse für die datengebende Partei sowie Auswertung und Interpretation der Analyseergebnisse zusammen mit den Datengebern (Abschnitt 6.4.7)

Die Datenanalyse ist wie die Datenvorverarbeitung im Wesentlichen ein sequentieller Prozess, d.h., ein Teilschritt erfordert die Bearbeitung aller vorherigen Schritte. Dieser Prozess ist aber auch iterativ, weil der Rücksprung in vorangegangene Teilschritte erforderlich sein kann.

## 6.4.1 Datenvorverarbeitung

Die Datenvorverarbeitung (siehe Leitfaden zu Modul Datenvorverarbeitung) ist eine wichtige Voraussetzung für eine fundierte Datenanalyse. Unter anderem hängt die Qualität der Analyseergebnisse stark von der Qualität (z.B. der Genauigkeit) der vorverarbeiteten / bereinigten Daten ab:

- Ist es möglich, mit Daten, die über Sensoren an den Maschinen gesammelt werden, Rückschlüsse auf die Prozesse zu ziehen?
- Ist die Qualität der Daten hoch genug, um präzise Voraussagen treffen zu können, d.h., wie hoch ist die Relevanz der Datenrelevanz für geeignete Vorhersagen?
- Genügen die Momentaufnahmen oder Zeitfenster der Datenaufzeichnungen, um ausreichend exakte Aussagen treffen zu können?
- Existieren Korrelationen im Datenraum und zwischen Daten und Prozessen, die Rückschlüsse erlauben?
- Können mit Datenanalysemethoden und -Werkzeugen diese Korrelationen abgebildet oder gelernt werden?
- 

## 6.4.2 Ausarbeitung von Fragestellungen

Bevor bereitgestellte Daten untersucht werden können, sind in Zusammenarbeit zwischen Datengebern und Datenanalysten konkrete Fragestellungen zu erarbeiten, bezüglich derer die Daten analysiert werden sollen.

Im Rahmen von EMuDig 4.0 konnten die nachstehenden Fragestellungen identifiziert und herausgearbeitet werden. Diese können gleichzeitig als Beispiele für typische Fragestellungen in der Massivumformungsbranche dienen:

1. Um die Qualität von Bauteilen zu beurteilen, werden üblicherweise für ausgewählte Bauteile verschiedene Gütemaße bestimmt.
2. Lassen sich Bauteile auf Basis von Gütemaßen in Gut- und Schlechteile einteilen?
3. Da die Bestimmung von Gütemaßen eines Bauteils in der Regel sehr zeitaufwändig und teuer ist und damit nicht für jedes Bauteil möglich ist, ist es wünschenswert, die Einteilung in Gut- und Schlechteile basierend auf Prozessgrößen durchzuführen, die einfacher aufzuzeichnen sind. Diese Prozessgrößen können z.B. die Position oder die Geschwindigkeit einer Presse sein. Für die gewünschte Prognose von Gütemaßen durch Prozessgrößen ist zunächst zu untersuchen, wie die Gütemaße von ausgewählten Prozessgrößen abhängen.
4. Wenn Zusammenhänge zwischen Prozessgrößen und Gütemaßen bestimmt werden konnten, dann kann untersucht werden, wie gut die Gütemaße auf Basis der Prozessgrößen vorhersagt werden können.
5. Die Qualität der eingesetzten Werkzeuge zum Bearbeiten von Bauteilen beeinflusst maßgeblich den Umformungsprozess und damit die Qualität des hergestellten Bauteils. Durch die andauernde Belastung unterliegen die Werkzeuge einem fortschreitenden Verschleiß, der zu einer immer schlechteren Qualität der gefertigten Bauteile führt. Üblicherweise müssen pro Fertigungsauftrag mehrere Werkzeuge aufgrund Verschleißes eingesetzt werden. Es ist daher wünschenswert, den Grad des Werkzeugverschleißes vor

einem Maschinenausfall / Werkzeugbruch erkennen zu können, um rechtzeitig Werkzeuge zu wechseln. Sind in ausgewählten Prozessgrößen Trends erkennbar, die auf einen fortschreitenden Werkzeugverschleiß hindeuten?

6. Basierend auf der Bestimmung von Trendindikatoren für einen fortschreitenden Werkzeugverschleiß ist zu untersuchen, wie zuverlässig Vorhersagen darüber getroffen können, ob es Werkzeugwechsel innerhalb eines vorab definierten Zeitfensters geben wird.

### 6.4.3 Statistische Beschreibung der Daten

Um die vorliegenden Daten besser verstehen zu können, sollten diese zu Beginn einer deskriptiven Analyse unterzogen werden. Die Daten können damit in ihrer Struktur näher beschrieben werden, u.a. hinsichtlich ihrer Wertebereiche und Ausprägungen. Verschiedene Sichtweisen auf die Daten lassen sich z.B. über die Berechnung statistischer Kenngrößen und die Erstellung von Tabellen und Diagrammen realisieren. Abhängig von der gewünschten Sicht auf die Daten, z.B. mit / ohne Berücksichtigung der Zeitkomponente (Längs- / Querschnitt), sind die Daten zuvor geeignet neu zu strukturieren und zu filtern. Empirische Verteilungen der vorliegenden Merkmale können z.B. durch Histogramme oder Kerndichtschätzungen visualisiert werden.

Durch die Bestimmung verschiedener Lagemaße wie des Minimums und des Maximums lässt sich der Wertebereich bestimmen, durch die Bestimmung von Quartilen (bzw. Quantilen) können Aussagen über die Verteilung eines Merkmals gemacht werden. So brauchen z.B. Merkmale, die über den gesamten Zeitraum konstant sind, oft nicht weiter berücksichtigt werden. Genauere Aussagen über die Verteilung eines Merkmals lassen sich durch die Berechnung von Streuungsmaßen in Verbindung mit den zuvor berechneten Lagemaßen machen. Speziell zur Identifikation und Darstellung von Extremwerten oder Ausreißern sind Boxplots geeignet. Mit diesen Ansätzen können erste Hypothesen aufgestellt werden, beispielsweise, ob die Daten bekannten Verteilungsmodellen folgen.

Bei der Darstellung von Variablen als Streudiagramm, aber auch bei Zeitreihenplots, besteht die Gefahr des Overplottings, insbesondere bei vielen Beobachtungen, diskreten Wertebereichen und kleinen Wertebereichen. Overplotting tritt auf, wenn viele Beobachtungen so nah nebeneinander bzw. übereinander liegen, dass im Diagramm nicht mehr erkennbar ist, wie viele Punkte tatsächlich geplottet wurden. Dieser Effekt kann u.a. durch Hexagonplots umgangen werden oder durch Verrauschen der Beobachtungen mittels eines zufälligen Störfaktors („Jittering“).

Zeitreihenplots (z.B. Zeit vs. Chargen-Id, zeitlicher Verlauf der Auftrags-Ids) liefern oft bereits wertvolle Informationen zu Aufträgen, Chargen und Bauteilen:

- Ist die Anzahl der Bauteile pro Auftrag bei allen Aufträgen ähnlich?
- Ist die Anzahl der Bauteile pro Charge bei allen Chargen ähnlich?
- Dominieren einige Chargen?
- Wie viele Bauteile wurden pro Auftrag gefertigt?
- Werden Chargen zeitlich komplett am Stück abgearbeitet oder wechseln sich Chargen-Ids öfter nacheinander ab?
- Ist der zeitliche Verlauf der internen Seriennummer monoton steigend, werden Bauteile also entsprechend der internen Seriennummer nacheinander abgearbeitet?

Um Abhängigkeiten zwischen verschiedenen Merkmalen zu charakterisieren und relevante Merkmale auszuwählen, können bivariate Zusammenhangsmaße (z.B. der Bravais-Pearson-

# EMuDig4.0

Korrelationskoeffizient) genutzt werden. Der Wert des Abhängigkeitsmaßes kann ein Kriterium bei der Auswahl relevanter Merkmale sein.

Autokorrelationen und Kreuzkorrelationen dienen der Visualisierung und Quantifizierung von ein- und mehrdimensionalen Abhängigkeiten im Längs- und Querschnitt.

- Streudiagramme mit Prozessgröße vs. Gütemaß geben erste Aufschlüsse zu Stärke und Richtung der Abhängigkeit.
- Die Stärke und Richtung von linearer Abhängigkeit zwischen Prozessgrößen und Gütemaßen kann z.B. über den Bravais-Pearson-Korrelationskoeffizient charakterisiert werden.

Häufig lassen sich bereits Prozess-bedingt kausale Zusammenhänge zwischen Gütemaß und bestimmten Prozessgrößen identifizieren. Handelt es sich beispielsweise beim Gütemaß um die gemessene Dicke eines Bauteils und bei der Prozessgröße um die Position der Presse, dann wird die Bauteildicke umso kleiner sein, je weiter die Presse nach unten bewegt wird.

## 6.4.4 Datenselektion und Dimensionsreduktion

In der Regel sind die vorliegenden Daten hochdimensional, d.h., es wurden während der Prozesse viele verschiedene Merkmale aufgezeichnet. Für die Reduktion der Daten auf wenige relevante Merkmale gibt es verschiedene Ansätze, z.B.:

- Exklusion konstanter und nicht-relevanter Merkmale
- Methode der Parallel-Koordinaten, auch zur Visualisierung hochdimensionaler Abhängigkeiten
- Hauptkomponentenanalyse (PCA)
- Aggregation von Zeitreihen, indem bestimmte Eigenschaften mit wenigen Maßzahlen quantifiziert werden (z.B. Trendkomponente, Zykluskomponente)

Die Relevanz von Merkmalen für die vorliegenden Fragestellungen sollte in der Datendokumentation vom Datengeber vermerkt werden. Die Entscheidung über die Relevanz der Merkmale sollte dabei von entsprechenden Prozessexperten getroffen werden.

Eine Dimensionsreduktion mittels der Hauptkomponentenanalyse bedeutet, dass nicht alle Hauptkomponenten im weiteren Analyseprozess berücksichtigt werden, sondern nur die ersten K Hauptkomponenten. Die Anzahl K sollte so gewählt werden, dass

- die gewählten Hauptkomponenten möglichst viel der Variabilität erklären,
- das adjustierte Bestimmtheitsmaß möglichst groß ist

Idealerweise existiert ein Knick im Graph Hauptkomponente vs. Variabilität (sog. Ellbogenplot oder Screeplot). Dieser zeigt an, dass die Hauptkomponenten ab dem Knick deutlich weniger der Variabilität erklären als die Hauptkomponenten davor. Dann können alle Hauptkomponenten bis zum Knick berücksichtigt und die weiteren Hauptkomponenten außer Acht gelassen werden. Ein Nachteil bei der Hauptkomponentenanalyse ist, dass die ursprünglichen Merkmale i.d.R. obskuriert werden und die resultierenden „künstlichen“ Variablen (= Hauptkomponenten) schwerer zu interpretieren sind.

## 6.4.5 Analysemethoden

Um geeignete Analysemethoden für die ausgearbeiteten Fragestellungen auszuwählen, sind die inhaltlichen Fragestellungen zunächst in datenanalytische Fragestellungen zu überführen und dann mathematisch zu formulieren.

Im Folgenden werden die in Abschnitt 6.4.2 beschriebenen Fragestellungen näher untersucht. Dabei wird der Weg bei der Bearbeitung der Fragestellungen aufgezeigt und begründet, warum welche Analysemethoden zum Einsatz gekommen sind. Es ist dabei zu berücksichtigen, dass die durchgeführten Datenanalysen aufgrund des knappen zeitlichen Rahmens im Forschungsprojekt lediglich als Voruntersuchung zu verstehen sind.

Die verwendeten mathematischen Modelle sollten dabei dem Prinzip der Parametersparsamkeit folgen. Dieser Aspekt erleichtert die inhaltliche Interpretation der Ergebnisse und kann damit helfen, das Prozessverständnis zu verbessern.

Bei der Wahl passender Analysemethoden zur Bearbeitung der Fragestellungen ist auch immer der Stichprobenumfang zu berücksichtigen, um zu beurteilen, inwiefern die gewählte Methode belastbare Ergebnisse liefern kann.

### 6.4.5.1 Klassifikation von Bauteilen in Gut- / Schlechtteile

Die in EMuDig 4.0 identifizierte Fragestellung

„Können gefertigte Bauteile auf Basis mehrerer Gütemerkmale in Gut- und Schlechtteile eingeteilt werden?“

kann als eine Klassifikationsaufgabe ohne Zielwerte modelliert werden. Die Definition der Klassen erfolgt dabei basierend auf den vorliegenden Gütemaßen. Nach der Klassenbildung können dann neue Bauteile in die zuvor definierten Klassen eingeteilt werden.

Für Klassifikationsaufgaben ohne bekannte Zielwerte sind Clusterverfahren als Methoden des unüberwachten maschinellen Lernens eine geeignete Wahl. Cluster lassen sich mittels verschiedener Methoden wie z.B. hierarchischer oder dichtebasierter Clusterverfahren bestimmen. Ein Vorteil von hierarchischen Clusterverfahren ist die geringe Anzahl an Verfahrensparametern, ein Nachteil ist die hohe Laufzeit: Bei  $n$  Datenpunkten beträgt die Komplexitätsordnung  $O(n^3)$ .

Im zwei- und dreidimensionalen Fall können graphische Darstellungen die Klassifikation der Daten unterstützen.

1. Bagplots verallgemeinern das Quantilskonzept auf den mehrdimensionalen Fall und sind bei zwei Gütemerkmalen das 2-D-Analogon zu Boxplots. Dabei werden bei zwei Gütemerkmalen die Beobachtungen basierend auf einer sogenannten Datenpunkttiefe in verschiedene Bereiche eingeteilt. Bauteile innerhalb des „Fences“ wurden als Gutteile definiert, Bauteile außerhalb des „Fences“ als Schlechtteile.
2. Ein Gütemaß kann in Quantilsbereiche (z.B. in 5-Prozent-Quantile) eingeteilt werden und für das andere Gütemaß für jeden Quantilsbereich ein Boxplot erstellt werden. Über diese graphische Darstellung können Hypothesen aufgestellt werden, in welchem Quantilsbereich

## EMuDig4.0

die Verteilung ein verändertes Verhalten aufweist. Diese Quantilsbereiche können dann z.B. als Klasse für Schlechteile gewählt werden.

3. Bei der Darstellung der Gütemaße über Streudiagramme können durch visuelle Analysen ebenfalls Hypothesen aufgestellt werden, welche Bereiche im Streudiagramm auffällig sind und als Klasse für Schlechteile gewählt werden können.

### 6.4.5.2 Zusammenhänge zwischen Bauteilgütemaßen und Prozessgrößen

Da die Bestimmung von Gütemaßen eines Bauteils in der Regel sehr aufwändig und teuer ist und daher üblicherweise nicht für jedes Bauteil möglich ist, ist es wünschenswert, die Klassifizierung von Bauteilen basierend auf Prozessgrößen durchzuführen, die einfacher aufzuzeichnen sind. Dies können z.B. die Position oder die Geschwindigkeit der Presse sein. Es ist daher zu untersuchen, wie die interessierenden Gütemaße von anderen Prozessgrößen basierend auf der vorherigen Klassifizierung abhängen. Höherdimensionale Abhängigkeiten zwischen Merkmalen lassen sich über Parallel-Koordinaten-Plots (PCP) grafisch darstellen.

### 6.4.5.3 Vorhersage von Bauteil-Gütemaßen

Die Fragestellung

*„Wie gut können Bauteil-Gütemaße durch Prozessgrößen wie z.B. Druck und Position an verschiedenen Messpunkten vorhergesagt werden?“*

kann als eine Regressionsaufgabe modelliert werden. Zur Vorhersage von Bauteil-Gütemaßen durch Prozessgrößen wie z.B. Druck und Position können daher klassische Regressionsverfahren und künstliche neuronale Netze verwendet werden. Die gewählten Vorhersagemethoden können z.B. über die Bestimmung der maximalen relativen Abweichung der vorhergesagten Werte zu den gemessenen Werten des Bauteil-Gütemaßes evaluiert werden. Typische Gütekriterien für die Bestimmung der Verfahrensgüte sind:

- Maximale absolute /relative Abweichung
- Mittlere absolute / relative Abweichung

Bei Vorliegen vergleichsweise kleiner Stichproben (z.B. mit nur einigen wenigen Hundert Bauteilen) kann die Evaluierung über einen Resampling-Ansatz erfolgen. Beim Resampling wird der Datensatz durch wiederholtes zufälliges Ziehen mehrmals neu in Test-, Validierungs- und Trainingsdatensatz aufgeteilt. Ein üblicher Quotient für das Verhältnis zwischen Trainings- und Validierungsdatensatz und dem Testdatensatz ist 80:20.

# EMuDig4.0

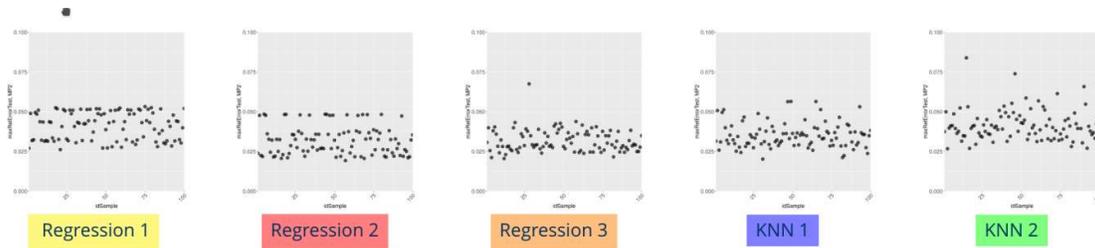


Abbildung 6.8: Vergleich der Gütemaße der verwendeten Verfahren für einen Messpunkt. Die X-Achse enthält die Resampling-Ide, die Y-Achse die zugehörigen durchschnittlichen, relativen Abweichungen.

Abbildung 6.8 zeigt die Abweichungen für drei Regressionsverfahren und zwei KNN-Verfahren bei einem Ansatz mit 100 Resamples. Hier liefert das Regressionsverfahren 3 (lineare Regression mit vorgeschalteter PCA und Berücksichtigung aller Kenngrößen) die im Durchschnitt kleinsten Abweichungen von zwei bis fünf Prozent.

Innerhalb eines Regressionsmodells wird der Zusammenhang zwischen dem Gütemaß (erklärte Variable, Regressand) und den Prozessgrößen (erklärende Variablen, Regressoren) in einer funktionalen Form, d.h. in Form mathematischer Gleichungen, dargestellt. Diese Methode liefert somit mathematische Gleichungen zur Vorhersage des Gütemaßes neuer Bauteile.

Unter Berücksichtigung des Prinzips der Parametersparsamkeit sollten zu Beginn bevorzugt Regressionsmodelle mit wenigen erklärenden Variablen verwendet werden. Dies bietet die Möglichkeit, vorhandenes Wissen über den betrachteten Prozess zu nutzen. Ebenso können weniger komplexe und inhaltlich gut interpretierbare Modelle dazu dienen, neue Einsichten in die bestehenden Prozesse zu gewinnen.

Speziell für die Vorhersageaufgabe sollten die verwendeten erklärenden Variablen auf Basis einer zuvor durchgeführten Zusammenhangsanalyse ausgewählt werden. Zusätzlich kann eine Hauptkomponentenanalyse zur Dimensionsreduktion durchgeführt werden, indem lediglich die wichtigsten Hauptkomponenten als erklärende Variablen Verwendung finden.

Nichtlineare Zusammenhänge lassen sich durch geeignete Variablentransformation auch innerhalb eines linearen Modells beschreiben.

Die Berücksichtigung weiterer Merkmale als erklärende Variablen verbessert oftmals die Vorhersage, führt aber zu einem höheren Aufwand bei der Messung. Zusätzlich wird die Interpretierbarkeit des Modells erschwert. In Abhängigkeit von der Zielstellung müssen diese Aspekte gegeneinander abgewogen werden.

Eine sehr komplexe Form von Regressionsmodellen stellen künstliche neuronale Netze (KNN) dar. Unter Berücksichtigung des Prinzips der Parametersparsamkeit wurden mit Feedforward Multi-Layer-Perceptrons einfache KNN gewählt. Die gewählten KNN bestehen aus einer Eingabeschicht (Input-Layer), einer Ausgabe-Schicht (Output-Layer) und ggfs. einer oder mehreren verborgene Schichten (Hidden-Layer). Als Knoten der Eingabeschicht können z.B. Kenngrößen basierend auf den Prozessgrößen gewählt werden, die zuvor zur Charakterisierung des Umformungsprozesses (=Zyklus') gewählt wurden. Als Knoten der Ausgabeschicht ist das vorherzusagende Gütemaß zu wählen.

Werden Hidden-Layers mit mehreren Knoten verwendet, dann kann die Anzahl der zu bestimmenden Parameter (=Kantengewichte) bei KNN schnell sehr groß werden. Um KNN sinnvoll einzusetzen, wird eine relativ große Anzahl von Daten benötigt. Falls die Datenmengen zu gering sind, besteht hier die

# EMuDig4.0

Gefahr der Überanpassung und die Verlässlichkeit der Modellaussagen über den vorliegenden Datensatz hinaus ist möglicherweise sehr gering. Typischerweise werden einige tausend Beobachtungen benötigt, um bei guter Generalisierbarkeit die unbekannt Parameter (sog. Gewichte) in einem KNN zu bestimmen.

## 6.4.5.4 Trendindikatoren für fortschreitenden Werkzeugverschleiß

Bei der Fragestellung

*„Sind Trends in den Daten erkennbar, die auf einen fortschreitenden Werkzeugverschleiß hinweisen?“*

handelt es sich um eine Trendanalyse. Eine Herausforderung bei der Bearbeitung der Fragestellung ist das Bestimmen eines „frischen Zustands“, d.h. eines Zustandes, bei dem alle Werkzeugstufen neu sind. In der Praxis kommt ein frischer Zustand der Werkzeuge so gut nie vor, da Werkzeugstufen unterschiedlich stark verschleifen und daher nicht alle Werkzeugstufen gleichzeitig gewechselt werden. Das bedeutet auch, dass eine bestimmte Wechselkombination ebenfalls nur selten vorkommt. Bei vier Werkzeugstufen gibt es z.B. insgesamt schon  $2^4 = 16$  verschiedene Wechselkombinationen.

Eine weitere Herausforderung ist die Berücksichtigung des Zeitpunkts der Werkzeugwechsel. So können verschiedene Maschinenbediener unterschiedliche Präferenzen bezüglich eines Werkzeugwechsels haben:

- Werkzeuge können gewechselt werden, bevor es wirklich nötig ist.
- Ebenso können Werkzeuge länger verwendet werden, als es der Qualität der Bauteile zuträglich ist.

Eine dritte wesentliche Herausforderung sind die aufgezeichneten Daten zum Werkzeugwechsel. Werkzeugwechsel werden in maschinenlesbarer Form aufgezeichnet, oft aber nicht, welche Werkzeugstufe gewechselt wurde. Der Datensatz sollte jedoch diese Information enthalten, um erforderliche Wechsel einzelner Werkzeugstufen belastbar vorhersagen zu können.

Zusätzliche Herausforderungen sind:

- Üblicherweise wird nicht der gleiche Artikel über einen längeren Zeitraum auf einer Presse hergestellt. Stattdessen wechselt sich die Produktion mehrerer Artikel in unregelmäßigen Abständen ab.
- Neben ungeplanten Werkzeugwechseln gibt es weitere Ereignisse (u.a. Kleinstörungen, geplante Wartungsfenster), bei denen die Maschine still steht.

Beim Abgleich der vorhergesagten Wechsel mit den tatsächlich durchgeführten Wechseln sind die oben angegebenen Aspekte zu berücksichtigen.

Zur Bestimmung von Trendindikatoren ist die Identifikation von Zyklen sinnvoll. Darauf basierend können charakterisierende Maßzahlen für jeden Zyklus gebildet und ausgewählt werden. Diese Maßzahlen können als erklärende Variablen (oder auch sog. Features) in Regressionsmodellen oder KNN verwendet werden.

# EMuDig4.0

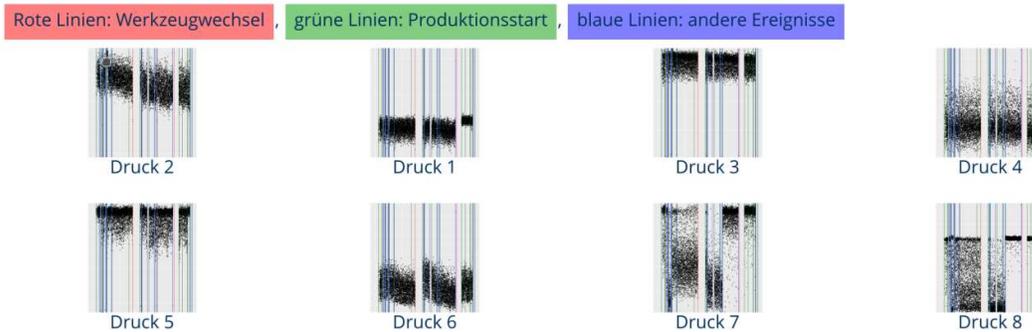


Abbildung 6.9: Trendindikatoren für Werkzeugwechsel

Abbildung 6.9 zeigt ausgewählte Features für Werkzeugwechsel im zeitlichen Verlauf. Die Features, bestimmt für jeden Pressvorgang, liefern in diesem Beispiel gute Trendindikatoren. So sind alle erfolgten Werkzeugwechsel durch die Trendindikatoren erkennbar: Beim ersten Werkzeugwechsel wurde die Werkzeugstufe 3 (entspricht Drücken 5 und 6) gewechselt, beim zweiten Werkzeugwechsel die Stufe 1 (entspricht Drücken 1 und 2) und die Stufe 3.

Neben der Bestimmung von geeigneten Features können die Zeitreihen der Merkmale in Trend- und Zykluskomponente zerlegt werden. Signifikanztests auf Vorliegen eines Trends können eingesetzt werden, um den Zufall von der Systematik abzugrenzen.

## 6.4.5.5 Vorhersage von Werkzeugwechseln

Bei der Entscheidungsfrage

„Gibt es in den nächsten N Minuten einen Werkzeugwechsel?“

handelt es sich um eine binäre Klassifikationsaufgabe. Das Ziel ist es, Ereignisse anhand der Prozessdaten vorhersagen. Vorhersagen können basierend auf den zuvor entwickelten Trendindikatoren z.B. durch Entscheidungsbäume oder KNNs durchgeführt werden. Da bei einem Werkzeugwechsel nicht alle Werkzeugstufen gleichzeitig gewechselt werden, ist bei Wahl von Entscheidungsbäumen als Vorhersagemethode für jede Werkzeugstufe ein eigener Entscheidungsbaum zu erstellen.

Enthalten die zugrunde liegenden Daten nicht Wechsel aller Werkzeugstufen, dann sollten in den Vorhersagen auch nur Wechsel der gewechselten Stufen berücksichtigt werden.

Mit Hilfe der sogenannten Wahrheitsmatrix („Confusion matrix“) können die Klassifikationsverfahren bewertet werden. Tabelle 6.1 zeigt die Darstellung der Wahrheitsmatrix für die zugrunde liegende Fragestellung.

Tabelle 6.1: Wahrheitsmatrix für Fragestellung

	Werkzeug gewechselt	Kein Werkzeug gewechselt
Wechsel vorhergesagt	True Positives (TP)	False Positives (FP)
Kein Wechsel vorhergesagt	False Negatives (FN)	True Negatives (TN)

# EMuDig4.0

Klassische Kriterien zur Bestimmung der Verfahrensgüte sind Sensitivity, Recall, Accuracy, und Balanced Accuracy.

Bei geringem Stichprobenumfang können Vorhersageergebnisse nur als Hypothesen aufgefasst werden. Für belastbare Vorhersagen wird ein Datensatz benötigt, der für alle Stufen eine ausreichende Zahl an Wechseln enthält.

## 6.4.6 Implementierung

Die ausgewählten und angepassten Methoden zur Datenvorverarbeitung und Datenanalyse können in verschiedenen Programmiersprachen implementiert werden. Die beiden am weitesten verbreiteten Sprachen im Bereich der Datenanalyse sind R und Python.

Es empfiehlt sich, die Software modular zu halten und Pakete zu erstellen, die u.a. Funktionalitäten zum Einlesen, Zusammenführen, Manipulieren, Analysieren und Visualisieren der Prozessdaten enthalten.

Eine Aufteilung der Software in mindestens zwei Pakete ist sinnvoll, um erstellte Funktionalitäten in einfacher Weise auch auf andere Anwendungsfälle anwenden zu können. So sollte ein Software-Paket allgemeine Funktionalitäten bereitstellen, ein weiteres Software-Paket Funktionalitäten, die speziell auf die Datenanalyse des Anwendungsfalls angepasst sind.

## 6.4.7 Aufbereitung und Evaluation der Ergebnisse

Analyseergebnisse sollten von Datenanalysten so aufbereitet werden, dass diese für die Datengeber verständlich und, wenn die Rahmenbedingungen es erlauben, abhängig von der Größe des Datensatzes, auch reproduzierbar sind.

In Zusammenarbeit zwischen Datenanalysten und Datengebern sind dabei die Anforderungen bzgl. der Ergebnisvisualisierung zu untersuchen. Anforderungen können sein:

- Interaktive Bedienung in einem Geräte-unabhängigen graphischen Nutzerinterface
- Einfache Verwendung komplexer Statistik- und Datenanalysemethoden, insbesondere „Kapselung“ des Expertenwissens vor dem Nutzer
- Integration in die vorhandene Datenanalyse-Infrastruktur

Die Bereitstellung der Analyseergebnisse kann auf verschiedene Arten erfolgen:

- Als statische Inhalte über Präsentationsfolien im PDF-Format
- Als dynamische Inhalte über interaktive Web-Applikationen (evtl. zusammen mit Rohdaten)
- Als interaktive Dashboards: Ein Dashboard ist eine graphische Anzeige, in der die wichtigsten Analyseergebnisse auf einer Webseite zusammengefasst werden. Mit Hilfe von Dashboards lassen sich z.B. auch Ergebnisse vom Anwender reproduzieren. Ebenso können Anwender ohne besondere Vorkenntnisse mit den Daten interagieren und eigene Untersuchungen durchführen.

Die Statistik-Software R bietet mehrere Erweiterungspakete für eine komfortable Darstellung von Ergebnissen im Webbrowser. Die R-Pakete flexdashboard und shinydashboard sind zwei der

# EMuDig4.0

bekanntesten Pakete. Werden R-Markdown-Dateien in Kombination mit shiny-Dashboards verwendet, dann erlaubt diese Kombination eine einfache interaktive Nutzung der Ergebnisse im Webbrowser.

Vorhandene Quellcodes aus R-Skripten und Funktionen können mit nur wenigen Modifikationen in R-Markdown-Dateien verwendet werden.

Das R-Paket shiny erlaubt die Erstellung moderner, interaktiver Webapplikationen in R. Abbildung 6.10 zeigt ausgewählte Möglichkeiten zur Datenanalyse in der entwickelten Webapplikation.

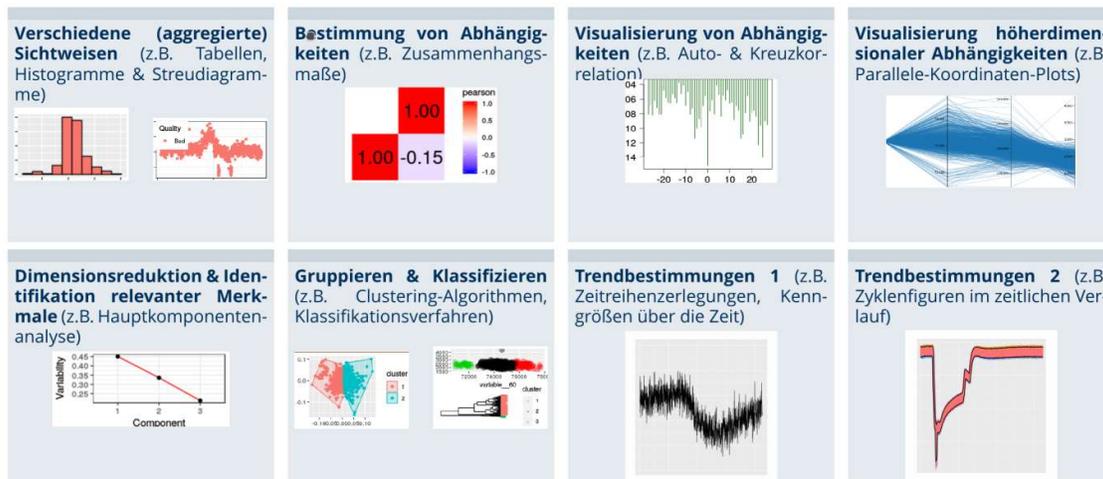


Abbildung 6.10: Analysemöglichkeiten in interaktiver Webapplikation, angepasst an Fragestellungen in EMuDig 4.0

Abschließend sind die Analyseergebnisse zu interpretieren. Dabei sind die Analyseergebnisse mit Fachexperten des Datengebers zu diskutieren. Fachexperten sind in der Regel die Fachingenieure der analysierten Prozesse.

# EMuDig4.0

## Literaturverzeichnis

Amazon Web Services. 2020. "AWS Amazon Elastic Compute Cloud (Amazon EC2)." <https://aws.amazon.com/de/ec2/>.

Apache Software Foundation. 2020. "Apache Zookeeper." <https://zookeeper.apache.org>.

ASSn. 2017. "Kappa Architecture: Pros and Cons." <http://www.datascienceassn.org/sites/default/files/Kappa%20Architecture%20%20Pros%20and%20Cons.pdf>.

Bill Ledingham. 2016. "Docker und Co. - wie sicher sind Ihre Container?" <https://www.computerwoche.de/a/docker-und-co-wie-sicher-sind-ihre-container,3312579>.

Bundesamt für Sicherheit in der Informationstechnik. 2020. "IT-Grundschutz." [https://www.bsi.bund.de/DE/Themen/ITGrundschutz/itgrundschutz\\_node.html](https://www.bsi.bund.de/DE/Themen/ITGrundschutz/itgrundschutz_node.html).

Contributors to Chronos. 2020. "Chronos A fault tolerant job scheduler for Mesos." <https://mesosphere.github.io/marathon/>.

Contributors to Eucalyptus. 2020. "Eucalyptus Cloud Platform." <https://github.com/eucalyptus/eucalyptus/>.

Contributors to CloudStack. 2020a. "Apache CloudStack: Open Source Cloud Computing." <http://cloudstack.apache.org/>.

———. 2020b. "Open source software for creating private and public clouds." <https://www.openstack.org/>.

Contributors to R. 2020. "The Comprehensive R Archive Network (CRAN)." <https://cran.r-project.org/>.

Docker Docs. 2020. "Docker overview." <https://docs.docker.com/engine/docker-overview/>.

Docker Docs. 2020. "Swarm mode overview." <https://docs.docker.com/engine/swarm/>.

Docker Inc. 2020. "Docker Documentation: Set up high availability." <https://docs.docker.com/ee/ucp/admin/configure/join-nodes/>.

Hitachi Vantara. 2020. "Pentaho: Data Integration and Analytics Platform." <https://www.hitachivantara.com/en-us/products/data-management-analytics/pentaho-platform.html>.

InfluxData Inc. 2020. "InfluxDB: Purpose-Built Open Source Time Series Database." <https://www.influxdata.com/>.

Karau, Holden, Andy Konwinski, Patrick Wendell, and Matei Zaharia. 2015. Learning Spark: Lightning-Fast Big Data Analytics. 1st ed. O'Reilly Media, Inc.

Kreps, Jay. 2014. "Questioning the Lambda Architecture." <https://www.oreilly.com/radar/questioning-the-lambda-architecture/>.

Marius Dorlöchter. 2017. "Docker Security Internals: Sind meine Container sicher und wenn ja warum?" <https://blog.ordix.de/technologien/sind-meine-container-sicher-und-wenn-ja-warum>.

Marz, Nathan, and James Warren. 2015. Big Data: Principles and Best Practices of Scalable Real-Time Data Systems. Manning. <https://www.amazon.de/Big-Data-Principles-practices-scalable/dp/1617290343>.

# EMuDig4.0

Mesosphere, Inc. 2020. "Marathon: A container orchestration platform for Mesos and DC/OS." <https://mesosphere.github.io/marathon/>.

Michael Hausenblas. 2015. "Hyperscaling Docker Swarm with Mesos [Mesosphere HackWeek]." <https://d2iq.com/blog/hyperscaling-docker-swarm-with-mesos-mesosphere-hackweek>.

Microsoft. 2020. "Cloud Computing Services." <https://azure.microsoft.com/de-de/>.

OpenNebula Systems. 2020. "OpenNebula - The open source Cloud Management Platform developed for the Enterprise." <https://opennebula.io/>.

Python Software Foundation. 2020. "Welcome to Python.org." <https://www.python.org/>.

Red Hat Inc. 2020. "Gluster: Storage for your Cloud." <https://www.gluster.org/>.

R Foundation. 2020. "The R Project for Statistical Computing." <https://www.r-project.org>.

Shruti Gupta. 2020. "rkafka: Using Apache 'Kafka' Messaging Queue Through 'R'." <https://cran.r-project.org/package=rkafka>.

The Apache Software Foundation. 2020a. "Apache Mesos." <https://mesos.apache.org/>.

———. 2020b. "Apache Spark – Unified analytics Engine for Big Data." <https://spark.apache.org/>.

———. 2020c. "HDFS Architecture Guide." [https://hadoop.apache.org/docs/r1.2.1/hdfs\\_design.html](https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html).

———. 2020d. "Spark becomes top-level Apache project." <https://spark.apache.org/news/spark-becomes-tlp.html>.

The Kubernetes Authors. 2020. "Production-Grade Container Orchestration – Kubernetes." <https://kubernetes.io/>.

Tim Mackey. 2018. "Container und Security: IT-Sicherheit gewährleisten!" <https://www.informatik-aktuell.de/entwicklung/methoden/container-und-security-it-sicherheit-gewaehrleisten.html>.

Topmonks. 2020. <https://swarmpit.io/>.

SerNet. 2020. "verinice: Das Open Source ISMS" <https://verinice.com/>.

VMware Inc. 2020. <https://www.vmware.com/>.